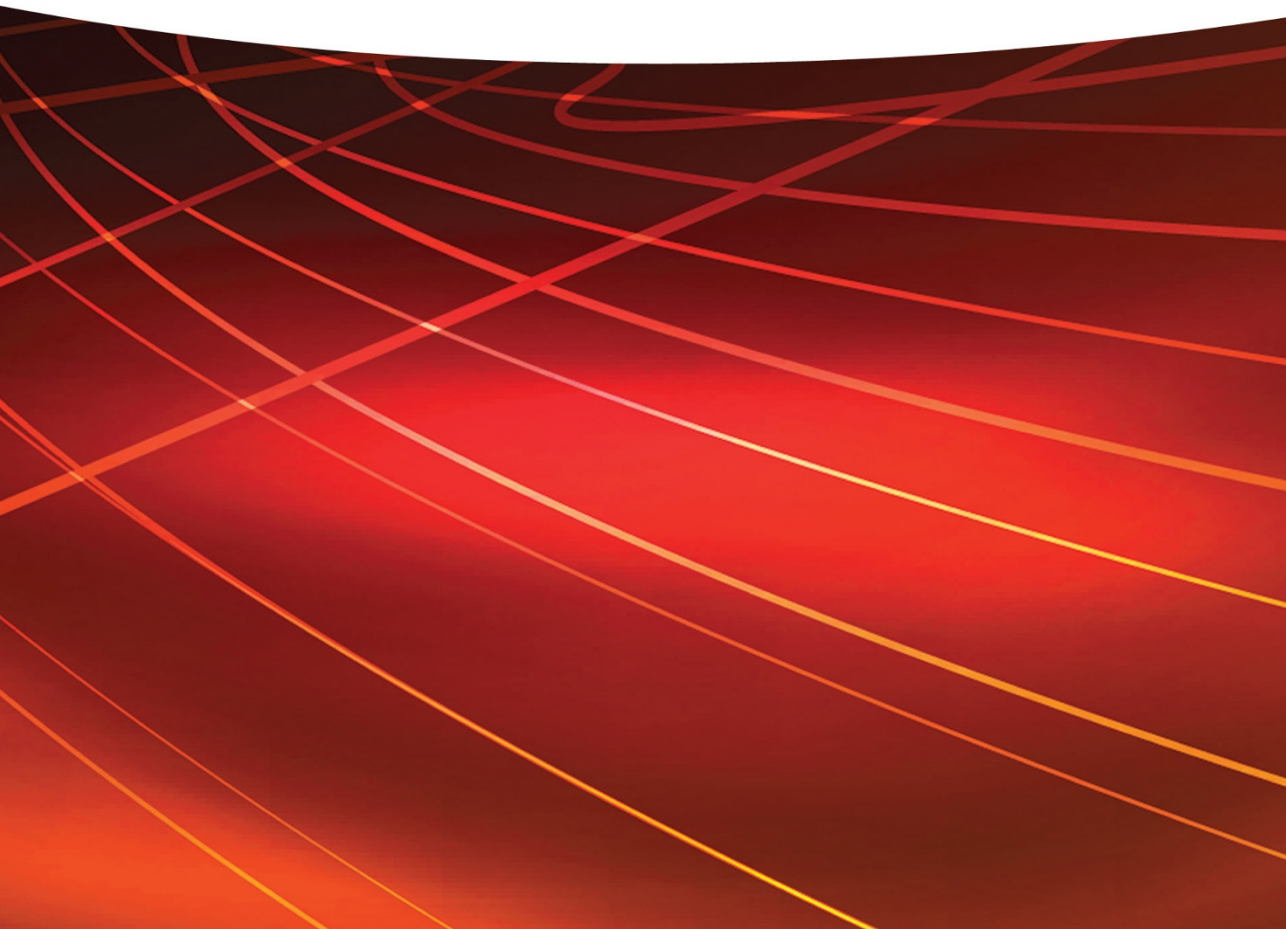
WILEY-VCH

Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë

Quantum Mechanics

Volume I: Basic Concepts, Tools, and Applications

Second Edition



QUANTUM MECHANICS

Volume I

Basic Concepts, Tools, and Applications

Claude Cohen-Tannoudji, Bernard Diu,
and Franck Laloë

Translated from the French by **Susan Reid Hemley, Nicole Ostrowsky, and Dan Ostrowsky**

Authors

Prof. Dr. Claude Cohen-Tannoudji
Laboratoire Kastler Brossel (ENS)
24 rue Lhomond
75231 Paris Cedex 05
France

Prof. Dr. Bernard Diu
4 rue du Docteur Roux
91440 Boures-sur-Yvette
France

Prof. Dr. Frank Laloë
Laboratoire Kastler Brossel (ENS)
24 rue Lhomond
75231 Paris Cedex 05
France

Cover Image
© antishock/Getty Images

Second Edition

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data:
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de.

© 2020 WILEY-VCH Verlag GmbH & Co. KGaA, Boschstr. 12, 69469 Weinheim, Germany

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Print ISBN 978-3-527-34553-3
ePDF ISBN 978-3-527-82270-6
ePub ISBN 978-3-527-82271-3

Cover Design Tata Consulting Services
Printing and Binding CPI Ebner & Spiegel

Printed on acid-free paper.

Directions for Use

This book is composed of chapters and their complements:

– *The chapters* contain the fundamental concepts. Except for a few additions and variations, they correspond to a course given in the last year of a typical undergraduate physics program (Volume I) or of a graduate program (Volumes II and III). The 21 chapters are *complete in themselves* and can be studied independently of the complements.

– *The complements* follow the corresponding chapter. Each is labelled by a letter followed by a subscript, which gives the number of the chapter (for example, the complements of Chapter V are, in order, A_V, B_V, C_V, etc.). They can be recognized immediately by the symbol • that appears at the top of each of their pages.

The complements vary in character. Some are intended to expand the treatment of the corresponding chapter or to provide a more detailed discussion of certain points. Others describe concrete examples or introduce various physical concepts. One of the complements (usually the last one) is a collection of exercises.

The *difficulty* of the complements varies. Some are very simple examples or extensions of the chapter. Others are more difficult and at the graduate level or close to current research. In any case, the reader should have studied the material in the chapter before using the complements.

The complements are generally independent of one another. *The student should not try to study all the complements of a chapter at once.* In accordance with his/her aims and interests, he/she should choose a small number of them (two or three, for example), plus a few exercises. The other complements can be left for later study. To help with the choise, the complements are listed at the end of each chapter in a "*reader's guide*", which discusses the difficulty and importance of each.

Some passages within the book have been set in small type, and these can be omitted on a first reading.

Foreword

Quantum mechanics is a branch of physics whose importance has continually increased over the last decades. It is essential for understanding the structure and dynamics of microscopic objects such as atoms, molecules and their interactions with electromagnetic radiation. It is also the basis for understanding the functioning of numerous new systems with countless practical applications. This includes lasers (in communications, medicine, milling, etc.), atomic clocks (essential in particular for the GPS), transistors (communications, computers), magnetic resonance imaging, energy production (solar panels, nuclear reactors), etc. Quantum mechanics also permits understanding surprising physical properties such as superfluidity or supraconductivity. There is currently a great interest in entangled quantum states whose non-intuitive properties of nonlocality and nonseparability permit conceiving remarkable applications in the emerging field of quantum information. Our civilization is increasingly impacted by technological applications based on quantum concepts. This why a particular effort should be made in the teaching of quantum mechanics, which is the object of these three volumes.

The first contact with quantum mechanics can be disconcerting. Our work grew out of the authors' experiences while teaching quantum mechanics for many years. It was conceived with the objective of easing a first approach, and then aiding the reader to progress to a more advance level of quantum mechanics. The first two volumes, first published more than forty years ago, have been used throughout the world. They remain however at an intermediate level. They have now been completed with a third volume treating more advanced subjects. Throughout we have used a progressive approach to problems, where no difficulty goes untreated and each aspect of the diverse questions is discussed in detail (often starting with a classical review).

This willingness to go further "without cheating or taking shortcuts" is built into the book structure, using two distinct linked texts: *chapters* and *complements*. As we just outlined in the "Directions for use", the chapters present the general ideas and basic concepts, whereas the complements illustrate both the methods and concepts just exposed.

Volume I presents a general introduction of the subject, followed by a second chapter describing the basic mathematical tools used in quantum mechanics. While this chapter can appear long and dense, the teaching experience of the authors has shown that such a presentation is the most efficient. In the third chapter the postulates are announced and illustrated in many of the complements. We then go on to certain important applications of quantum mechanics, such as the harmonic oscillator, which lead to numerous applications (molecular vibrations, phonons, etc.). Many of these are the object of specific complements.

Volume II pursues this development, while expanding its scope at a slightly higher level. It treats collision theory, spin, addition of angular momenta, and both time-dependent and time-independent perturbation theory. It also presents a first approach to the study of identical particles. In this volume as in the previous one, each theoretical concept is immediately illustrated by diverse applications presented in the complements. Both volumes I and II have benefited from several recent corrections, but there have also been additions. Chapter XIII now contains two sections §§ D and E that treat random perturbations, and a complement concerning relaxation has been added.

Volume III extends the two volumes at a slightly higher level. It is based on the use of the creation and annihilation operator formalism (second quantization), which is commonly used in quantum field theory. We start with a study of systems of identical particles, fermions or bosons. The properties of ideal gases in thermal equilibrium are presented. For fermions, the Hartree-Fock method is developed in detail. It is the base of many studies in chemistry, atomic physics and solid state physics, etc. For bosons, the Gross-Pitaevskii equation and the Bogolubov theory are discussed. An original presentation that treats the pairing effect of both fermions and bosons permits obtaining the BCS (Bardeen-Cooper-Schrieffer) and Bogolubov theories in a unified framework. The second part of volume III treats quantum electrodynamics, its general introduction, the study of interactions between atoms and photons, and various applications (spontaneous emission, multiphoton transitions, optical pumping, etc.). The dressed atom method is presented and illustrated for concrete cases. A final chapter discusses the notion of quantum entanglement and certain fundamental aspects of quantum mechanics, in particular the Bell inequalities and their violations.

Finally note that we have not treated either the philosophical implications of quantum mechanics, or the diverse interpretations of this theory, despite the great interest of these subjects. We have in fact limited ourselves to presenting what is commonly called the "orthodox point of view". It is only in Chapter XXI that we touch on certain questions concerning the foundations of quantum mechanics (nonlocality, etc.). We have made this choice because we feel that one can address such questions more efficiently after mastering the manipulation of the quantum mechanical formalism as well as its numerous applications. These subjects are addressed in the book *Do we really understand quantum mechanics?* (F. Laloë, Cambridge University Press, 2019); see also section 5 of the bibliography of volumes I and II.

Acknowledgments:

Volumes I and II:

The teaching experience out of which this text grew were group efforts, pursued over several years. We wish to thank all the members of the various groups and particularly Jacques Dupont-Roc and Serge Haroche, for their friendly collaboration, for the fruitful discussions we have had in our weekly meetings and for the ideas for problems and exercises that they have suggested. Without their enthusiasm and valuable help, we would never have been able to undertake and carry out the writing of this book.

Nor can we forget what we owe to the physicists who introduced us to research, Alfred Kastler and Jean Brossel for two of us and Maurice Levy for the third. It was in the context of their laboratories that we discovered the beauty and power of quantum mechanics. Neither have we forgotten the importance to us of the modern physics taught at the C.E.A. by Albert Messiah, Claude Bloch and Anatole Abragam, at a time when graduate studies were not yet incorporated into French university programs.

We wish to express our gratitude to Ms. Aucher, Baudrit, Boy, Brodschi, Emo, Heywaerts, Lemirre, Touzeau for preparation of the mansucript.

Volume III:

We are very grateful to Nicole and Daniel Ostrowsky, who, as they translated this Volume from French into English, proposed numerous improvements and clarifications. More recently, Carsten Henkel also made many useful suggestions during his translation of the text into German; we are very grateful for the improvements of the text that resulted from this exchange. There are actually many colleagues and friends who greatly contributed, each in his own way, to finalizing this book. All their complementary remarks and suggestions have been very helpful and we are in particular thankful to:

Pierre-François Cohadon
Jean Dalibard
Sébastien Gleyzes
Markus Holzmann
Thibaut Jacqmin
Philippe Jacquier
Amaury Mouchet
Jean-Michel Raimond
Félix Werner

Some delicate aspects of Latex typography have been resolved thanks to Marco Picco, Pierre Cladé and Jean Hare. Roger Balian, Edouard Brézin and William Mullin have offered useful advice and suggestions. Finally, our sincere thanks go to Geneviève Tastevin, Pierre-François Cohadon and Samuel Deléglise for their help with a number of figures.

Volume I

Table of contents **vii**

I WAVES AND PARTICLES. INTRODUCTION TO THE BASIC IDEAS OF QUANTUM MECHANICS **1**
A Electromagnetic waves and photons 3
B Material particles and matter waves 10
C Quantum description of a particle. Wave packets 13
D Particle in a time-independent scalar potential 23

READER'S GUIDE FOR COMPLEMENTS **33**

A_I Order of magnitude of the wavelengths associated with material particles **35**

B_I Constraints imposed by the uncertainty relations **39**
1 Macroscopic system 39
2 Microscopic system 40

C_I Heisenberg relation and atomic parameters **41**

D_I An experiment illustrating the Heisenberg relations **45**

E_I A simple treatment of a two-dimensional wave packet **49**
1 Introduction 49
2 Angular dispersion and lateral dimensions 49
3 Discussion 51

F_I The relationship between one- and three-dimensional problems **53**
1 Three-dimensional wave packet 53
2 Justification of one-dimensional models 56

G_I One-dimensional Gaussian wave packet: spreading of the wave packet **57**
1 Definition of a Gaussian wave packet 57
2 Calculation of Δx and Δp; uncertainty relation 58
3 Evolution of the wave packet 59

H_I Stationary states of a particle in one-dimensional square potentials **63**
1 Behavior of a stationary wave function $\varphi(x)$ 63
2 Some simple cases 65

J_I Behavior of a wave packet at a potential step **75**
1 Total reflection: $E < V_0$ 75
2 Partial reflection: $E > V_0$ 79

K_I Exercises **83**

****

II THE MATHEMATICAL TOOLS OF QUANTUM MECHANICS 87
A Space of the one-particle wave function 88
B State space. Dirac notation 102
C Representations in state space 116
D Eigenvalue equations. Observables 126
E Two important examples of representations and observables 139
F Tensor product of state spaces 147

READER'S GUIDE FOR COMPLEMENTS 159

A_{II} The Schwarz inequality 161

B_{II} Review of some useful properties of linear operators 163
1 Trace of an operator 163
2 Commutator algebra 165
3 Restriction of an operator to a subspace 165
4 Functions of operators 166
5 Derivative of an operator 169

C_{II} Unitary operators 173
1 General properties of unitary operators 173
2 Unitary transformations of operators 177
3 The infinitesimal unitary operator 178

D_{II} A more detailed study of the r and p representations 181
1 The $\{|\mathbf{r}\rangle\}$ representation 181
2 The $\{|\mathbf{p}\rangle\}$ representation 184

E_{II} Some general properties of two observables, Q and P, whose commutator is equal to $i\hbar$ 187
1 The operator $S(\lambda)$: definition, properties 187
2 Eigenvalues and eigenvectors of Q 188
3 The $\{|\mathbf{q}\rangle\}$ representation 189
4 The $\{|\boldsymbol{p}\rangle\}$ representation. The symmetric nature of the $\boldsymbol{P}$ and $\boldsymbol{Q}$ observables 190

F_{II} The parity operator 193
1 The parity operator 193
2 Even and odd operators 196
3 Eigenstates of an even observable $\boldsymbol{B}_+$ 199
4 Application to an important special case 199

G_{II} An application of the properties of the tensor product: the two-dimensional infinite well 201
1 Definition; eigenstates 201
2 Study of the energy levels 202

H_{II} Exercises 205

III THE POSTULATES OF QUANTUM MECHANICS 213
A Introduction . . . 214
B Statement of the postulates . . . 215
C The physical interpretation of the postulates concerning observables and their measurement . . . 226
D The physical implications of the Schrödinger equation . . . 237
E The superposition principle and physical predictions . . . 253

READER'S GUIDE FOR COMPLEMENTS 267

A_{III} Particle in an infinite one-dimensional potential well 271
1 Distribution of the momentum values in a stationary state . . . 271
2 Evolution of the particle's wave function . . . 275
3 Perturbation created by a position measurement . . . 279

B_{III} Study of the probability current in some special cases 283
1 Expression for the current in constant potential regions . . . 283
2 Application to potential step problems . . . 284
3 Probability current of incident and evanescent waves, in the case of reflection from a two-dimensional potential step . . . 285

C_{III} Root mean square deviations of two conjugate observables 289
1 The Heisenberg relation for P and Q . . . 289
2 The "minimum" wave packet . . . 290

D_{III} Measurements bearing on only one part of a physical system 293
1 Calculation of the physical predictions . . . 293
2 Physical meaning of a tensor product state . . . 295
3 Physical meaning of a state that is not a tensor product . . . 296

E_{III} The density operator 299
1 Outline of the problem . . . 299
2 The concept of a statistical mixture of states . . . 299
3 The pure case. Introduction of the density operator . . . 301
4 A statistical mixture of states (non-pure case) . . . 304
5 Use of the density operator: some applications . . . 308

F_{III} The evolution operator 313
1 General properties . . . 313
2 Case of conservative systems . . . 315

G_{III} The Schrödinger and Heisenberg pictures 317

H_{III} Gauge invariance 321
1 Outline of the problem: scalar and vector potentials associated with an electromagnetic field; concept of a gauge . . . 321
2 Gauge invariance in classical mechanics . . . 322
3 Gauge invariance in quantum mechanics . . . 327

J_{III} Propagator for the Schrödinger equation **335**
1 Introduction . . . 335
2 Existence and properties of a propagator $K(2,1)$ . . . 336
3 Lagrangian formulation of quantum mechanics . . . 339

K_{III} Unstable states. Lifetime **343**
1 Introduction . . . 343
2 Definition of the lifetime . . . 344
3 Phenomenological description of the instability of a state . . . 345

L_{III} Exercises **347**

M_{III} Bound states in a "potential well" of arbitrary shape **359**
1 Quantization of the bound state energies . . . 359
2 Minimum value of the ground state energy . . . 363

N_{III} Unbound states of a particle in the presence of a potential well or barrier **367**
1 Transmission matrix $M(k)$ . . . 368
2 Transmission and reflection coefficients . . . 372
3 Example . . . 373

O_{III} Quantum properties of a particle in a one-dimensional periodic structure **375**
1 Passage through several successive identical potential barriers . . . 376
2 Discussion: the concept of an allowed or forbidden energy band . . . 381
3 Quantization of energy levels in a periodic potential; effect of boundary conditions . . . 383

****

IV APPLICATIONS OF THE POSTULATES TO SIMPLE CASES: SPIN 1/2 AND TWO-LEVEL SYSTEMS **393**
A Spin 1/2 particle: quantization of the angular momentum . . . 394
B Illustration of the postulates in the case of a spin 1/2 . . . 401
C General study of two-level systems . . . 411

READER'S GUIDE FOR COMPLEMENTS **423**

A_{IV} The Pauli matrices **425**
1 Definition; eigenvalues and eigenvectors . . . 425
2 Simple properties . . . 426
3 A convenient basis of the 2×2 matrix space . . . 427

B_{IV} Diagonalization of a 2×2 Hermitian matrix **429**
1 Introduction . . . 429
2 Changing the eigenvalue origin . . . 429
3 Calculation of the eigenvalues and eigenvectors . . . 430

C_{IV} Fictitious spin 1/2 associated with a two-level system **435**
1 Introduction . . . 435
2 Interpretation of the Hamiltonian in terms of fictitious spin . . . 435
3 Geometrical interpretation . . . 437

D_{IV} System of two spin 1/2 particles **441**
1 Quantum mechanical description . . . 441
2 Prediction of the measurement results . . . 444

E_{IV} Spin 1/2 density matrix **449**
1 Introduction . . . 449
2 Density matrix of a perfectly polarized spin (pure case) . . . 449
3 Example of a statistical mixture: unpolarized spin . . . 450
4 Spin 1/2 at thermodynamic equilibrium in a static field . . . 452
5 Expansion of the density matrix in terms of the Pauli matrices . . . 453

F_{IV} Spin 1/2 particle in a static and a rotating magnetic fields: magnetic resonance **455**
1 Classical treatment; rotating reference frame . . . 455
2 Quantum mechanical treatment . . . 458
3 Relation between the classical treatment and the quantum mechanical treatment: evolution of $\langle \mathbf{M} \rangle$ . . . 463
4 Bloch equations . . . 463

G_{IV} A simple model of the ammonia molecule **469**
1 Description of the model . . . 469
2 Eigenfunctions and eigenvalues of the Hamiltonian . . . 471
3 The ammonia molecule considered as a two-level system . . . 477

H_{IV} Effects of a coupling between a stable state and an unstable state **485**
1 Introduction. Notation . . . 485
2 Influence of a weak coupling on states of different energies . . . 486
3 Influence of an arbitrary coupling on states of the same energy . . . 487

J_{IV} Exercises **491**

****

V THE ONE-DIMENSIONAL HARMONIC OSCILLATOR **497**
A Introduction . . . 497
B Eigenvalues of the Hamiltonian . . . 503
C Eigenstates of the Hamiltonian . . . 510
D Discussion . . . 518

READER'S GUIDE FOR COMPLEMENTS **525**

A_V Some examples of harmonic oscillators **527**
1 Vibration of the nuclei of a diatomic molecule . . . 527
2 Vibration of the nuclei in a crystal . . . 534
3 Torsional oscillations of a molecule: ethylene . . . 536
4 Heavy muonic atoms . . . 541

B_V Study of the stationary states in the x representation. Hermite polynomials **547**
1 Hermite polynomials . . . 547
2 The eigenfunctions of the harmonic oscillator Hamiltonian . . . 550

C_V Solving the eigenvalue equation of the harmonic oscillator by the polynomial method **555**
1 Changing the function and the variable . . . 555
2 The polynomial method . . . 557

D_V Study of the stationary states in the momentum representation **563**
1 Wave functions in momentum space . . . 563
2 Discussion . . . 565

E_V The isotropic three-dimensional harmonic oscillator **569**
1 The Hamiltonian operator . . . 569
2 Separation of the variables in Cartesian coordinates . . . 570
3 Degeneracy of the energy levels . . . 572

F_V A charged harmonic oscillator in a uniform electric field **575**
1 Eigenvalue equation of $H'(\mathscr{E})$ in the $\{|x\rangle\}$ representation . . . 575
2 Discussion . . . 577
3 Use of the translation operator . . . 579

G_V Coherent "quasi-classical" states of the harmonic oscillator **583**
1 Quasi-classical states . . . 584
2 Properties of the $|\alpha\rangle$ states . . . 588
3 Time evolution of a quasi-classical state . . . 594
4 Example: quantum mechanical treatment of a macroscopic oscillator . . . 596

H_V Normal vibrational modes of two coupled harmonic oscillators **599**
1 Vibration of the two coupled in classical mechanics . . . 599
2 Vibrational states of the system in quantum mechanics . . . 605

J_V Vibrational modes of an infinite linear chain of coupled harmonic oscillators; phonons **611**
1 Classical treatment . . . 612
2 Quantum mechanical treatment . . . 622
3 Application to the study of crystal vibrations: phonons . . . 626

K_V Vibrational modes of a continuous physical system. Photons 631
1 Outline of the problem . . . 631
2 Vibrational modes of a continuous mechanical system: example of a vibrating string . . . 632
3 Vibrational modes of radiation: photons . . . 639

L_V One-dimensional harmonic oscillator in thermodynamic equilibrium at a temperature T 647
1 Mean value of the energy . . . 648
2 Discussion . . . 650
3 Applications . . . 651
4 Probability distribution of the observable X . . . 655

M_V Exercises 661

****

VI GENERAL PROPERTIES OF ANGULAR MOMENTUM IN QUANTUM MECHANICS 667
A Introduction: the importance of angular momentum . . . 667
B Commutation relations characteristic of angular momentum . . . 669
C General theory of angular momentum . . . 671
D Application to orbital angular momentum . . . 685

READER'S GUIDE FOR COMPLEMENTS 703

A_{VI} Spherical harmonics 705
1 Calculation of spherical harmonics . . . 705
2 Properties of spherical harmonics . . . 710

B_{VI} Angular momentum and rotations 717
1 Introduction . . . 717
2 Brief study of geometrical rotations $\mathscr{R}$ . . . 718
3 Rotation operators in state space. Example: a spinless particle . . . 720
4 Rotation operators in the state space of an arbitrary system . . . 727
5 Rotation of observables . . . 730
6 Rotation invariance . . . 734

C_{VI} Rotation of diatomic molecules 739
1 Introduction . . . 739
2 Rigid rotator. Classical study . . . 740
3 Quantization of the rigid rotator . . . 741
4 Experimental evidence for the rotation of molecules . . . 746

D_{VI} Angular momentum of stationary states of a two-dimensional harmonic oscillator 755
1 Introduction . . . 755
2 Classification of the stationary states by the quantum numbers n_x and n_y . . . 759
3 Classification of the stationary states in terms of their angular momenta . . . 761

4 Quasi-classical states 765

E_{VI} A charged particle in a magnetic field: Landau levels 771
1 Review of the classical problem 771
2 General quantum mechanical properties of a particle in a magnetic field . . . 775
3 Case of a uniform magnetic field 779

F_{VI} Exercises 795

****

VII PARTICLE IN A CENTRAL POTENTIAL, HYDROGEN ATOM 803
A Stationary states of a particle in a central potential 804
B Motion of the center of mass and relative motion for a system of two interacting particles 812
C The hydrogen atom 818

READER'S GUIDE FOR COMPLEMENTS 831

A_{VII} Hydrogen-like systems 833
1 Hydrogen-like systems with one electron 834
2 Hydrogen-like systems without an electron 839

B_{VII} A soluble example of a central potential: The isotropic three-dimensional harmonic oscillator 841
1 Solving the radial equation 842
2 Energy levels and stationary wave functions 845

C_{VII} Probability currents associated with the stationary states of the hydrogen atom 851
1 General expression for the probability current 851
2 Application to the stationary states of the hydrogen atom 852

D_{VII} The hydrogen atom placed in a uniform magnetic field. Paramagnetism and diamagnetism. The Zeeman effect 855
1 The Hamiltonian of the problem. The paramagnetic term and the diamagnetic term 856
2 The Zeeman effect 862

E_{VII} Some atomic orbitals. Hybrid orbitals 869
1 Introduction 869
2 Atomic orbitals associated with real wave functions 870
3 *sp* hybridization 876
4 sp^2 hybridization 878
5 sp^3 hybridization 882

F_{VII} Vibrational-rotational levels of diatomic molecules **885**
1 Introduction 885
2 Approximate solution of the radial equation 886
3 Evaluation of some corrections 892

G_{VII} Exercises **899**
1 Particle in a cylindrically symmetric potential 899
2 Three-dimensional harmonic oscillator in a uniform magnetic field 899

INDEX **901**

****

Volume II

VOLUME II **923**

Table of contents **v**

VIII AN ELEMENTARY APPROACH TO THE QUANTUM THEORY OF SCATTERING BY A POTENTIAL **923**

READER'S GUIDE FOR COMPLEMENTS **957**

A_{VIII} The free particle: stationary states with well-defined angular momentum **959**

B_{VIII} Phenomenological description of collisions with absorption **971**

C_{VIII} Some simple applications of scattering theory **977**

****

IX ELECTRON SPIN **985**

READER'S GUIDE FOR COMPLEMENTS **999**

A_{IX} Rotation operators for a spin 1/2 particle **1001**

B_{IX} Exercises **1009**

****

X ADDITION OF ANGULAR MOMENTA **1015**

READER'S GUIDE FOR COMPLEMENTS **1041**

A_X Examples of addition of angular momenta **1043**

B_X Clebsch-Gordan coefficients **1051**

C_X Addition of spherical harmonics **1059**

D_X Vector operators: the Wigner-Eckart theorem **1065**

E_X Electric multipole moments **1077**

F_X Two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ coupled by an interaction $a\mathbf{J}_1 \cdot \mathbf{J}_2$ **1091**

G_X Exercises **1107**

****

XI STATIONARY PERTURBATION THEORY 1115

READER'S GUIDE FOR COMPLEMENTS 1129

A_{XI} A one-dimensional harmonic oscillator subjected to a perturbing potential in x, x^2, x^3 1131

B_{XI} Interaction between the magnetic dipoles of two spin 1/2 particles 1141

C_{XI} Van der Waals forces 1151

D_{XI} The volume effect: the influence of the spatial extension of the nucleus on the atomic levels 1162

E_{XI} The variational method 1169

F_{XI} Energy bands of electrons in solids: a simple model 1177

G_{XI} A simple example of the chemical bond: the H_2^+ ion 1189

H_{XI} Exercises 1221

****

XII AN APPLICATION OF PERTURBATION THEORY: THE FINE AND HYPERFINE STRUCTURE OF THE HYDROGEN ATOM 1231

READER'S GUIDE FOR COMPLEMENTS 1265

A_{XII} The magnetic hyperfine Hamiltonian 1267

B_{XII} Calculation of the average values of the fine-structure Hamiltonian in the $1s$, $2s$ and $2p$ states 1276

C_{XII} The hyperfine structure and the Zeeman effect for muonium and positronium 1281

D_{XII} The influence of the electronic spin on the Zeeman effect of the hydrogen resonance line 1289

E_{XII} The Stark effect for the hydrogen atom 1298

****

XIII APPROXIMATION METHODS FOR TIME-DEPENDENT PROBLEMS 1303

READER'S GUIDE FOR COMPLEMENTS **1337**

A_{XIII} Interaction of an atom with an electromagnetic wave 1339

B_{XIII} Linear and non-linear responses of a two-level system subject to a sinusoidal perturbation 1357

C_{XIII} Oscillations of a system between two discrete states under the effect of a sinusoidal resonant perturbation 1374

D_{XIII} Decay of a discrete state resonantly coupled to a continuum of final states 1378

E_{XIII} Time-dependent random perturbation, relaxation 1390

F_{XIII} Exercises 1409

****

XIV SYSTEMS OF IDENTICAL PARTICLES 1419

READER'S GUIDE FOR COMPLEMENTS **1457**

A_{XIV} Many-electron atoms. Electronic configurations 1459

B_{XIV} Energy levels of the helium atom. Configurations, terms, multiplets 1467

C_{XIV} Physical properties of an electron gas. Application to solids 1481

D_{XIV} Exercises 1496

****

APPENDICES 1505

I Fourier series and Fourier transforms 1505

II The Dirac δ-"function" 1515

III Lagrangian and Hamiltonian in classical mechanics 1527

BIBLIOGRAPHY OF VOLUMES I AND II 1545

INDEX 1569

****

Volume III

VOLUME III **1591**

Table of contents **v**

XV CREATION AND ANNIHILATION OPERATORS FOR IDENTICAL PARTICLES **1591**

READER'S GUIDE FOR COMPLEMENTS **1617**

A_{XV} Particles and holes **1621**

B_{XV} Ideal gas in thermal equilibrium; quantum distribution functions **1625**

C_{XV} Condensed boson system, Gross-Pitaevskii equation **1643**

D_{XV} Time-dependent Gross-Pitaevskii equation **1657**

E_{XV} Fermion system, Hartree-Fock approximation **1677**

F_{XV} Fermions, time-dependent Hartree-Fock approximation **1701**

G_{XV} Fermions or Bosons: Mean field thermal equilibrium **1711**

H_{XV} Applications of the mean field method for non-zero temperature **1733**

****

XVI FIELD OPERATOR **1751**

READER'S GUIDE FOR COMPLEMENTS **1767**

A_{XVI} Spatial correlations in an ideal gas of bosons or fermions **1769**

B_{XVI} Spatio-temporal correlation functions, Green's functions **1781**

C_{XVI} Wick's theorem **1799**

****

XVII PAIRED STATES OF IDENTICAL PARTICLES **1811**

READER'S GUIDE FOR COMPLEMENTS **1843**

A_{XVII} Pair field operator for identical particles 1845

B_{XVII} Average energy in a paired state 1869

C_{XVII} Fermion pairing, BCS theory 1889

D_{XVII} Cooper pairs 1927

E_{XVII} Condensed repulsive bosons 1933

****

XVIII REVIEW OF CLASSICAL ELECTRODYNAMICS 1957

READER'S GUIDE FOR COMPLEMENTS 1977

A_{XVIII} Lagrangian formulation of electrodynamics 1979

****

XIX QUANTIZATION OF ELECTROMAGNETIC RADIATION 1997

READER'S GUIDE FOR COMPLEMENTS 2017

A_{XIX} Momentum exchange between atoms and photons 2019

B_{XIX} Angular momentum of radiation 2043

C_{XIX} Angular momentum exchange between atoms and photons 2055

****

XX ABSORPTION, EMISSION AND SCATTERING OF PHOTONS BY ATOMS 2067

READER'S GUIDE FOR COMPLEMENTS 2095

A_{XX} A multiphoton process: two-photon absorption 2097

B_{XX} Photoionization 2109

C_{XX} Two-level atom in a monochromatic field. Dressed-atom method 2129

D_{XX} Light shifts: a tool for manipulating atoms and fields 2151

E_{XX} Detection of one- or two-photon wave packets, interference 2163

****

XXI QUANTUM ENTANGLEMENT, MEASUREMENTS, BELL'S INEQUALITIES 2187

READER'S GUIDE FOR COMPLEMENTS 2215

A_{XXI} Density operator and correlations; separability 2217

B_{XXI} GHZ states, entanglement swapping 2227

C_{XXI} Measurement induced relative phase between two condensates 2237

D_{XXI} Emergence of a relative phase with spin condensates; macroscopic non-locality and the EPR argument 2253

****

APPENDICES 2267

IV Feynman path integral 2267

V Lagrange multipliers 2281

VI Brief review of Quantum Statistical Mechanics 2285

VII Wigner transform 2297

BIBLIOGRAPHY OF VOLUME III 2325

INDEX 2333

Chapter I

Waves and particles. Introduction to the fundamental ideas of quantum mechanics

A Electromagnetic waves and photons 3
A-1 Light quanta and the Planck-Einstein relations 3
A-2 Wave-particle duality 3
A-3 The principle of spectral decomposition 7
B Material particles and matter waves 10
B-1 The de Broglie relations 10
B-2 Wave functions. Schrödinger equation 11
C Quantum description of a particle. Wave packets 13
C-1 Free particle 14
C-2 Form of the wave packet at a given time 15
C-3 Heisenberg relations 19
C-4 Time evolution of a free wave packet 20
D Particle in a time-independent scalar potential 23
D-1 Separation of variables. Stationary states 24
D-2 One-dimensional "square" potentials. Qualitative study . . . 26

In the present state of scientific knowledge, quantum mechanics plays a fundamental role in the description and understanding of natural phenomena. In fact, phenomena that occur on a very small (atomic or subatomic) scale cannot be explained outside the framework of quantum physics. For example, the existence and the properties of atoms, the chemical bond and the propagation of an electron in a crystal cannot be understood

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

in terms of classical mechanics. Even when we are concerned only with macroscopic physical objects (that is, whose dimensions are comparable to those encountered in everyday life), it is necessary, in principle, to begin by studying the behavior of their various constituent atoms, ions, electrons, in order to arrive at a complete scientific description. Actually, there are many phenomena that reveal, on a macroscopic scale, the quantum behaviour of nature. It is in this sense that it can be said that quantum mechanics is the basis of our present understanding of all natural phenomena, including those traditionally treated in chemistry, biology, etc...

From a historical point of view, quantum ideas contributed to a remarkable unification of the concepts of fundamental physics by treating material particles and radiation on the same footing. At the end of the nineteenth century, people distinguished between two entities in physical phenomena: matter and radiation. Completely different laws were used for each one. To predict the motion of material bodies, the laws of *Newtonian mechanics* (*cf.* Appendix III) were utilized. Their success, though of long standing, was none the less impressive. With regard to radiation, the *theory of electromagnetism*, thanks to the introduction of Maxwell's equations, had produced a unified interpretation of a set of phenomena which had previously been considered as belonging to different domains: electricity, magnetism and optics. In particular, the electromagnetic theory of radiation had been spectacularly confirmed experimentally by the discovery of Hertzian waves. Finally, *interactions between radiation and matter* were well explained by the Lorentz force. This set of laws had brought physics to a point which could be considered satisfactory, in view of the experimental data at the time.

However, at the beginning of the twentieth century, physics was to be marked by the profound upheaval that led to the introduction of relativistic mechanics and quantum mechanics. The relativistic "revolution" and the quantum "revolution" were, to a large extent, independent, since they challenged classical physics on different points. Classical laws cease to be valid for material bodies travelling at very high speeds, comparable to that of light (relativistic domain). In addition, they are also found to be wanting on an atomic or subatomic scale (quantum domain). However, it is important to note that classical physics, in both cases, can be seen as an approximation of the new theories, an approximation which is valid for most phenomena on an everyday scale. For example, Newtonian mechanics enables us to predict correctly the motion of a solid body, providing it is non-relativistic (speeds much smaller than that of light) and macroscopic (dimensions much greater than atomic ones). Nevertheless, from a fundamental point of view, quantum theory remains indispensable. It is the only theory which enables us to understand the very existence of a solid body and the values of the macroscopic parameters (density, specific heat, elasticity, etc...) associated with it. It is possible to develop a theory that is at the same time quantum and relativistic, but such a theory is relatively complex. However, most atomic and molecular phenomena are well explained by the *non-relativistic quantum mechanics* that we intend to examine here.

This chapter is an introduction to quantum ideas and "vocabulary". *No attempt is made here to be rigorous or complete.* The essential goal is to awaken the curiosity of the reader. Phenomena will be described which unsettle ideas as firmly anchored in our intuition as the concept of a trajectory. We want to render the quantum theory "plausible" for the reader by showing simply and qualitatively how it enables us to solve the problems which are encountered on an atomic scale. We shall later return to the various ideas introduced in this chapter and go into further detail, either from the point

of view of the mathematical formalism (Chap. II) or from the physical point of view (Chap. III).

In the first section (§ A), we introduce the basic quantum ideas (wave-particle duality, the measurement process), relying on well-known optical experiments. Then we show (§ B) how these ideas can be extended to material particles (wave function, Schrödinger equation). We next study in more detail the characteristics of the "wave packet" associated with a particle, and we introduce the Heisenberg relations (§ C). Finally, we discuss some simple examples of typical quantum effects (§ D).

A. Electromagnetic waves and photons

A-1. Light quanta and the Planck-Einstein relations

Newton considered light to be a beam of particles, able, for example, to bounce back upon reflection from a mirror. During the first half of the nineteenth century, the wavelike nature of light was demonstrated (interference, diffraction). This later enabled optics to be integrated into electromagnetic theory. In this framework, the speed of light, c, is related to electric and magnetic constants and light polarization phenomena can be interpreted as manifestations of the vectorial character of the electric field.

However, the study of *blackbody* radiation, which electromagnetic theory could not explain, led Planck to suggest the hypothesis of the *quantization of energy* (1900): for an electromagnetic wave of frequency ν, the only possible energies are integral multiples of the quantum $h\nu$, where h is a new fundamental constant. Generalizing this hypothesis, Einstein proposed a return to the particle theory (1905): light consists of a beam of *photons*, each possessing an energy $h\nu$. Einstein showed how the introduction of photons made it possible to understand, in a very simple way, certain as yet unexplained characteristics of the photoelectric effect. Twenty years had to elapse before the photon was actually shown to exist, as a distinct entity, by the Compton effect (1924).

These results lead to the following conclusion: the interaction of an electromagnetic wave with matter occurs by means of *elementary indivisible processes*, in which the radiation appears to be composed of particles, the photons. Particle parameters (the energy E and the momentum $\mathbf{p}$ of a photon) and wave parameters (the angular frequency $\omega = 2\pi\nu$ and the wave vector $\mathbf{k}$, where $|\mathbf{k}| = 2\pi/\lambda$, with ν the frequency and λ the wavelength) are linked by the fundamental relations:

$$\boxed{\begin{array}{l} E = h\nu = \hbar\omega \\ \mathbf{p} = \hbar\,\mathbf{k} \end{array}} \qquad \text{(Planck-Einstein relations)} \tag{A-1}$$

where $\hbar = h/2\pi$ is defined in terms of the Planck constant h:

$$h \simeq 6.62\ \ 10^{-34}\ \ \text{Joule} \times \text{second} \tag{A-2}$$

During each elementary process, energy and total momentum must be conserved.

A-2. Wave-particle duality

Thus we have returned to a particle conception of light. Does this mean that we must abandon the wave theory? Certainly not. We shall see that typical wave

phenomena such as interference and diffraction could not be explained in a purely particle framework. Analyzing Young's well-known double-slit experiment will lead us to the following conclusion: a complete interpretation of the phenomena can be obtained only by conserving *both* the wave aspect and the particle aspect of light (although they seem *a priori* irreconcilable). We shall then show how this paradox can be resolved by the introduction of the fundamental quantum concepts.

A-2-a. Analysis of Young's double-slit experiment

The device used in this experiment is shown schematically in Figure 1. The monochromatic light emitted by the source $\mathscr{S}$ falls on an opaque screen $\mathscr{P}$ pierced by two narrow slits F_1 and F_2, which illuminate the observation screen $\mathscr{E}$ (a photographic plate, for example). If we block F_2, we obtain on $\mathscr{E}$ a light intensity distribution $I_1(x)$ which is the diffraction pattern of F_1. In the same way, when F_1 is obstructed, the diffraction pattern of F_2 is described by $I_2(x)$. When the two slits F_1 and F_2 are open at the same time, we observe a system of interference fringes on the screen. In particular, we note that the corresponding intensity $I(x)$ is not the sum of the intensities produced by F_1 and F_2 separately:

$$I(x) \neq I_1(x) + I_2(x) \tag{A-3}$$

How could one conceive of explaining, in terms of a particle theory (seen, in the preceding section, to be necessary), the experimental results just described? The existence of a diffraction pattern when only one of the two slits is open could, for example, be explained as being due to photon collisions with the edges of the slit. Such an explanation would, of course, have to be developed more precisely, and a more detailed study would show it to be insufficient. Instead, let us concentrate on the interference phenomenon. We could attempt to explain it by an interaction between the photons which pass through the slit F_1 and those which pass through the slit F_2. Such an explanation would lead to the following prediction: if the intensity of the source $\mathscr{S}$ (the number of photons emitted per second) is diminished until the photons strike the screen practically one by one, the interaction between the photons must diminish and, eventually, vanish. The interference fringes should therefore disappear.

Before we indicate the answer given by experiment, recall that the wave theory provides a completely natural interpretation of the fringes. The light intensity at a point of the screen $\mathscr{E}$ is proportional to the square of the amplitude of the electric field at this point. If $E_1(x)$ and $E_2(x)$ represent, in complex notation, the electric fields produced at x by slits F_1 and F_2 respectively (the slits behave like secondary sources), the total resultant field at this point when F_1 and F_2 are both open is[1]:

$$E(x) = E_1(x) + E_2(x) \tag{A-4}$$

Using complex notation, we then have:

$$I(x) \propto |E(x)|^2 = |E_1(x) + E_2(x)|^2 \tag{A-5}$$

Since the intensities $I_1(x)$ and $I_2(x)$ are proportional, respectively, to $|E_1(x)|^2$ and $|E_2(x)|^2$, formula (A-5) shows that $I(x)$ differs from $I_1(x) + I_2(x)$ by an interference

[1] Since the experiment studied here is performed with unpolarized light, the vectorial character of the electric field does not play an essential role. For the sake of simplicity, we ignore it in this paragraph.

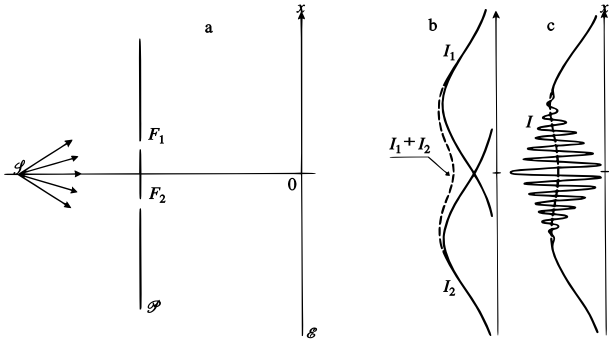


Figure 1: Diagram of Young's double-slit light interference experiment (fig. a). Each of the slits F_1 and F_2 produces a diffraction pattern on the screen $\mathcal{E}$. The corresponding intensities are $I_1(x)$ and $I_2(x)$ (solid lines in figure b). When the two slits F_1 and F_2 are open simultaneously, the intensity $I(x)$ observed on the screen is not the sum $I_1(x) + I_2(x)$ (dashed lines in figure b), but shows oscillations due to the interference between the electric fields radiated by F_1 and F_2 (solid line in figure c).

term which depends on the phase difference between E_1 and E_2 and whose presence explains the fringes. The wave theory thus predicts that diminishing the intensity of the source $\mathscr{S}$ will simply cause the fringes to diminish in intensity but not vanish.

What actually happens when $\mathscr{S}$ emits photons practically one by one? *Neither the predictions of the wave theory nor those of the particle theory are verified.* In fact:

(*i*) If we cover the screen $\mathscr{E}$ with a photographic plate and increase the exposure time so as to capture a large number of photons on each photograph, we observe when we develop them that the *fringes have not disappeared. Therefore, the purely corpuscular interpretation*, according to which the fringes are due to an interaction between photons, *must be rejected.*

(*ii*) On the other hand, we can expose the photographic plate during a time so short that it can only receive a few photons. We then observe that each photon produces a *localized impact* on $\mathscr{E}$ and not a very weak interference pattern. *Therefore, the purely wave interpretation must also be rejected.*

In reality, as more and more photons strike the photographic plate, the following phenomenon occurs. Their individual impacts seem to be distributed in a *random manner*, and only when a great number of them have reached $\mathscr{E}$ does the distribution of the impacts begin to have a continuous aspect. The density of the impacts at each point of $\mathscr{E}$ corresponds to the interference fringes: maximum on a bright fringe and zero on a dark fringe. It can thus be said that the photons, as they arrive, build up the interference pattern.

The result of this experiment therefore leads, apparently, to a paradox. Within the framework of the particle theory, for example, it can be expressed in the following way. Since photon-photon interactions are excluded, each photon must be considered separately. But then it is not clear why the phenomena should change drastically according to whether only one slit or both slits are open. For a photon passing through one of the slits, why should the fact that the other is open or closed have such a critical

importance?

Before we discuss this problem, note that in the preceding experiment we did not seek to determine through which slit each photon passed before it reached the screen. In order to obtain this information, we can imagine placing detectors (photomultipliers) behind F_1 and F_2. It will then be observed that, if the photons arrive one by one, each one passes through a well-determined slit (a signal is recorded either by the detector placed behind F_1 or by the one covering F_2 but not by both at once). But, obviously, the photons detected in this way are absorbed and do not reach the screen. Remove the photomultiplier which blocks F_1, for example. The one which remains behind F_2 tells us that, out of a large number of photons, about half pass through F_2. We conclude that the others (which can continue as far as the screen) pass through F_1. But the pattern that they gradually construct on the screen is not an interference pattern, since F_2 is blocked. It is only the diffraction pattern of F_1.

A-2-b. Quantum unification of the two aspects of light

The preceding analysis shows that it is impossible to explain all the phenomena observed if only one of the two aspects of light, wave or particle, is considered. Now these two aspects seem to be mutually exclusive. To overcome this difficulty, it thus becomes indispensable to reconsider in a critical way the concepts of classical physics. We must accept the possibility that these concepts, although our everyday experience leads us to consider them well-founded, may not be valid in the new ("microscopic") domain which we are entering. For example, an essential characteristic of this new domain appeared when we placed counters behind Young's slits: *when one performs a measurement on a microscopic system, one disturbs it in a fundamental fashion.* This is a new property since, in the macroscopic domain, we always have the possibility of conceiving measurement devices whose influence on the system is practically as weak as one might wish. This critical revision of classical physics is imposed by experiment and must of course be guided by experiment.

Let us reconsider the "paradox" stated above concerning the photon which passes through one slit but behaves differently depending on whether the other slit is open or closed. We saw that if we try to detect the photons when they cross the slits, we prevent them from reaching the screen. More generally, a detailed experimental analysis shows that *it is impossible to observe the interference pattern and to know at the same time through which slit each photon has passed* (*cf.* Complement $\mathrm{D_I}$). Thus it is necessary, in order to resolve the paradox, to give up the idea that a photon inevitably passes through a particular slit. We are then led to question the concept, which is a fundamental one of classical physics, of a particle's trajectory.

Moreover, as the photons arrive one by one, their impacts on the screen gradually build up the interference pattern. This implies that, for a particular photon, we are not certain in advance where it will strike the screen. Now these photons are all emitted under the same conditions. Thus another classical idea has been destroyed: that the initial conditions completely determine the subsequent motion of a particle. We can only say, when a photon is emitted, that the probability of its striking the screen at x is proportional to the intensity $I(x)$ calculated using wave theory, that is, to $|E(x)|^2$.

After many tentative efforts that we shall not describe here, the concept of *wave-particle duality* was formulated. We can summarize it schematically as follows[2]:

[2]It is worth noting that this interpretation of physical phenomena, generally considered to be "ortho-

(i) The particle and wave aspects of light are inseparable. *Light behaves simultaneously like a wave and like a flux of particles, the wave enabling us to calculate the probability of the manifestation of a particle.*

(ii) Predictions about the behavior of a photon can only be probabilistic.

(iii) The information about a photon at time t is given by the wave $E(\mathbf{r}, t)$, which is a solution of Maxwell's equations. We say that this wave characterizes the state of the photons at time t; $E(\mathbf{r}, t)$ is interpreted as the *probability amplitude* of a photon appearing, at time t, at the point $\mathbf{r}$. This means that the corresponding probability is proportional to $|E(\mathbf{r}, t)|^2$.

Comments:

(i) Since Maxwell's equations are linear and homogeneous, we can use a *superposition principle*: if E_1 and E_2 are two solutions of these equations, then $E = \lambda_1 E_1 + \lambda_2 E_2$, where λ_1 and λ_2 are constants, is also a solution. It is this superposition principle which explains wave phenomena in classical optics (interference, diffraction). In quantum physics, the interpretation of $E(\mathbf{r}, t)$ as a probability amplitude is thus essential to the persistence of such phenomena.

(ii) The theory merely allows one to calculate the probability of the occurence of a given event. Experimental verifications must thus be founded on the repetition of a large number of identical experiments. In the above experiment, a large number of photons, all produced in the same way, are emitted successively and build up the interference pattern, according to the calculated probabilities.

(iii) We are talking here about "the photon state" so as to be able to develop in § B an analogy between $E(\mathbf{r}, t)$ and the wave function $\psi(\mathbf{r}, t)$ that characterizes the quantum state of a material particle. This "optical analogy" is very fruitful. In particular, as we shall see in § D, it allows us to understand, simply and without recourse to calculation, various quantum properties of material particles. However, we should not push it too far and let it lead us to believe that it is rigorously correct to consider $E(\mathbf{r}, t)$ as characterizing the quantum state of a photon.

Furthermore, we shall see that the fact that $\psi(\mathbf{r}, t)$ is complex is essential in quantum mechanics, while the complex notation $E(\mathbf{r}, t)$ is used in optics purely for convenience (only its real part has a physical meaning). The precise definition of the (complex) quantum state of radiation can only be given in the framework of quantum electrodynamics, a theory which is simultaneously quantum mechanical and relativistic. We shall not consider these problems here (they will briefly be discussed in Complement $\mathrm{K_V}$).

A-3. The principle of spectral decomposition

Armed with the ideas introduced in § A-2, we are now going to discuss another simple optical experiment, whose subject is the polarization of light. This will permit us to introduce the fundamental concepts which concern the measurement of physical quantities.

dox", is not unique; other interpretations have been proposed, which are still discussed among physicists.

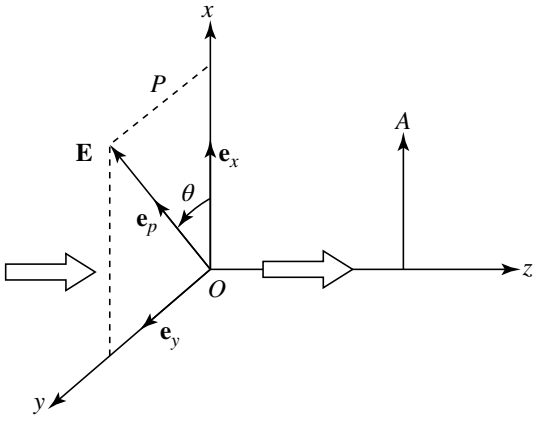


Figure 2: A simple measurement experiment relating to the polarization of a light wave. A beam of light propagates along the direction Oz and crosses successively the polarizer P and the analyzer A; θ is the angle between Ox and the electric field of the wave transmitted by P. The vibrations transmitted by A are parallel to Ox.

The experiment consists of directing a polarized monochromatic plane light wave onto an analyzer A. Oz designates the direction of propagation of this wave and $\mathbf{e}_p$, the unit vector describing its polarization (*cf.* Fig. 2). The analyzer A transmits light polarized parallel to Ox and absorbs light polarized parallel to Oy.

The classical description of this experiment (a description which is valid for a sufficiently intense light beam) is the following. The polarized plane wave is characterized by an electric field of the form:

$$\mathbf{E}(\mathbf{r},t) = E_0\ \mathbf{e}_p\ \mathrm{e}^{i(kz-\omega t)} \tag{A-6}$$

where E_0 is a constant. The light intensity I is proportional to $|E_0|^2$. After its passage through the analyzer A, the plane wave is polarized along Ox:

$$\mathbf{E}'(\mathbf{r},t) = E'_0\ \mathbf{e}_x\ \mathrm{e}^{i(kz-\omega t)} \tag{A-7}$$

and its intensity I', proportional to $|E'_0|^2$, is given by *Malus' law*:

$$I' = I\ \cos^2\ \theta \tag{A-8}$$

[$\mathbf{e}_x$ is the unit vector of the Ox axis and θ is the angle between $\mathbf{e}_x$ and $\mathbf{e}_p$].

What will happen on the quantum level, that is, when I is weak enough for the photons to reach the analyzer one by one? (We then place a photon detector behind this analyser.) First of all, the detector never registers a "fraction of a photon". Either the photon crosses the analyzer or it is entirely absorbed by it. Next (except in special cases that we shall examine in a moment), we cannot predict with certainty whether a given incident photon will pass or be absorbed. We can only know the corresponding

probabilities. Finally, if we send out a large number N of photons one after the other, the result will correspond to the classical law, in the sense that about $N \cos^2 \theta$ photons will be detected after the analyzer.

We shall retain the following ideas from this description:

(*i*) The measurement device (the analyzer, in this case) can give only certain privileged results, which we shall call *eigen* (or proper) *results*[3]. In the above experiment, there are only two possible results: the photon crosses the analyzer or it is stopped. One says that there is quantization of the result of the measurement, in contrast to the classical case [*cf.* formula (A-8)] where the transmitted intensity I' can vary continuously, according to the value of θ, between 0 and I.

(*ii*) To each of these eigen results corresponds an *eigenstate*. Here, the two eigenstates are characterized by:

$$\begin{array}{ll} & \mathbf{e}_p = \mathbf{e}_x \\ \text{or} & \mathbf{e}_p = \mathbf{e}_y \end{array} \tag{A-9}$$

($\mathbf{e}_y$ is the unit vector of the Oy axis). If $\mathbf{e}_p = \mathbf{e}_x$, we know with certainty that the photon will traverse the analyzer; if $\mathbf{e}_p = \mathbf{e}_y$, it will, on the contrary, definitely be stopped. The correspondence between eigen results and eigenstates is therefore the following. If the particle is, before the measurement, in one of the eigenstates, the result of this measurement is certain: it can only be the associated eigen result.

(*iii*) When the state before the measurement is arbitrary, only the probabilities of obtaining the different eigen results can be predicted. To find these probabilities, one decomposes the state of the particles into a linear combination of the various eigenstates. Here, for an arbitrary $\mathbf{e}_p$, we write:

$$\mathbf{e}_p = \mathbf{e}_x \cos \theta + \mathbf{e}_y \sin \theta \tag{A-10}$$

The probability of obtaining a given eigen result is then proportional to the square of the absolute value of the coefficient of the corresponding eigenstate.

The proportionality factor is determined by the condition that the sum of all these probabilities must be equal to 1. We thus deduce from (A-10) that each photon has a probability $\cos^2 \theta$ of traversing the analyzer and a probability $\sin^2 \theta$ of being absorbed by it (we know that $\cos^2 \theta + \sin^2 \theta = 1$). This is indeed what was stated above. This rule is called in quantum mechanics the *principle of spectral decomposition*. Note that the decomposition to be performed depends on the type of measurement device being considered, since one must use the eigenstates which correspond to it: in formula (A-10), the choice of the axes Ox and Oy is fixed by the analyzer.

(*iv*) After passing through the analyzer, the light is completely polarized along $\mathbf{e}_x$. If we place, after the first analyzer A, a second analyzer A', having the same axis, all the photons which traversed A will also traverse A'. According to what we have just seen in point (*ii*), this means that, after they have crossed A, the state of the photons is the eigenstate characterized by $\mathbf{e}_x$. There has therefore been an abrupt change in the state of the particles. Before the measurement, this state was defined by a vector $\mathbf{E}(\mathbf{r}, t)$ which was collinear with $\mathbf{e}_p$. After the measurement, we possess an additional piece of information (the photon has passed) which is incorporated by describing the state by a different vector, which is now collinear with $\mathbf{e}_x$. This expresses the fact, already pointed

[3]The reason for this name will appear in Chapter III.

out in § A-2, that *the measurement disturbs the microscopic system* (here, the photon) in *a fundamental fashion.*

Comment:

The certain prediction of the result when $\mathbf{e}_p = \mathbf{e}_x$ or $\mathbf{e}_p = \mathbf{e}_y$ is only a special case. The probability of one of the possible events is then indeed equal to 1. But, in order to verify this prediction, one must perform a large number of experiments. One must be sure that *all* the photons pass (or are stopped), since the fact that a particular photon crosses the analyzer (or is absorbed) is not characteristic of $\mathbf{e}_p = \mathbf{e}_x$ (or $\mathbf{e}_p = \mathbf{e}_y$).

B. Material particles and matter waves

B-1. The de Broglie relations

Parallel to the discovery of photons, the study of atomic emission and absorption spectra uncovered a fundamental fact, which classical physics was unable to explain: these spectra are composed of *narrow lines.* In other words, a given atom emits or absorbs only photons having well-determined frequencies (that is, energies). This fact can be interpreted very easily if one accepts that *the energy of the atom is quantized*, that is, it can take on only certain discrete values E_i ($i = 1, 2, ..., n, ...$): the emission or absorption of a photon is then accompanied by a "jump" in the energy of the atom from one permitted value E_i to another E_j. Conservation of energy implies that the photon has a frequency ν_{ij} such that:

$$h\,\nu_{ij} = |E_i - E_j| \tag{B-1}$$

Only frequencies which obey (B-1) can therefore be emitted or absorbed by the atom.

The existence of such discrete energy levels was confirmed independently by the Franck-Hertz experiment. Bohr interpreted this in terms of privileged electronic orbits and stated, with Sommerfeld, an empirical rule which permitted the calculation of these orbits for the case of the hydrogen atom. But the fundamental origin of these quantization rules remained mysterious.

In 1923, however, de Broglie put forth the following hypothesis: *material particles, just like photons, can have a wavelike aspect.* He then derived the Bohr-Sommerfeld quantization rules as a consequence of this hypothesis, the various permitted energy levels appearing as analogues of the normal modes of a vibrating string. Electron diffraction experiments (Davisson and Germer, 1927) strikingly confirmed the existence of a wavelike aspect of matter by showing that *interference patterns could be obtained with material particles such as electrons.*

One therefore associates with a material particle of energy E and momentum $\mathbf{p}$, a wave whose angular frequency $\omega = 2\pi\nu$ and wave vector $\mathbf{k}$ are given by the same relations as for photons (*cf.* § A-1):

$$\begin{cases} E = h\nu = \hbar\omega \\ \mathbf{p} = \hbar\mathbf{k} \end{cases} \tag{B-2}$$

In other words, the corresponding wavelength is:

$$\lambda = \frac{2\pi}{|\mathbf{k}|} = \frac{h}{|\mathbf{p}|} \qquad \text{(de Broglie relation)} \tag{B-3}$$

Comment:

The very small value of the Planck constant h explains why the wavelike nature of matter is very difficult to demonstrate on a macroscopic scale. Complement A_I of this chapter discusses the orders of magnitude of the de Broglie wavelengths associated with various material particles.

B-2. Wave functions. Schrödinger equation

In accordance with de Broglie's hypothesis, we shall apply the ideas introduced in § A for the case of the photon to all material particles. Recalling the conclusions of this paragraph, we are led to the following formulation:

(*i*) For the classical concept of a position and momentum, we must substitute the concept of a time-varying *state*. The quantum state of a particle such as the electron[4] is characterized by a *wave function* $\psi(\mathbf{r}, t)$, which contains all the information it is possible to obtain about the particle.

(*ii*) $\psi(\mathbf{r}, t)$ is interpreted as a *probability amplitude of the particle's presence*. Since the possible positions of the particle form a continuum, the probability $\mathrm{d}\mathscr{P}(\mathbf{r}, t)$ of the particle being, at time t, in a volume element $\mathrm{d}^3r = \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z$ situated at the point $\mathbf{r}$ must be proportional to d^3r. It is therefore infinitesimal, and $|\psi(\mathbf{r}, t)|^2$ is interpreted as the corresponding *probability density*, with:

$$\mathrm{d}\mathscr{P}(\mathbf{r}, t) = C\ |\psi(\mathbf{r}, t)|^2\ \mathrm{d}^3r \tag{B-4}$$

where C is a normalization constant [see comment (*i*) at the end of § B-2].

(*iii*) *The principle of spectral decomposition* applies to the measurement of an arbitrary physical quantity:

– The result found must belong to a set of eigen results $\{a\}$.

– With each eigenvalue a is associated an eigenstate, that is, an eigenfunction $\psi_a(\mathbf{r})$. This function is such that, if $\psi(\mathbf{r}, t_0) = \psi_a(\mathbf{r})$ (where t_0 is the time at which the measurement is performed), the measurement will always yield a.

– For any $\psi(\mathbf{r}, t)$, the probability $\mathscr{P}_a$ of finding the eigenvalue a for a measurement at time t_0 is found by decomposing $\psi(\mathbf{r}, t_0)$ in terms of the functions $\psi_a(\mathbf{r})$:

$$\psi(\mathbf{r}, t_0) = \sum_a c_a\ \psi_a(\mathbf{r}) \tag{B-5}$$

Then:

$$\mathscr{P}_a = \frac{|c_a|^2}{\sum_a |c_a|^2} \tag{B-6}$$

[4] We shall not take into account here the existence of electron spin (*cf.* Chap. IX).

$\Big($the presence of the denominator ensures that the total probability is equal to unity: $\sum_a \mathscr{P}_a = 1\Big)$.

– If the measurement indeed yields a, the wave function of the particle immediately after the measurement is:

$$\psi'(\mathbf{r}, t_0) = \psi_a(\mathbf{r}) \tag{B-7}$$

(*iv*) The equation describing the evolution of the function $\psi(\mathbf{r}, t)$ remains to be written. It is possible to introduce it in a very natural way, using the Planck and de Broglie relations. Nevertheless, we have no intention of proving this fundamental equation, which is called the *Schrödinger equation*. We shall simply assume it. Later, we shall discuss some of its consequences (whose experimental verification will prove its validity). Besides, we shall consider this equation in much more detail in Chapter III.

When the particle (of mass m) is subjected to the influence of a potential[5] $V(\mathbf{r}, t)$, the Schrödinger equation takes on the form:

$$\boxed{i\hbar \frac{\partial}{\partial t}\, \psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m}\, \Delta\psi(\mathbf{r}, t) + V(\mathbf{r}, t)\, \psi(\mathbf{r}, t)} \tag{B-8}$$

where Δ is the Laplacian operator $\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$.
We notice immediately that this equation is linear and homogeneous in ψ. Consequently, for material particles, there exists a superposition principle which, combined with the interpretation of ψ as a probability amplitude, is the source of wavelike effects. Note, moreover, that the differential equation (B-8) is first-order with respect to time. This condition is necessary if the state of the particle at a time t_0, characterized by $\psi(\mathbf{r}, t_0)$, is to determine its subsequent state.

Thus there exists a fundamental analogy between matter and radiation: in both cases, a correct description of the phenomena necessitates the introduction of quantum concepts, and, in particular, the idea of wave-particle duality.

Comments:

(*i*) For a system composed of only one particle, the total probability of finding the particle anywhere in space, at time t, is equal to 1:

$$\int \mathrm{d}\mathscr{P}(\mathbf{r}, t) = 1 \tag{B-9}$$

Since $\mathrm{d}\mathscr{P}(\mathbf{r}, t)$ is given by formula (B-4), we conclude that the *wave function* $\psi(\mathbf{r}, t)$ *must be square-integrable*:

$$\int |\psi(\mathbf{r}, t)|^2\, \mathrm{d}^3 r \quad \text{is finite} \tag{B-10}$$

[5] $V(\mathbf{r}, t)$ designates a potential energy here. For example, it may be the product of an electric potential and the particle's charge. In quantum mechanics, $V(\mathbf{r}, t)$ is commonly called a potential.

The normalization constant C that appears in (B-4) is then given by the relation:

$$\frac{1}{C} = \int |\psi(\mathbf{r}, t)|^2 \, \mathrm{d}^3 r \tag{B-11}$$

(we shall later see that the form of the Schrödinger equation implies that C is time-independent). One often uses wave functions which are normalized, such that:

$$\int |\psi(\mathbf{r}, t)|^2 \, \mathrm{d}^3 r = 1 \tag{B-12}$$

The constant C is then equal to 1.

(*ii*) Note the important difference between the concepts of classical states and quantum states. The classical state of a particle is determined at time t by the specification of six parameters characterizing its position and its velocity at time t: x, y, z; v_x, v_y, v_z. The quantum state of a particle is determined by an *infinite number* of parameters: the values at the various points in space of the wave function $\psi(\mathbf{r}, t)$ which is associated with it. For the classical idea of a trajectory (the succession in time of the various states of the classical particle), we must substitute the idea of the propagation of the wave associated with the particle. Consider, for example, Young's double-slit experiment, previously described for the case of photons, but which in principle can also be performed with material particles such as electrons. When the interference pattern is observed, it makes no sense to ask through which slit each particle has passed, since the wave associated with it passed through both.

(*iii*) It is worth noting that, unlike photons, which can be emitted or absorbed during an experiment, material particles can neither be created nor destroyed. The electrons emitted by a heated filament for example already existed in the filament. In the same way, an electron absorbed by a counter does not disappear; it becomes part of an atom or an electric current. Actually, the theory of relativity shows that it is possible to create and annihilate material particles: for example, a photon having sufficient energy, passing near an atom, can materialize into an electron-positron pair. Inversely, the positron, when it collides with an electron, annihilates with it, emitting photons. However, we pointed out in the beginning of this chapter that we would limit ourselves here to the non-relativistic quantum domain, and we have indeed treated time and space coordinates asymmetrically. In the framework of non-relativistic quantum mechanics, material particles can neither be created nor annihilated. This conservation law, as we shall see, plays a role of primary importance. The need to abandon it is one of the important difficulties encountered when one tries to construct a relativistic quantum mechanics.

C. Quantum description of a particle. Wave packets

In the preceding paragraph, we introduced the fundamental concepts necessary for the quantum description of a particle. In this paragraph, we are going to familiarize ourselves with these concepts and deduce from them several very important properties. We start with the very simple case of a free particle.

C-1. Free particle

Consider a particle whose potential energy is zero (or has a constant value) at every point in space. The particle is thus not subjected to any force; it is said to be free.

When $V(\mathbf{r}, t) = 0$, the Schrödinger equation becomes:

$$i\hbar \frac{\partial}{\partial t}\psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}, t) \tag{C-1}$$

This differential equation is obviously satisfied by solutions of the form:

$$\psi(\mathbf{r}, t) = A\, \mathrm{e}^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{C-2}$$

(where A is a constant), on the condition that $\mathbf{k}$ and ω satisfy the relation:

$$\omega = \frac{\hbar\mathbf{k}^2}{2m} \tag{C-3}$$

Observe that, according to the de Broglie relations [see (B-2)], condition (C-3) expresses the fact that the energy E and the momentum $\mathbf{p}$ of a free particle satisfy the equation, which is well-known in classical mechanics:

$$E = \frac{\mathbf{p}^2}{2m} \tag{C-4}$$

We shall come back later (§ C-3) to the physical interpretation of a state of the form (C-2). We already see that, since

$$|\psi(\mathbf{r}, t)|^2 = |A|^2 \tag{C-5}$$

a plane wave of this type represents a particle whose probability of presence is uniform throughout all space (see comment below).

The principle of superposition tells us that every linear combination of plane waves satisfying (C-3) will also be a solution of equation (C-1). Such a superposition can be written:

$$\psi(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}}\int g(\mathbf{k})\, \mathrm{e}^{i[\mathbf{k}\cdot\mathbf{r}-\omega(k)t]}\, \mathrm{d}^3k \tag{C-6}$$

(d^3k represents, by definition, the infinitesimal volume element in $\mathbf{k}$-space: $\mathrm{d}k_x\, \mathrm{d}k_y\, \mathrm{d}k_z$). $g(\mathbf{k})$, which can be complex, must be sufficiently regular to allow differentiation inside the integral. It can be shown, moreover, that any square-integrable solution can be written in the form (C-6).

A wave function such as (C-6), a superposition of plane waves, is called a three-dimensional "wave packet". For the sake of simplicity, we shall often be led to study the case of a one-dimensional wave packet[6] of plane waves all propagating parallel to Ox. The wave function then depends only on x and t:

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} g(k)\, \mathrm{e}^{i[kx-\omega(k)t]}\, \mathrm{d}k \tag{C-7}$$

[6] A simple model of a two-dimensional wave packet is presented in Complement $\mathrm{E_I}$. Some general properties of three-dimensional wave packets are studied in Complement $\mathrm{F_I}$, which also shows how, in certain cases, a three-dimensional problem can be reduced to several one-dimensional problems.

In the following paragraph, we shall be interested in the form of the wave packet at a given instant. If we choose this instant as the time origin, the wave function is written:

$$\psi(x,0) = \frac{1}{\sqrt{2\pi}} \int g(k)\, \mathrm{e}^{ikx}\, \mathrm{d}k \tag{C-8}$$

We see that $g(k)$ is simply the Fourier transform (*cf.* Appendix I) of $\psi(x,0)$:

$$g(k) = \frac{1}{\sqrt{2\pi}} \int \psi(x,0)\, \mathrm{e}^{-ikx}\, \mathrm{d}x \tag{C-9}$$

Consequently, the validity of formula (C-8) is not limited to the case of the free particle: whatever the potential, $\psi(x,0)$ can always be written in this form. The consequences that we shall derive from this in §§ C-2 and C-3 below are thus perfectly general. It is not until § C-4 that we shall return explicitly to the free particle.

Comment:

A plane wave of type (C-2), whose modulus is constant throughout all space [*cf.* (C-5)], is not square-integrable. Therefore, rigorously, it cannot represent a physical state of the particle (in the same way as, in optics, a monochromatic plane wave is not physically realizable). On the other hand, a superposition of plane waves like (C-7) can be square-integrable.

C-2. Form of the wave packet at a given time

The form of the wave packet is given by the x-dependence of $\psi(x,0)$ defined by equation (C-8). Imagine that $|g(k)|$ has the shape depicted in Figure 3, with a pronounced peak situated at $k = k_0$ and a width (defined, for example, at half its maximum value) of Δk.

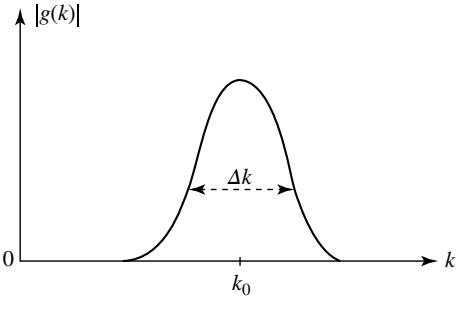


Figure 3: Shape of the function $|g(k)|$, modulus of the Fourier transform of $\psi(x,0)$: we assume that it is centered at $k = k_0$, where it reaches a maximum, and has a width of Δk.

Let us begin by trying to understand qualitatively the behavior of $\psi(x,0)$ through the study of a very simple special case. Let $\psi(x,0)$, instead of being the superposition of an infinite number of plane waves e^{ikx} as in formula (C-8), be the sum of only three plane waves. The wave vectors of these plane waves are k_0, $k_0 - \dfrac{\Delta k}{2}$, $k_0 + \dfrac{\Delta k}{2}$, and their

amplitudes are proportional, respectively, to 1, 1/2 and 1/2. We then have:

$$\psi(x) = \frac{g(k_0)}{\sqrt{2\pi}} \left[\mathrm{e}^{ik_0 x} + \frac{1}{2}\,\mathrm{e}^{i\left(k_0 - \frac{\Delta k}{2}\right)x} + \frac{1}{2}\,\mathrm{e}^{i\left(k_0 + \frac{\Delta k}{2}\right)x} \right]$$
$$= \frac{g(k_0)}{\sqrt{2\pi}}\ \mathrm{e}^{ik_0 x} \left[1 + \cos\left(\frac{\Delta k}{2} x \right) \right] \qquad \text{(C-10)}$$

We see that $|\psi(x)|$ is maximum when $x = 0$. This result is due to the fact that, when x takes on this value, the three waves are in phase and interfere constructively, as shown in Figure 4. As one moves away from the value $x = 0$, the waves become more and more out of phase, and $|\psi(x)|$ decreases. The interference becomes completely destructive when the phase shift between $\mathrm{e}^{ik_0 x}$ and $\mathrm{e}^{i(k_0 \mp \Delta k/2)x}$ is equal to $\pm\,\pi$: $\psi(x)$ goes to zero when $x = \pm\,\Delta x/2$, Δx being given by:

$$\Delta x\ \Delta k = 4\pi \qquad \text{(C-11)}$$

This formula shows that the smaller the width Δk of the function $|g(k)|$, the larger the width Δx of the function $|\psi(x)|$ (the distance between two zeros of $|\psi(x)|$).

Comment:

Formula (C-10) shows that $|\psi(x)|$ is periodic in x and therefore has a series of maxima and minima. This arises from the fact that $\psi(x)$ is the superposition of a finite number of waves (here, three). For a continuous superposition of an infinite number of waves, as in formula (C-8), such a phenomenon does not occur, and $|\psi(x, 0)|$ can have only one maximum.

Let us now return to the general wave packet of formula (C-8). Its form also results from an interference phenomenon: $|\psi(x, 0)|$ is maximum when the different plane waves interfere constructively.

Let $\alpha(k)$ be the argument of the function $g(k)$:

$$g(k) = \ |g(k)|\,\mathrm{e}^{i\alpha(k)} \qquad \text{(C-12)}$$

Assume that $\alpha(k)$ varies sufficiently smoothly within the interval $\left[k_0 - \dfrac{\Delta k}{2},\ k_0 + \dfrac{\Delta k}{2} \right]$ where $|g(k)|$ is appreciable; then, when Δk is sufficiently small, one can expand $\alpha(k)$ in the neighborhood of $k = k_0$:

$$\alpha(k) \simeq \alpha\,(k_0) + (k - k_0) \left[\frac{\mathrm{d}\alpha}{\mathrm{d}k} \right]_{k=k_0} \qquad \text{(C-13)}$$

which enables us to rewrite (C-8) in the form:

$$\psi(x, 0) \simeq \frac{\mathrm{e}^{i[k_0 x + \alpha(k_0)]}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |g(k)|\,\mathrm{e}^{i(k - k_0)(x - x_0)}\,\mathrm{d}k \qquad \text{(C-14)}$$

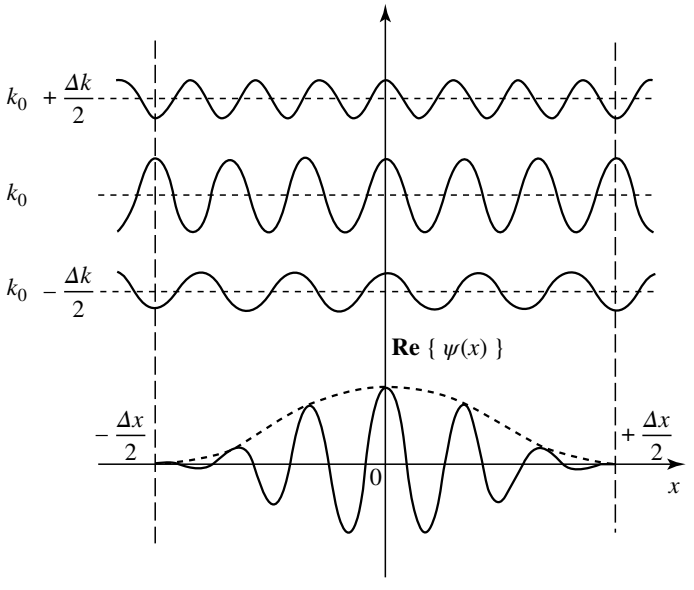


Figure 4: The real parts of the three waves whose sum gives the function $\psi(x)$ of (C-10). At $x = 0$, the three waves are in phase and interfere constructively. As one moves away from $x = 0$, they go out of phase and interfere destructively for $x = \pm\Delta x/2$.
In the lower part of the figure, $Re\{\psi(x)\}$ is shown. The dashed-line curve corresponds to the function $\left[1 + \cos\left(\frac{\Delta k}{2}x\right)\right]$, which, according to (C-10), gives $|\psi(x)|$ (and therefore, the form of the wave packet).

with:

$$x_0 = -\left[\frac{\mathrm{d}\alpha}{\mathrm{d}k}\right]_{k=k_0} \tag{C-15}$$

The form (C-14) is useful for studying the variations of $|\psi(x,0)|$ in terms of x. When $|x - x_0|$ is large, the function of k which is to be integrated oscillates a very large number of times within the interval Δk. We then see (*cf.* Fig. 5-a, in which the real part of this function is depicted) that the contributions of the successive oscillations cancel each other out, and the integral over k becomes negligible. In other words, when x is fixed at a value far from x_0, the phases of the various waves which make up $\psi(x,0)$ vary very rapidly in the domain Δk, and these waves destroy each other by interference. On the other hand, if $x \simeq x_0$, the function to be integrated over k oscillates hardly at all (*cf.* Fig. 5-b), and $|\psi(x,0)|$ is maximum.

The position $x_M(0)$ of the center of the wave packet is therefore:

$$x_M(0) = x_0 = -\left[\frac{\mathrm{d}\alpha}{\mathrm{d}k}\right]_{k=k_0} \tag{C-16}$$

Actually the result (C-16) can be obtained very easily. An integral such as the one appearing in (C-8) will be maximum (in absolute value) when the waves having the

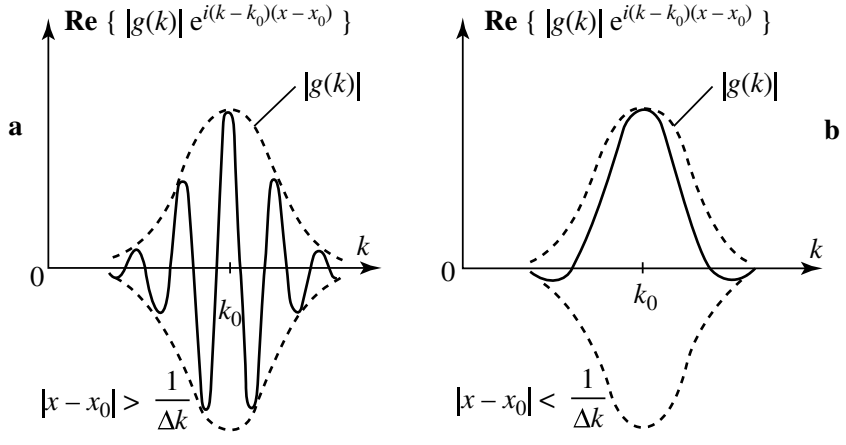


Figure 5: Variations with respect to k of the function to be integrated over k in order to obtain $\psi(x,0)$. In figure (a), x is fixed at a value such that $|x-x_0| > 1/\Delta k$, and the function to be integrated oscillates several times within the interval Δk. In figure (b), x is fixed such that $|x-x_0| < 1/\Delta k$, and the function to be integrated hardly oscillates, so that its integral over k takes on a relatively large value. Consequently, the center of the wave packet [point where $|\psi(x,0)|$ is maximum] is situated at $x = x_0$.

largest amplitude (those with k close to k_0) interfere constructively. This occurs when the k-dependent phases of these waves vary only slightly around $k = k_0$. To obtain the center of the wave packet, one then imposes (*stationary phase* condition) that the derivative with respect to k of the phase is zero for $k = k_0$. In the particular case which we are studying, the phase of the wave corresponding to k is $kx+\alpha(k)$. Therefore, $x_M(0)$ is that value of x for which the derivative $x + \mathrm{d}\alpha/\mathrm{d}k$ is zero at $k = k_0$.

When x moves away from the value x_0, $|\psi(x,0)|$ decreases. This decrease becomes appreciable if $\mathrm{e}^{i(k-k_0)(x-x_0)}$ oscillates approximately once when k traverses the domain Δk, that is, when:

$$\Delta k \;.\; (x - x_0) \simeq 1 \tag{C-17}$$

If Δx is the approximate width of the wave packet, we therefore have:

$$\Delta k \;.\; \Delta x \gtrsim 1 \tag{C-18}$$

We are thus brought back to a classical relation between the widths of two functions which are Fourier transforms of each other. The important fact is that the product $\Delta x \;.\; \Delta k$ has a lower bound; the exact value of this bound clearly depends on the precise definition of the widths Δx and Δk.

A wave packet such as (C-7) thus represents the state of a particle whose probability of presence, at the time $t = 0$, is practically zero outside an interval of approximate width Δx centered at the value x_0.

Comment:

The preceding argument could lead one to believe that the product $\Delta x \,.\, \Delta k$ is always of the order of 1 [*cf.* (C-17)]. Let us stress the fact that this is a lower limit. Although it is impossible to construct wave packets for which the product $\Delta x \,.\, \Delta k$ is negligible compared to 1, it is perfectly possible to construct packets for which this product is as large as desired [see, for example, Complement G_I, especially comment (*ii*) of § 3-c]. This is why (C-18) is written in the form of an inequality.

C-3. Heisenberg relations

In quantum mechanics, inequality (C-18) has extremely important physical consequences. We intend to discuss these now (we shall stay, for simplicity, within the framework of a one-dimensional model).

We have seen that a plane wave $e^{i(k_0 x - \omega_0 t)}$ corresponds to a constant probability density for the particle's presence along the Ox axis, for all values of t. This result can be roughly expressed by saying that the corresponding value of Δx is infinite. On the other hand, only one angular frequency ω_0 and one wave vector k_0 are involved. According to the de Broglie relations, this means that the energy and momentum of the particle are well-defined: $E = \hbar\omega_0$ and $p = \hbar k_0$. Such a plane wave can, moreover, be considered to be a special case of (C-7), for which $g(k)$ is a "delta function" (*cf.* Appendix II):

$$g(k) = \delta(k - k_0) \tag{C-19}$$

The corresponding value of Δk is then zero.

This property can also be interpreted in the following manner, using the principle of spectral decomposition (*cf.* §§ A-3 and B-1). To say that a particle, described at $t = 0$ by the wave function $\psi(x, 0) = A\, e^{ikx}$, has a well-determined momentum, is to say that a measurement of the momentum at this time will definitely yield $p = \hbar k$. From this we deduce that e^{ikx} *characterizes the eigenstate corresponding to* $p = \hbar k$. Since there exists a plane wave for every real value of k, the eigenvalues which one can expect to find in a measurement of the momentum on an arbitrary state include all real values. In this case, there is no quantization of the possible results: as in classical mechanics, all values of the momentum are allowed.

Now consider formula (C-8). In this formula, $\psi(x, 0)$ appears as a linear superposition of the momentum eigenfunctions in which the coefficient of e^{ikx} is $g(k)$. We are thus led to interpret $|g(k)|^2$ (to within a constant factor) as the probability of finding $p = \hbar k$ if one measures, at $t = 0$, the momentum of a particle whose state is described by $\psi(x, t)$. In reality, the possible values of p, like those of x, form a continuous set, and $|g(k)|^2$ is proportional to a *probability density*: the probability $\overline{\mathrm{d}\mathscr{P}}(k)$ of obtaining a value between $\hbar k$ and $\hbar(k + \mathrm{d}k)$ is, to within a constant factor, $|g(k)|^2\, \mathrm{d}k$. More precisely, if we rewrite formula (C-8) in the form:

$$\psi(x, 0) = \frac{1}{\sqrt{2\pi\hbar}} \int \overline{\psi}(p)\, e^{ipx/\hbar}\, \mathrm{d}p \tag{C-20}$$

we know that $\overline{\psi}(p)$ and $\psi(x, 0)$ satisfy the Parseval-Plancherel relation (*cf.* Appendix I):

$$\int_{-\infty}^{+\infty} |\psi(x,0)|^2 \, \mathrm{d}x = \int_{-\infty}^{+\infty} |\overline{\psi}(p)|^2 \, \mathrm{d}p \tag{C-21}$$

If the common value of these integrals is C, $\mathrm{d}\mathscr{P}(x) = \dfrac{1}{C} |\psi(x,0)|^2 \, \mathrm{d}x$ is the probability of the particle being found, at $t = 0$, between x and $x + \mathrm{d}x$. In the same way:

$$\overline{\mathrm{d}\mathscr{P}}(p) = \frac{1}{C} |\overline{\psi}(p)|^2 \, \mathrm{d}p \tag{C-22}$$

is the probability that the measurement of the momentum will yield a result included between p and $p + \mathrm{d}p$ [relation (C-21) then insures that the total probability of finding any value is indeed equal to 1].

Now let us go back to the inequality (C-18). We can write it as:

$$\Delta x \, . \, \Delta p \gtrsim \hbar \tag{C-23}$$

($\Delta p = \hbar \Delta k$ is the width of the curve representing $|\overline{\psi}(p)|$). Consider a particle whose state is defined by the wave packet (C-20). We know that its position probability at $t = 0$, is appreciable only within a region of width Δx about x_0: its position is known within an uncertainty Δx. If one measures the momentum of this particle at the same time, one will find a value between $p_0 + \dfrac{\Delta p}{2}$ and $p_0 - \dfrac{\Delta p}{2}$, since $|\overline{\psi}(p)|^2$ is practically zero outside this interval: the uncertainty in the momentum is therefore Δp. The interpretation of relation (C-23) is then as follows: it is impossible to define at a given time both the position of the particle and its momentum to an arbitrary degree of accuracy. When the lower limit imposed by (C-23) is reached, increasing the accuracy in the position (decreasing Δx) implies that the accuracy in the momentum diminishes (Δp increases), and vice versa. This relation is called the *Heisenberg relation* (or sometimes *uncertainty relation*, for historical reasons).

We know of nothing like this in classical mechanics. The limitation expressed by (C-23) arises from the fact that h is not zero. It is the very small value of h on the macroscopic scale which renders this limitation totally negligible in classical mechanics (an example is discussed in detail in Complement $\mathrm{B_I}$).

Comment:

The inequality (C-18) with which we started is not an inherently quantum mechanical principle. It merely expresses a general property of Fourier transforms, numerous applications of which can be found in classical physics. For example, it is well known from electromagnetic theory that there exists no packet of electromagnetic waves for which one can define the position and the wavelength with infinite accuracy at the same time. Quantum mechanics enters when one associates a wave with a material particle and requires that the wavelength and the momentum satisfy de Broglie's relation.

C-4. Time evolution of a free wave packet

Until now, we have been concerned only with the form of a wave packet at a given instant; in this paragraph, we are going to study its time evolution. Let us return,

therefore, to the case of a free particle whose state is described by the one-dimensional wave packet (C-7).

A given plane wave $e^{i(kx-\omega t)}$ propagates along the Ox axis with the velocity:

$$V_\varphi(k) = \frac{\omega}{k} \tag{C-24}$$

since it depends on x and t only through $\left(x - \frac{\omega}{k}\,t\right)$; $V_\varphi(k)$ is called the *phase velocity* of the plane wave.

We know that in the case of an electromagnetic wave propagating in a vacuum, V_φ is independent of k and equal to the speed of light c. All the waves making up a wave packet move at the same velocity, so that the packet as a whole also moves with the same velocity, without changing its shape. On the other hand, we know that this is not true in a dispersive medium, where the phase velocity is given by:

$$V_\varphi(k) = \frac{c}{n(k)} \tag{C-25}$$

$n(k)$ being the index of the medium, which varies with the wavelength.

The case that we are considering here corresponds to a dispersive medium, since the phase velocity is equal to [*cf.* equation (C-3)]:

$$V_\varphi(k) = \frac{\hbar k}{2m} \tag{C-26}$$

We shall see that when the different waves have unequal phase velocities, the velocity of the maximum x_M of the wave packet is not the average phase velocity $\frac{\omega_0}{k_0} = \frac{\hbar k_0}{2m}$, contrary to what one might expect.

As we did before, we shall begin by trying to understand qualitatively what happens, before taking a more general point of view. Therefore, let us return to the superposition of three waves considered in § C-2. For arbitrary t, $\psi(x,t)$ is given by:

$$\begin{aligned}\psi(x,t) &= \frac{g(k_0)}{\sqrt{2\pi}}\left\{e^{i[k_0x-\omega_0t]} + \frac{1}{2}\,e^{i\left[\left(k_0-\frac{\Delta k}{2}\right)x-\left(\omega_0-\frac{\Delta\omega}{2}\right)t\right]}\right.\\ &\qquad\left. + \frac{1}{2}\,e^{i\left[\left(k_0+\frac{\Delta k}{2}\right)x-\left(\omega_0+\frac{\Delta\omega}{2}\right)t\right]}\right\}\\ &= \frac{g(k_0)}{\sqrt{2\pi}}\,e^{i[k_0x-\omega_0t]}\left[1+\cos\left(\frac{\Delta k}{2}\,x - \frac{\Delta\omega}{2}\,t\right)\right]\end{aligned} \tag{C-27}$$

We see that the maximum of $|\psi(x,t)|$, which was at $x=0$ at $t=0$, is now at point:

$$x_M(t) = \frac{\Delta\omega}{\Delta k}\,t \tag{C-28}$$

and not at point $x = \frac{\omega_0}{k_0}\,t$. The physical origin of this result appears in Figure 6.

Part *a*) of this figure represents the position at time $t=0$ of three adjacent maxima (1), (2), (3), for the real parts of each of the three waves. Since the maxima denoted by the index (2) coincide at $x=0$, there is constructive interference at this point,

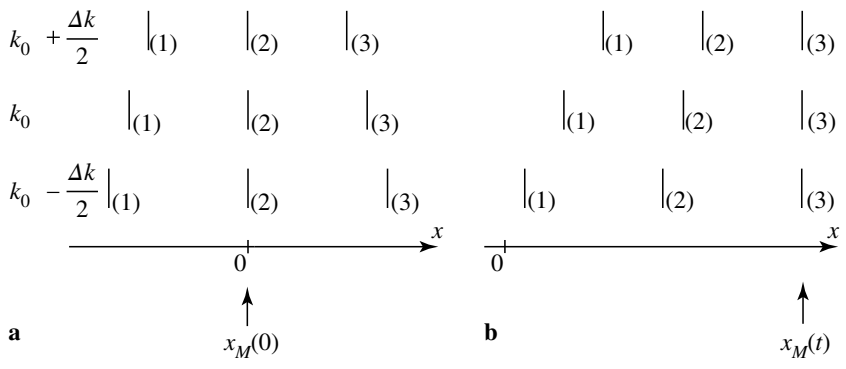


Figure 6: Positions of the maxima of the three waves of Figure 4 at time $t = 0$ (fig. a) and at a subsequent time t (fig. b). At time $t = 0$, it is the maxima (2), situated at $x = 0$, which interfere constructively: the position of the center of the wave packet is $x_M(0) = 0$. At time t, the three waves have advanced with different phase velocities V_φ. It is then the maxima (3) which interfere constructively and the center of the wave packet is situated at $x = x_M(t)$. We thus see that the velocity of the center of the wave packet (group velocity) is different from the phase velocities of the three waves.

which thus corresponds to the position of the maximum of $|\psi(x,0)|$. Since the phase velocity increases with k [formula (C-26)], the maximum (3) of the wave $\left(k_0 + \frac{\Delta k}{2}\right)$ will gradually catch up with that of the wave (k_0), which will in turn catch up with that of the wave $\left(k_0 - \frac{\Delta k}{2}\right)$. After a certain time, we shall thus have the situation shown in Figure 6-*b*: it will be the maxima (3) which coincide and thus determine the position of the maximum $x_M(t)$ of $|\psi(x,t)|$. We clearly see in the figure that $x_M(t)$ is not equal to $\frac{\omega_0}{k_0}\,t$, and a simple calculation again yields (C-28).

The shift of the center of the wave packet (C-7) can be found in an analogous fashion, by applying the "stationary phase" method. It can be seen from the form (C-7) of the free wave packet that, in order to go from $\psi(x,0)$ to $\psi(x,t)$, all we need to do is change $g(k)$ to $g(k)\,\mathrm{e}^{-i\omega(k)t}$. The reasoning of § C-2 thus remains valid, on the condition that we replace the argument $\alpha(k)$ of $g(k)$ by:

$$\alpha(k) - \omega(k)t \tag{C-29}$$

The condition (C-16) then gives:

$$x_M(t) = \left[\frac{\mathrm{d}\omega}{\mathrm{d}k}\right]_{k=k_0} t \;-\; \left[\frac{\mathrm{d}\alpha}{\mathrm{d}k}\right]_{k=k_0} \tag{C-30}$$

We are thus brought back to result (C-28): the velocity of the maximum of the wave packet is:

$$V_G(k_0) = \left[\frac{\mathrm{d}\omega}{\mathrm{d}k}\right]_{k=k_0} \tag{C-31}$$

$V_G(k_0)$ is called the *group velocity*indexGroup velocity of the wave packet. With the dispersion relation given in (C-3), we obtain:

$$V_G(k_0) = \frac{\hbar k_0}{m} = 2V_\varphi(k_0) \tag{C-32}$$

This result is important, for it enables us to retrieve the classical description of the free particle, for the cases where this description is valid. For example, when one is dealing with a macroscopic particle (and the example of the dust particle discussed in Complement B_I shows how small it can be), the uncertainty relation does not introduce an observable limit on the accuracy with which its position and momentum are known. This means that we can construct, in order to describe such a particle in a quantum mechanical way, a wave packet whose characteristic widths Δx and Δp are negligible. We would then speak, in classical terms, of the position $x_M(t)$ and the momentum p_0 of the particle. But then its velocity must be $v = \frac{p_0}{m}$. This is indeed what is implied by formula (C-32), obtained in the quantum description: in the cases where Δx and Δp can both be made negligible, the maximum of the wave packet moves like a particle that obeys the laws of classical mechanics.

Comment:

We have stressed here the motion of the center of the free wave packet. It is also possible to study the way in which its form evolves in time. It is then easy to show that, if the width Δp is a constant of the motion, Δx varies over time and, for sufficiently long times, increases without limit (spreading of the wave packet). The discussion of this phenomenon is given in Complement G_I, where the special case of a Gaussian wave packet is treated.

D. Particle in a time-independent scalar potential

We have seen, in § C, how the quantum mechanical description of a particle reduces to the classical description when Planck's constant h can be considered to be negligible. In the classical approximation, the wavelike character does not appear because the wavelength $\lambda = \frac{h}{p}$ associated with the particle is much smaller than the characteristic lengths of its motion. This situation is analogous to the one encountered in optics. Geometrical optics, which ignores the wavelike properties of light, constitutes a good approximation when the corresponding wavelength can be neglected compared to the lengths with which one is concerned. Classical mechanics thus plays, with respect to quantum mechanics, the same role played by geometrical optics with respect to wave optics.

In this paragraph, we are going to be concerned with a particle in a time-independent potential. What we have just said implies that typically quantum effects (that is, those of wave origin) should arise when the potential varies appreciably over distances shorter than the wavelength, which cannot then be neglected. This is why we are going to study the behavior of a quantum particle placed in various "square potentials", that is, "step potentials", as shown in Figure 7-a. Such a potential, which is discontinuous, clearly varies considerably over intervals of the order of the wavelength, however small it is: quantum effects must therefore always appear. Before beginning this investigation, we

shall discuss some important properties of the Schrödinger equation when the potential is not time-dependent.

D-1. Separation of variables. Stationary states

The wave function of a particle whose polential energy $V(\mathbf{r})$ is not time-dependent must satisfy the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\,\psi(\mathbf{r},t) = -\frac{\hbar^2}{2m}\,\Delta\psi(\mathbf{r},t) + V(\mathbf{r})\,\psi(\mathbf{r},t) \tag{D-1}$$

D-1-a. Existence of stationary states

Let us see if there exist solutions of this equation of the form:

$$\psi(\mathbf{r},t) = \varphi(\mathbf{r})\,\chi(t) \tag{D-2}$$

Substituting (D-2) into (D-1), we obtain:

$$i\hbar\,\varphi(\mathbf{r})\frac{\mathrm{d}\chi(t)}{\mathrm{d}t} = \chi(t)\left[-\frac{\hbar^2}{2m}\,\Delta\varphi(\mathbf{r})\right] + \chi(t)\,V(\mathbf{r})\,\varphi(\mathbf{r}) \tag{D-3}$$

If we divide both sides by the product $\varphi(\mathbf{r})\,\chi(t)$, we find:

$$\frac{i\hbar}{\chi(t)}\,\frac{\mathrm{d}\chi(t)}{\mathrm{d}t} = \frac{1}{\varphi(\mathbf{r})}\left[-\frac{\hbar^2}{2m}\,\Delta\varphi(\mathbf{r})\right] + V(\mathbf{r}) \tag{D-4}$$

This equation equates a function of t only (left-hand side) and a function of $\mathbf{r}$ only (right-hand side). This equality is only possible if each of these functions is in fact a constant, which we shall set equal to $\hbar\omega$, where ω has the dimensions of an angular frequency.

Setting the left-hand side equal to $\hbar\omega$, we obtain for $\chi(t)$ a differential equation which can easily be integrated to give:

$$\chi(t) = A\,\mathrm{e}^{-i\omega t} \tag{D-5}$$

In the same way, $\varphi(\mathbf{r})$ must satisfy the equation:

$$-\frac{\hbar^2}{2m}\,\Delta\varphi(\mathbf{r}) + V(\mathbf{r})\,\varphi(\mathbf{r}) = \hbar\omega\,\varphi(\mathbf{r}) \tag{D-6}$$

If we set $A = 1$ in equation (D-5) [which is possible if we incorporate, for example, the constant A in $\varphi(\mathbf{r})$], we achieve the following result: the function

$$\psi(\mathbf{r},t) = \varphi(\mathbf{r})\,\mathrm{e}^{-i\omega t} \tag{D-7}$$

is a solution of the Schrödinger equation, on the condition that $\varphi(\mathbf{r})$ is a solution of (D-6). The time and space variables are said to have been separated.

A wave function of the form (D-7) is called a *stationary solution of the Schrödinger equation*: it leads to a time-independent probability density $|\psi(\mathbf{r},t)|^2 = |\varphi(\mathbf{r})|^2$. In a stationary function, only one angular frequency ω appears; according to the Planck-Einstein relations, *a stationary state is a state with a well-defined energy* $E = \hbar\omega$ (energy eigenstate). In classical mechanics, when the potential energy is time-independent, the total

energy is a constant of the motion; in quantum mechanics, there exist well-determined energy states.

Equation (D-6) can therefore be written:

$$\left[-\frac{\hbar^2}{2m}\,\Delta + V(\mathbf{r})\right]\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}) \tag{D-8}$$

or:

$$\boxed{H\,\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r})} \tag{D-9}$$

where H is the differential operator:

$$\boxed{H = -\frac{\hbar^2}{2m}\,\Delta + V(\mathbf{r})} \tag{D-10}$$

H is a linear operator since, if λ_1 and λ_2 are constants, we have:

$$H\left[\lambda_1\varphi_1(\mathbf{r}) + \lambda_2\,\varphi_2(\mathbf{r})\right] = \lambda_1 H\varphi_1(\mathbf{r}) + \lambda_2 H\varphi_2(\mathbf{r}) \tag{D-11}$$

Equation (D-9) is thus the *eigenvalue equation* of the linear operator H: the application of H to the "eigenfunction" $\varphi(\mathbf{r})$ yields the same function, multiplied by the corresponding "eigenvalue" E. *The allowed energies are therefore the eigenvalues of the operator H*. We shall see later that equation (D-9) has square-integrable solutions $\varphi(\mathbf{r})$ only for certain values of E (*cf.* § D-2-c and § 2-c of Complement $\mathrm{H_I}$): this is the origin of *energy quantization*.

Comment:

Equation (D-8) [or (D-9)] is sometimes called the "time-independent Schrödinger equation", as opposed to the "time-dependent Schrödinger equation" (D-1). We stress their essential difference: equation (D-1) is a general equation which gives the evolution of the wave function, whatever the state of the particle; on the other hand, the eigenvalue equation (D-9) enables us to find, amongst all the possible states of the particle, those which are stationary.

D-1-b. Superposition of stationary states

In order to distinguish between the various possible values of the energy E (and the corresponding eigenfunctions $\varphi(\mathbf{r})$), we label them with an index n. Thus we have:

$$H\,\varphi_n(\mathbf{r}) = E_n\,\varphi_n(\mathbf{r}) \tag{D-12}$$

and the stationary states of the particle have as wave functions:

$$\psi_n(\mathbf{r}, t) = \varphi_n(\mathbf{r})\,\mathrm{e}^{-iE_n t/\hbar} \tag{D-13}$$

$\psi_n(\mathbf{r}, t)$ is a solution of the Schrödinger equation (D-1). Since this equation is linear, it has a whole series of other solutions of the form:

$$\psi(\mathbf{r}, t) = \sum_n c_n \varphi_n(\mathbf{r})\,\mathrm{e}^{-iE_n t/\hbar} \tag{D-14}$$

where the coefficients c_n are arbitrary complex constants. In particular, we have:

$$\psi(\mathbf{r},0) = \sum_n c_n \varphi_n(\mathbf{r}) \tag{D-15}$$

Inversely, assume that we know $\psi(\mathbf{r},0)$, the state of the particle at $t=0$. We shall see later that any function $\psi(\mathbf{r},0)$ can always be decomposed in terms of eigenfunctions of H, as in (D-15). The coefficients c_n are therefore determined by $\psi(\mathbf{r},0)$. The corresponding solution $\psi(\mathbf{r},t)$ of the Schrödinger equation is then given by (D-14). It is simply obtained by multiplying each term of (D-15) by the factor $\mathrm{e}^{-iE_nt/\hbar}$, where E_n is the eigenvalue associated with $\varphi_n(\mathbf{r})$. We stress the fact that these phase factors differ from one term to another. It is only in the case of stationary states that the t-dependence involves only one exponential [formula (D-13)].

D-2. One-dimensional "square" potentials. Qualitative study

We said at the beginning of § D that in order to display quantum effects we were going to consider potentials that varied considerably over small distances. We shall limit ourselves here to a qualitative study, so as to concentrate on the simple physical ideas. A more detailed study is presented in the complements of this chapter (Complement $\mathrm{H_I}$). To simplify the problem, we shall consider a one-dimensional model, in which the potential energy depends only on x (the justification for such a model is given in Complement $\mathrm{F_I}$).

D-2-a. Physical meaning of a square potential

We consider a one-dimensional problem with a potential of the type shown in Figure 7-a. The Ox axis is divided into a certain number of constant-potential regions. At the border of two adjacent regions the potential makes an abrupt jump (discontinuity). Actually, such a function cannot really represent a physical potential, which must be continuous. We use it to represent schematically a potential energy $V(x)$ which actually has the shape shown in Figure 7-b: there are no discontinuities, but $V(x)$ varies very rapidly in the neighborhood of certain values of x. When the intervals over which these variations occur are much smaller than all other distances involved in the problem (in particular, the wavelength associated with the particle), we can replace the true potential by the square potential of Figure 7-a. This is an approximation, which would cease to be valid, for example, for a particle having too high an energy, whose wavelength would be very short.

The predictions of classical mechanics concerning the behavior of a particle in a potential such as that of Figure 7 are easy to determine. For example, imagine that $V(x)$ is the gravitational potential energy. Figure 7-b then represents the real profile of the terrain on which the particle moves: the corresponding discontinuities are sharp slopes, separated by horizontal plateaus. Notice that, if we fix the total energy E of the particle, the domains of the Ox axis where $V > E$ are forbidden to it (its kinetic energy $E_c = E - V$ must be positive).

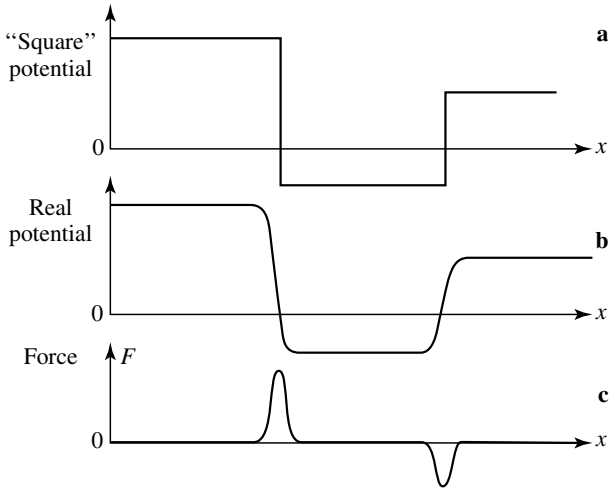


Figure 7: Square potential (fig. a) which schematically represents a real potential (fig. b) for which the force has the shape shown in figure c.

Comment:

The force exerted on the particle is $F(x) = -\dfrac{\mathrm{d}V(x)}{\mathrm{d}x}$. In Figure 7-c, we have depicted this force, obtained from the potential $V(x)$ of Figure 7-b. It can be seen that this particle, in all the regions where the potential is constant, is not subjected to any force. Its velocity is then constant. It is only in the frontier zones between these plateaus that a force acts on the particle and, depending on the case, accelerates it or slows it down.

D-2-b. Optical analogy

We are going to consider the stationary states (§ D-1) of a particle in a one-dimensional "square" potential.

In a region where the potential has a constant value V, the eigenvalue equation (D-9) is written:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V\right]\varphi(x) = E\varphi(x) \tag{D-16}$$

or:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{2m}{\hbar^2}(E - V)\right]\varphi(x) = 0 \tag{D-17}$$

Now, in optics, there exists a completely analogous equation. Consider a transparent medium whose index n depends neither on $\mathbf{r}$ nor on time. In this medium, there can be electromagnetic waves whose electric field $\mathbf{E}(\mathbf{r}, t)$ is independent of y and z and has the form:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{e}\, E(x)\, \mathrm{e}^{-i\Omega t} \tag{D-18}$$

where **e** is a unit vector perpendicular to Ox; $E(x)$ must then satisfy:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{n^2\Omega^2}{c^2}\right] E(x) = 0 \tag{D-19}$$

We see that equations (D-17) and (D-19) become identical if we set:

$$\frac{2m}{\hbar^2}(E - V) = \frac{n^2\Omega^2}{c^2} \tag{D-20}$$

Moreover, at a point x where the potential energy V [and, consequently, the index n given by (D-20)] is discontinuous, the boundary conditions for $\varphi(x)$ and $E(x)$ are the same: these two functions, as well as their first derivatives, must remain continuous (*cf.* Complement $\mathrm{H_I}$, § 1-b). The structural analogy between the two equations (D-17) and (D-19) thus enables us to associate with a quantum mechanical problem, corresponding to the potential of Figure 7-a, an optical problem: the propagation of an electromagnetic wave of angular frequency Ω in a medium whose index n has discontinuities of the same type. According to (D-20), the relation between the optical and mechanical parameters is:

$$n(\Omega) = \frac{1}{\hbar\Omega}\sqrt{2mc^2(E - V)} \tag{D-21}$$

For the light wave, a region where $E > V$ corresponds to a transparent medium whose index is real. The wave is then of the form e^{ikx}.

What happens when $V > E$? Formula (D-20) gives a pure imaginary index. In relation (D-19), n^2 is negative and the solution is of the form $\mathrm{e}^{-\rho x}$: it is the analogue of an "evanescent wave". Certain aspects of the situation recall the propagation of an electromagnetic wave in a metallic medium[7].

Thus we can transpose the well-known results of wave optics to the problems which we are studying here. It is important, however, to realize that this is merely an analogy. The interpretation that we give for the wave function is fundamentally different from that which classical wave optics attributes to the electromagnetic wave.

D-2-c. Examples

α. Potential step and barrier

Consider a particle of energy E which, coming from the region of negative x, arrives at the potential "step" of height V_0 shown in Figure 8.

If $E > V_0$ (the case in which the classical particle clears the potential step and continues towards the right with a smaller velocity), the optical analogy is the following: a light wave propagates from left to right in a medium of index n_1:

$$n_1 = \frac{c}{\hbar\Omega}\sqrt{2mE} \tag{D-22}$$

at $x = x_1$, there is a discontinuity, and the index, for $x > x_1$, is:

$$n_2 = \frac{c}{\hbar\Omega}\sqrt{2m(E - V_0)} \tag{D-23}$$

[7] This analogy should not be pushed too far, since the index n of a metallic medium has both a real and a complex part (in a metal, an optical wave continues to oscillate as it damps out).

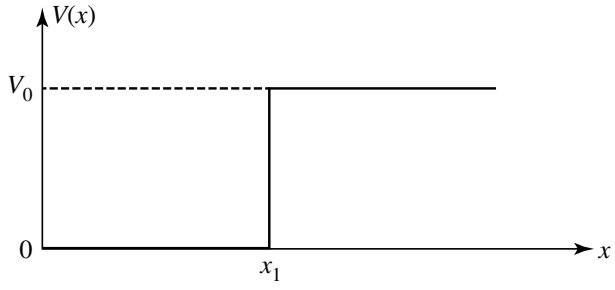


Figure 8: Potential step.

We know that the incident wave coming from the left splits into a reflected wave and a transmitted wave. Let us transpose this result to quantum mechanics: *the particle has a certain probability $\mathscr{P}$ of being reflected*, and only the probability $1 - \mathscr{P}$ of pursuing its course towards the right. This result is contrary to what is predicted by classical mechanics.

When $E < V_0$, the index n_2, which corresponds to the region $x > x_1$, becomes pure imaginary, and the incident light wave is totally reflected. The quantum prediction therefore coincides at this point with that of classical mechanics. Nevertheless, the existence, for $x > x_1$, of an evanescent wave, shows that the quantum particle has a non-zero probability of being found in this region.

The role of this evanescent wave is more striking in the case of a potential barrier (Fig. 9). For $E < V_0$, a classical particle would always turn back. But, in the corresponding optical problem, we would have a layer of finite thickness, with an imaginary index, surrounded by a transparent medium. If this thickness is not much greater than the range $1/\rho$ of the evanescent wave, part of the incident wave is transmitted into the region $x > x_2$. Therefore, even for $E < V_0$, we find a *nonzero probability of the particle crossing the barrier*. This is called the "tunnel effect".

β. *Potential well*

The function $V(x)$ now has the form shown in Figure 10. The predictions of classical mechanics are the following: when the particle has a negative energy E (but greater than $-V_0$), it can only oscillate between x_1 and x_2, with kinetic energy $E_c = E + V_0$; when the particle has a positive energy and arrives from the left, it undergoes an

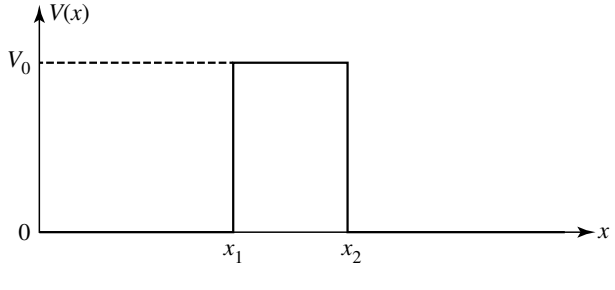


Figure 9: Potential barrier.

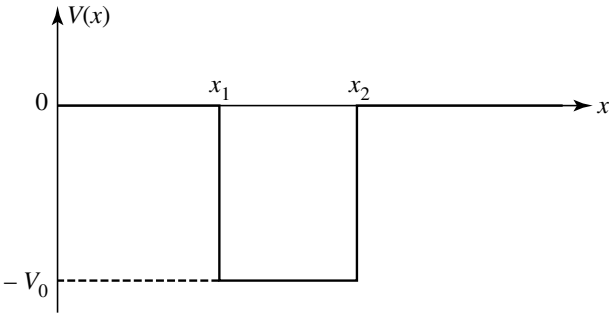


Figure 10: Potential well.

abrupt acceleration at x_1, then an equal deceleration at x_2, and then continues towards the right.

In the optical analogue of the case $-V_0 < E < 0$, the indices n_1 and n_3, which correspond to the regions $x < x_1$ and $x > x_2$, are imaginary, while the index n_2, which characterizes the interval $[x_1, x_2]$, is real. Thus we have the equivalent of a layer of air, for example, between two reflecting media. The different waves reflected successively at x_1 and x_2 destroy each other through interference, except for certain well-determined frequencies ("normal modes") that allow stable stationary waves to be established. From the quantum point of view, this implies that *the negative energies are quantized*[8], while, classically, all values included between $-V_0$ and 0 are possible.

For $E > 0$, the indices n_1, n_2 and n_3 are real:

$$n_1 = n_3 = \frac{c}{\Omega}\frac{1}{\hbar}\sqrt{2mE} \tag{D-24}$$

$$n_2 = \frac{c}{\Omega}\frac{1}{\hbar}\sqrt{2m(E+V_0)} \tag{D-25}$$

Since n_2 is greater than n_1 and n_3, the situation is analogous to that of a layer of glass in air. In order to obtain the reflected wave for $x < x_1$, or the transmitted wave in the region $x > x_2$, it is necessary to superpose an infinite number of waves that arise from successive reflections at x_1 and x_2 (multiple wave interferometer analogous to a Fabry-Perot). We then find that, for certain incident frequencies, the wave is entirely transmitted. From the quantum point of view, the particle thus has, in general, a certain probability of being reflected. However, there exist energy values, called *resonance energies*, for which the probability of transmission is 1 and, consequently, the probability of reflection is 0.

These few examples show how much the predictions of quantum mechanics can differ from those of classical mechanics. They also clearly stress the primordial role of potential discontinuities (which represent, schematically, rapid variations).

CONCLUSION

In this chapter we have introduced and discussed, in a qualitative and intuitive manner, certain fundamental ideas of quantum mechanics. We shall later return to

[8]The allowed energy values are not given by the well-known condition: $x_2 - x_1 = k\lambda_2/2$, for it is necessary to take into account the existence of evanescent waves, which introduce a phase shift upon reflection at $x = x_1$ and $x = x_2$ (*cf.* Complement H_I, § 2-c).

these ideas (Chap. III) so as to present them in a more precise and systematic way. Nevertheless, it is already clear that the quantum description of physical systems differs radically from the one given by classical mechanics (although the latter constitutes, in numerous cases, an excellent approximation). We have limited ourselves in this chapter to the case of physical systems composed of only one particle. The description of these systems at a given time is, in classical mechanics, founded on the specification of six parameters, which are the components of the position $\mathbf{r}(t)$ and the velocity $\mathbf{v}(t)$ of the particle. All the dynamical variables (energy, linear momentum, angular momentum) are determined by the specification of $\mathbf{r}(t)$ and $\mathbf{v}(t)$. Newton's laws enable us to calculate $\mathbf{r}(t)$ through the solution of second-order differential equations with respect to time. Consequently, they fix the values of $\mathbf{r}(t)$ and $\mathbf{v}(t)$ for any time t when they are known for the initial time.

Quantum mechanics uses a more complicated description of phenomena. The dynamic state of a particle, at a given time, is characterized by a wave function. It no longer depends on only six parameters, but on an infinite number [the values of $\psi(\mathbf{r}, t)$ at all points $\mathbf{r}$ of space]. Moreover, the predictions of the measurement results are now only probabilistic (they yield only the probability of obtaining a given result in the measurement of a dynamical variable). The wave function is a solution of the Schrödinger equation, which enables us to calculate $\psi(\mathbf{r}, t)$ from $\psi(\mathbf{r}, 0)$. This equation implies a principle of superposition which leads to wave effects.

This upheaval in our conception of mechanics was imposed by experiment. The structure and behavior of matter on an atomic level are incomprehensible in the framework of classical mechanics. The theory has thereby lost some of its simplicity, but it has gained a great deal of unity, since matter and radiation are described in terms of the same general scheme (wave-particle duality). We stress the fact that this general scheme, although it runs counter to our ideas and habits drawn from the study of the macroscopic domain, is perfectly consistent. No one has ever succeeded in imagining an experiment that could violate the uncertainty principle (*cf.* Complement D_I of this chapter). In general, no observation has, to date, contradicted the fundamental principles of quantum mechanics. Nevertheless, at present, there is no global theory including quantum phenomena within general relativity (gravity) and, of course, nothing prevents the possibility of a new upheaval.

References and suggestions for further reading:

Description of physical phenomena which demonstrate the necessity of introducing quantum mechanical concepts: see the subsection "Introductory work – quantum physics" of section 1 of the bibliography; in particular, Wichmann (1.1) and Feynman III (1.2), Chaps. 1 and 2.

History of the development of quantum mechanical concepts: references of section 4 of the bibliography, in particular, Jammer (4.8); also see references (5.11) and (5.12), which contain numerous references to the original articles.

Fundamental experiments: references to the original articles can be found in section 3 of the bibliography.

The problem of interpretation in quantum mechanics: section 5 of the bibliography; in particular, the "Resource Letter" (5.11), which contains many references.

Analogies and differences between matter waves and electromagnetic waves: Bohm (5.1), Chap. 4; in particular, the table "Summary on Probabilities" at the end of the chapter.

See also the articles by Schrödinger (1.25), Gamow (1.26), Born and Biem (1.28), Scully and Sargent (1.30).

A_I : ORDER OF MAGNITUDE OF THE WAVELENGTHS ASSOCIATED WITH MATERIAL PARTICLES **B_I : CONSTRAINTS IMPOSED BY THE UNCERTAINTY RELATION** **C_I : HEISENBERG RELATION AND ATOMIC PARAMETERS**	These three complements provide very simple but fundamental comments on the order of magnitude of quantum parameters.
D_I : AN EXPERIMENT ILLUSTRATING THE HEISENBERG RELATIONS	Discussion of a simple thought-experiment that attempts to invalidate the complementarity between the particle and the wave aspects of light (easy, but could be reserved for further study).
E_I : A SIMPLE TREATMENT OF A TWO-DIMENSIONAL WAVE PACKET **F_I : THE RELATION BETWEEN ONE- AND THREE-DIMENSIONAL PROBLEMS** **G_I : ONE-DIMENSIONAL GAUSSIAN WAVE PACKET: SPREADING OF THE WAVE PACKET**	Complements on wave packets (§ C of Chapter I): E_I : reveals in a simple, qualitative way the relation that exists between the lateral extension of a two-dimensional wave packet and the angular dispersion of the wave vectors (easy). F_I : generalization to three dimensions of the results of § C of Chapter I; shows how the study of a particle in three-dimensional space can, in certain cases, be reduced to one-dimensional problems (a little more difficult). G_I : treats in detail a special case of wave packets for which one can calculate exactly the properties and the evolution (some difficulties in the calculation, but conceptually simple).
H_I : STATIONARY STATES OF A PARTICLE IN ONE-DIMENSIONAL SQUARE POTENTIALS	Takes up in a more quantitative way the ideas of § D-2 of Chapter I. Strongly recommended, since square potentials are often used to illustrate simply the implications of quantum mechanics (numerous complements and excercizes proposed later in this book rely on the results of H_I).
J_I : BEHAVIOR OF A WAVE PACKET AT A POTENTIAL STEP	More precise study, for a special case, of the quantum behavior of a particle in a square potential. Since the particle is sufficiently well localized in space (wave packet), one can follow its "motion" (average difficulty; important for the physical interpretation of the results).
K_I : EXERCISES	

Complement A_I

Order of magnitude of the wavelengths associated with material particles

De Broglie's relation:

$$\lambda = \frac{h}{p} \tag{1}$$

shows that, for a particle of mass m and speed v, the smaller m and v, the longer the corresponding wavelength.

To show that the wave properties of matter are impossible to detect in the macroscopic domain, take as an example a dust particle, of diameter 1 μ and mass $m \simeq 10^{-15}$ kg. Even for such a small mass and a speed of $v \simeq 1$ mm/s, formula (1) gives:

$$\lambda \simeq \frac{6.6 \times 10^{-34}}{10^{-15} \times 10^{-3}} \text{ meter} = 6.6 \times 10^{-16} \text{ meter} = 6.6 \times 10^{-6} \text{ Å} \tag{2}$$

Such a wavelength is completely negligible on the scale of the dust particle.

Consider, on the other hand, a thermal neutron, that is, a neutron ($m_n \simeq 1.67 \times 10^{-27}$ kg) with a speed v corresponding to the average thermal energy at the (absolute) temperature T. v is given by the relation:

$$\frac{1}{2} m_n v^2 = \frac{p^2}{2m_n} \simeq \frac{3}{2} k_B T \tag{3}$$

where k_B is the Boltzmann constant ($k_B \simeq 1.38 \times 10^{-23}$ joule/degree). The wavelength corresponding to such a speed is:

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{3 m_n k_B T}} \tag{4}$$

For $T \simeq 300\,^\circ$K, we find:

$$\lambda \simeq 1.4 \text{ Å} \tag{5}$$

that is, a wavelength which is of the order of the distance between atoms in a crystal lattice. A beam of thermal neutrons falling on a crystal therefore gives rise to diffraction phenomena analogous to those observed with X-rays.

Let us now examine the order of magnitude of the de Broglie wavelengths associated with electrons ($m_e \simeq 0.9 \times 10^{-30}$ kg). If one accelerates an electron beam through a potential difference V (expressed in volts), one gives the electrons a kinetic energy:

$$E = qV = 1.6 \times 10^{-19}\, V \text{ joule} \tag{6}$$

($q = 1.6 \times 10^{-19}$ coulomb is the electron charge.) Since $E = \dfrac{p^2}{2m_e}$, the associated wavelength is equal to:

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2 m_e E}} \tag{7}$$

that is, numerically:

$$\lambda = \frac{6.6 \times 10^{-34}}{\sqrt{2 \times 0.9 \times 10^{-30} \times 1.6 \times 10^{-19}\, V}} \text{ meter} \simeq \frac{12.3}{\sqrt{V}} \text{ Å} \tag{8}$$

With potential differences of several hundreds of volts, one again obtains wavelengths comparable to those of X-rays, and electron diffraction phenomena can be observed with crystals or crystalline powders.

The large accelerators currently available are able to impart considerable energy to particles. This takes us out of the non-relativistic domain to which we have thus far confined ourselves. For example, electron beams are easily obtained for which the energy[1] exceeds 1 GeV = 10^9 eV (1 eV = 1 electron-volt = 16×10^{-19} joule), while the electron rest mass is equal to $m_e c^2 \simeq 05 \times 10^6$ eV. This means that the corresponding speed is very close to the speed of light c. Consequently, the non-relativistic quantum mechanics which we are studying here does not apply. However, the relations:

$$E = h\nu \tag{9-a}$$

$$\lambda = \frac{h}{p} \tag{9-b}$$

remain valid in the relativistic domain. On the other hand, relation (7) must be modified since, relativistically, the energy E of a particle of rest mass m_0 is no longer $p^2/2m_0$, but instead:

$$E = \sqrt{p^2 c^2 + m_0^2 c^4} \tag{10}$$

In the example considered above (an electron of energy 1 GeV), $m_e c^2$ is negligible compared to E, and we obtain:

$$\lambda \simeq \frac{hc}{E} = \frac{6.6 \times 10^{-34} \times 3 \times 10^8}{1.6 \times 10^{-10}} \text{ m} = 1.2 \times 10^{-15} \text{ m} = 1.2 \text{ fermi} \tag{11}$$

(1 fermi = 10^{-15} m). With electrons accelerated in this way, one can explore the structure of atomic nuclei and, in particular, the structure of the proton; nuclear dimensions are of the order of a fermi.

Comments:

(i) We want to point out a common error in the calculation of the wavelength of a material particle of mass $m_0 \neq 0$, whose energy E is known. This error consists of calculating the frequency ν using (9-a) and, then, by analogy with electromagnetic waves, of taking c/ν for the de Broglie wavelength. Obviously, the correct reasoning consists of calculating, for example from (10) (or, in the non-relativistic domain, from the relation $E = p^2/2m$) the momentum p associated with the energy E and then using (9-b) to find λ.

[1] Translator's note: In the United States, the unit Gev is sometimes written BeV.

(*ii*) According to (9-a), the frequency ν depends on the origin chosen for the energies. The same is true for the phase velocity $V_\varphi = \dfrac{\omega}{k} = \nu\lambda$. Note, on the other hand, that the group velocity $V_G = \dfrac{\mathrm{d}\omega}{\mathrm{d}k} = 2\pi\dfrac{\mathrm{d}\nu}{\mathrm{d}k}$ does not depend on the choice of the energy origin. This is important in the physical interpretation of V_G.

(*iii*) Strictly speaking, relation (1) would predict a wavelength that always tends to infinity when the velocity of the particle tends to zero, whatever its mass is. Let us nevertheless consider again the dust particle introduced above. Even when that particle has a velocity as low as $v \simeq 10^{-9}$ mm/s $= 10^{-6}\ \mu/s$, its wavelength is already of the order of its diameter (1 μ). It is obviously impossible to ensure that the velocity of a dust particle does not exceed such a low limit; therefore, even if it is at rest, its wavelength remains negligible with respect to its size. By contrast, for a particle such as the electron or the neutron, the quantum effects can appear much more easily.

References and suggestions for further reading:

Wichmann (1.1), Chap. 5; Eisberg and Resnick (1.3), § 3.1.

Complement B_I

Constraints imposed by the uncertainty relations

1 **Macroscopic system** **39**
2 **Microscopic system** **40**

We saw in § C-3 of Chapter I that the position and momentum of a particle cannot be simultaneously defined with arbitrary precision: the corresponding uncertainties Δx and Δp must satisfy the uncertainty relation:

$$\Delta x \cdot \Delta p \gtrsim \hbar \tag{1}$$

Here we intend to evaluate numerically the importance of this constraint. We shall show that it is completely negligible in the macroscopic domain and that it becomes, on the other hand, crucial on the microscopic level.

1. Macroscopic system

Let us take up again the example of a dust particle (*cf.* Complement A_I), whose diameter is on the order of 1 μ and whose mass $m \simeq 10^{-15}$ kg, having a speed $v = 10^{-3}$ m/sec. Its momentum is then equal to:

$$p = mv \simeq 10^{-18} \text{ joule sec/m} \tag{2}$$

If its position is measured to within $0.01\,\mu$, for example, the uncertainty Δp in the momentum must satisfy:

$$\Delta p \simeq \frac{\hbar}{\Delta x} \simeq \frac{10^{-34}}{10^{-8}} = 10^{-26} \text{ joule sec/m} \tag{3}$$

Thus the uncertainty relation introduces practically no restrictions in this case since, in practice, a momentum measurement device is incapable of attaining the required relative accuracy of 10^{-8}.

In quantum terms, the dust particle is described by a wave packet whose group velocity is $v = 10^{-3}$ m/sec and whose average momentum is $p = 10^{-18}$ joule sec/m. But one can then choose such a small spatial extension Δx and momentum dispersion Δp that they are both totally negligible. The maximum of the wave packet then represents *the* position of the dust particle, and its motion is identical to that of the classical particle.

2. Microscopic system

Now let us consider an atomic electron. The Bohr model describes it as a classical particle. The allowed orbits are defined by quantization rules which are assumed *a priori*: for example, the radius r of a circular orbit and the momentum $p = mv$ of the electron travelling in it must satisfy:

$$pr = n\hbar \tag{4}$$

where n is an integer.

For us to be able to speak in this way of an electron trajectory in classical terms, the uncertainties in its position and momentum must be negligible compared to r and p respectively:

$$\Delta x \ll r \tag{5-a}$$

$$\Delta p \ll p \tag{5-b}$$

which would mean that:

$$\frac{\Delta x}{r} \cdot \frac{\Delta p}{p} \ll 1 \tag{6}$$

Now the uncertainty relation imposes:

$$\frac{\Delta x}{r} \cdot \frac{\Delta p}{p} \gtrsim \frac{\hbar}{rp} \tag{7}$$

If we use formula (4) to replace rp by $n\hbar$ on the right-hand side, this inequality can be written as:

$$\frac{\Delta x}{r} \cdot \frac{\Delta p}{p} \gtrsim \frac{1}{n} \tag{8}$$

We then see that (8) is incompatible with (6), unless $n \gg 1$. The uncertainty relation thus makes us reject the semi-classical picture of the Bohr orbits (see § C-2 of Chapter VII).

References and suggestions for further reading:

Bohm (5.1), Chap. 5, § 14.

Complement C_I

Heisenberg relation and atomic parameters

The Bohr orbit has no physical reality when coupled with the Heisenberg relation (*cf.* Complement B_I). We shall study later (Chap. VII) the quantum theory of the hydrogen atom. However, we can already show how the Heisenberg uncertainty relation enables one to understand the stability of atoms and even to derive simply the order of magnitude of the dimensions and the energy of the hydrogen atom in its ground state.

Let us consider an electron in the Coulomb field of a proton, which we shall assume to be stationary at the origin of the coordinate system. When the two particles are separated by a distance r, the potential energy of the electron is:

$$V(r) = -\frac{q^2}{4\pi\varepsilon_0}\frac{1}{r} \tag{1}$$

where q is its charge (exactly opposite to that of the proton). We shall set:

$$\frac{q^2}{4\pi\varepsilon_0} = e^2 \tag{2}$$

Assume that the state of the electron is described by a spherically symmetric wave function whose spatial extent is characterized by r_0 (this means that the probability of presence is practically zero beyond $2r_0$ or $3r_0$). The potential energy corresponding to this state is then on the order of:

$$\overline{V} \simeq -\frac{e^2}{r_0} \tag{3}$$

For it to be as low as possible, it is necessary to take r_0 as small as possible. This means that the wave function must be as concentrated as possible about the proton.

But it is also necessary to take the kinetic energy into account. This is where the uncertainty principle comes in: if the electron is confined within a volume of linear dimension r_0, the uncertainty Δp in its momentum is at least of the order of $\hbar/r_0$. In other words, even if the average momentum is zero, the kinetic energy $\overline{T}$ associated with the state under consideration is not zero:

$$\overline{T} \gtrsim \overline{T}_{min} = \frac{1}{2m}(\Delta p)^2 \simeq \frac{\hbar^2}{2mr_0^2} \tag{4}$$

If we reduce r_0 in order to decrease the potential energy, the minimum kinetic energy (4) increases.

The lowest total energy compatible with the uncertainty relation is thus the minimum of the function:

$$E_{min} = \overline{T}_{min} + \overline{V} = \frac{\hbar^2}{2mr_0^2} - \frac{e^2}{r_0} \tag{5}$$

This minimum is obtained for:

$$r_0 = a_0 = \frac{\hbar^2}{me^2} \tag{6}$$

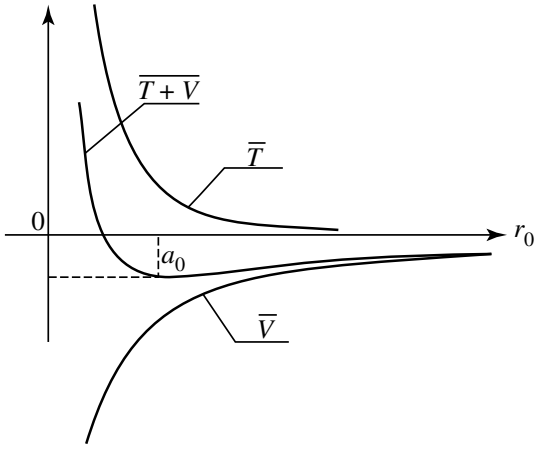


Figure 1: Variation with respect to r_0 (extension of the wave function) of the potential energy $\overline{V}$, the kinetic energy $\overline{T}$, and the total energy $\overline{T}+\overline{V}$ of a hydrogen atom. The functions $\overline{T}$ and $\overline{V}$ vary inversely, so the total energy passes through a minimum value for some value of $\overline{T}$ and $\overline{V}$. The corresponding value a_0 of r_0 gives the order of magnitude of the hydrogen atom's size.

and is equal to:

$$E_0 = -\frac{me^4}{2\hbar^2} \tag{7}$$

Expression (6) is the one found in the Bohr model for the radius of the first orbit, and (7) gives correctly the energy of the ground state of the hydrogen atom (see Chap. VII; the wave function of the ground state is indeed e^{-r/a_0}). Such quantitative agreement can only be accidental, since we have been reasoning on the basis of orders of magnitude. However, the preceding calculation reveals an important physical idea: because of the uncertainty relations, the smaller the extension of the wave function, the greater the kinetic energy of the electron. The ground state of the atom results from a compromise between the kinetic energy and the potential energy.

We stress the fact that this compromise, based on the uncertainty relations, is totally different from what would be expected in classical mechanics. If the electron moved in a classical circular orbit of radius r_0, its potential energy would be equal to:

$$V_{cl} = -\frac{e^2}{r_0} \tag{8}$$

The corresponding kinetic energy is obtained by equating the electrostatic force and the centrifugal force[1]:

$$\frac{e^2}{r_0^2} = m\frac{v^2}{r_0} \tag{9}$$

[1] In fact, the laws of classical electromagnetism indicate that an accelerated electron radiates, which already forbids the existence of stable orbits.

which gives:

$$T_{cl} = \frac{1}{2}mv^2 = \frac{1}{2}\frac{e^2}{r_0} \tag{10}$$

The total energy would then be equal to:

$$E_{cl} = T_{cl} + V_{cl} = -\frac{1}{2}\frac{e^2}{r_0} \tag{11}$$

The most favorable energetic situation would occur at $r_0 = 0$, which would give an infinite binding energy. Thus, we can say that it is the uncertainty relations that enables us to understand, as it were, the existence of atoms.

References and suggestions for further reading:

Feynman III (1.2), § 2-4. The same type of reasoning applied to molecules: Schiff (1.18), first section of § 49.

Complement D_I

An experiment illustrating the Heisenberg relations

Young's double-slit experiment, which we analyzed in § A-2 of Chapter I, led us to the following conclusions: both wave and particle aspects of light are needed to explain the observed phenomena; but they seem to be mutually exclusive, in the sense that *it is impossible to determine through which slit each photon has passed without destroying, by this very operation, the interference pattern.* The wave and particle aspects are sometimes said to be *complementary.*

We are going to consider Young's double-slit experiment again to demonstrate how complementarity and uncertainty relations are intimately related. To try to cast doubt on the uncertainty relation, one can imagine more subtle devices than the one of Chapter I, which used photomultipliers placed behind the slits. We shall now analyze one of these devices.

Assume that the plate $\mathscr{P}$, in which the slits are pierced, is mounted so that it can move vertically in the same plane. Thus, it is possible to measure the vertical momentum transferred to it. Consider (Fig. 1) a photon that strikes the observation screen $\mathscr{E}$ at point M (for simplicity, we choose a source $\mathscr{S}$ at infinity). The momentum of this photon changes when it crosses $\mathscr{P}$. Conservation of momentum implies that the plate $\mathscr{P}$ absorbs the difference. But the momentum thus transferred to $\mathscr{P}$ depends on the path of the photon; depending on whether it passed through F_1 or F_2, the vertical transferred momentum is equal to:

$$p_1 = -\frac{h\nu}{c}\sin\theta_1 \tag{1}$$

or:

$$p_2 = -\frac{h\nu}{c}\sin\theta_2 \tag{2}$$

$\Big($ $\frac{h\nu}{c}$ is the photon's momentum, θ_1 and θ_2 are the angles made by F_1M and F_2M with the incident direction. $\Big)$.

We then allow the photons to arrive one by one and gradually construct the interference pattern on the screen $\mathscr{E}$. For each one, we determine through which slit it has passed by measuring the momentum acquired by the plate $\mathscr{P}$. It therefore seems that interference phenomena can still be observed on $\mathscr{E}$ although we know through which slit each photon has passed.

Actually, we shall see that the interference fringes are not visible with this device. The error in the preceding argument consists of assuming that only the photons have a quantum character. In reality, it must not be forgotten that quantum mechanics also applies to the plate $\mathscr{P}$ (macroscopic object). If we want to know through which hole a photon has passed, the uncertainty Δp in the vertical momentum of $\mathscr{P}$ must be sufficiently small for us to be able to measure the difference between p_1 and p_2:

$$\Delta p \ll |\,p_2 - p_1|\tag{3}$$

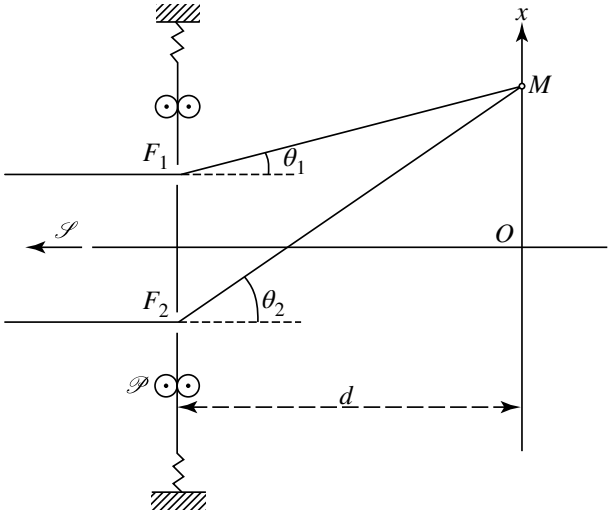


Figure 1: Diagram of a device using a movable plate $\mathscr{P}$ whose momentum is measured before and after the passage of the photon to determine whether the photon passed through F_1 or through F_2 before arriving at point M on the screen.

But then the uncertainty relation implies that the position of $\mathscr{P}$ is only known to within Δx, with:

$$\Delta x \gtrsim \frac{h}{|p_2 - p_1|} \tag{4}$$

If we designate by a the separation of the two slits and by d the distance between the plate $\mathscr{P}$ and the screen $\mathscr{E}$, and if we assume that θ_1 and θ_2 are small ($d/a \gg 1$), we find (Fig. 1):

$$\begin{aligned} \sin\theta_1 \simeq \theta_1 &\simeq \frac{x - a/2}{d} \\ \sin\theta_2 \simeq \theta_2 &\simeq \frac{x + a/2}{d} \end{aligned} \tag{5}$$

(x denotes the position of the point of impact M on $\mathscr{E}$). Formulas (1) and (2) then give:

$$|p_2 - p_1| \simeq \frac{h\nu}{c}\,|\theta_2 - \theta_1| \simeq \frac{h}{\lambda}\frac{a}{d} \tag{6}$$

where $\lambda = \dfrac{c}{\nu}$ is the wavelength of light. Substituting this value into formula (4), we obtain:

$$\Delta x \gtrsim \frac{\lambda d}{a} \tag{7}$$

But $\frac{\lambda d}{a}$ is precisely the fringe separation we expect to find on $\mathscr{E}$. If the vertical position of the slits F_1 and F_2 is defined only to within an uncertainty greater than the fringe separation, it is impossible to observe the interference pattern.

The preceding discussion clearly shows that it is impossible to construct a quantum theory that is valid for light and not for material systems without running into serious contradictions. Thus, in the above example, if we could treat the plate $\mathscr{P}$ as a classical material system, we could invalidate the complementarity of the two aspects of light, and, consequently, the quantum theory of radiation. Inversely, a quantum theory of matter alone would come up against analogous difficulties. In order to obtain an overall coherence, we must apply quantum ideas to all physical systems.

References and suggestions for further reading:

Bohm (5.1), Chaps. 5 and 6; Messiah (1.17), Chap. IV § III; Schiff (1.18), § 4; Jammer (5.12), Chaps. 4 and 5; also see reference (5.7).

Complement E_I

A simple treatment of a two-dimensional wave packet

1	**Introduction**	**49**
2	**Angular dispersion and lateral dimensions**	**49**
3	**Discussion**	**51**

1. Introduction

In § C-2 of Chapter I, we studied the shape of one-dimensional wave packets, obtained by superposing plane waves which all propagate in the same direction [formula (C-7)]. If this direction is that of the Ox axis, the resulting function is independent of y and z. It has a finite extension along Ox, but is not limited in the perpendicular directions: its value is the same at all points of a plane parallel to yOz.

We intend to examine here another simple type of wave packets: the plane waves which we are going to combine have coplanar wave vectors, which are (nearly) equal in magnitude but have slightly different directions. The goal is to show how the angular dispersion leads to a limitation of the wave packet in the directions perpendicular to the average wave vector.

We saw in § C-2 of Chapter I how, by studying the superposition of three specific waves of the one-dimensional packet, one can understand the most important aspects of the phenomena. In particular, this led to the fundamental relation (C-18) of that chapter. We are going to limit ourselves here to a simplified model of this type. The generalization of the results which we are going to find can be carried out in the same way as in Chapter I (see also Complement F_I).

2. Angular dispersion and lateral dimensions

Consider three plane waves, whose wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$ and $\mathbf{k}_3$ are shown in Figure 1. All three are in the xOy plane; $\mathbf{k}_1$ is directed along Ox; $\mathbf{k}_2$ and $\mathbf{k}_3$ are symmetric with respect to $\mathbf{k}_1$, the angle between each of them and $\mathbf{k}_1$ being $\Delta\theta$, which we assume to be small. Finally, the projections of $\mathbf{k}_1$, $\mathbf{k}_2$ and $\mathbf{k}_3$ on Ox are equal:

$$k_{1x} \simeq k_{2x} = k_{3x} \simeq |\mathbf{k}_1| = k \tag{1}$$

The magnitudes of these three vectors differ only by terms which are second order in $\Delta\theta$, which we shall neglect. Their components along the Oy axis are:

$$\begin{cases} k_{1y} = 0 \\ k_{2y} = -k_{3y} \simeq k\,\Delta\theta \end{cases} \tag{2}$$

We shall choose, as in § C-2 of Chapter I, real amplitudes $g(\mathbf{k})$ for waves, which satisfy

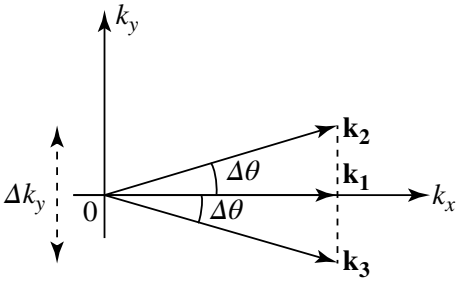


Figure 1: Arrangement of the wave vectors $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$ associated with three plane waves that will be superposed to construct a two-dimensional wave packet.

the relations:

$$g(\mathbf{k}_2) = g(\mathbf{k}_3) = \frac{1}{2} g(\mathbf{k}_1) \tag{3}$$

This model represents schematically a more complex situation, in which one would have a real wave packet, as in equation (C-6) of Chapter I, with the following characteristics: all the wave vectors are perpendicular to Oz and have the same projection on Ox (only the component along Oy varies); the function $|g(\mathbf{k})|$ has, with respect to this single variable k_y, the shape shown in Figure 2; its width Δk_y is related very simply to the angular dispersion $2\Delta\theta$:

$$\Delta k_y = 2k\Delta\theta \tag{4}$$

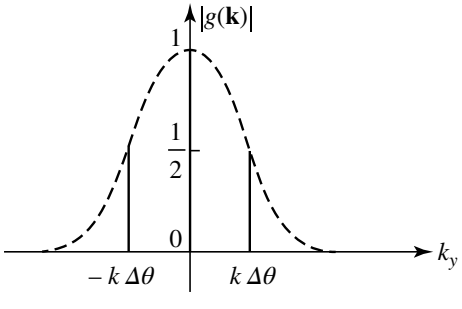


Figure 2: The three values chosen for k_y represent very schematically a peaked function $|g(\boldsymbol{k})|$ (dashed line).

The superposition of the three waves defined above yields:

$$\begin{aligned}\psi(x,y) &= \sum_{i=1}^{3} g(\mathbf{k}_i)\, \mathrm{e}^{i\mathbf{k}_i\cdot\mathbf{r}} \\ &= g(\mathbf{k}_1)\left[\mathrm{e}^{ikx} + \frac{1}{2}\,\mathrm{e}^{i(kx + k\,\Delta\theta\, y)} + \frac{1}{2}\mathrm{e}^{i(kx - k\,\Delta\theta\, y)}\right] \\ &= g(\mathbf{k}_1)\,\mathrm{e}^{ikx}\,[1 + \cos(k\,\Delta\theta\, y)]\end{aligned} \tag{5}$$

(there is no z-dependence, which is why this is called a two-dimensional wave packet). In order to understand what happens, we can use Figure 3, where we represent, for each of

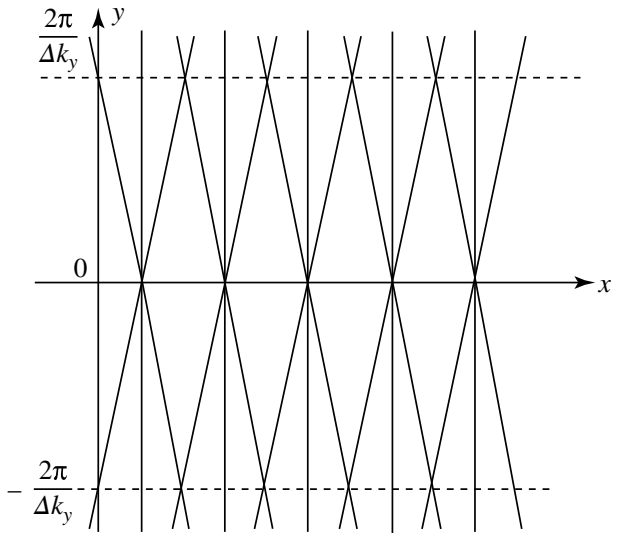


Figure 3: Equal phase planes of the three waves associated with the three $\boldsymbol{k}$ *vectors of Figure 1: these waves are in phase at* $y = 0$*, but interfere destructively at* $y = \pm 2\pi/\Delta k_y$*.*

the three components, the successive wave fronts corresponding to phase differences of 2π. The function $|\psi(x, y)|$ has a maximum at $y = 0$: the three waves interfere constructively on the Ox axis. When we move away from this axis, $|\psi(x, y)|$ decreases (the phase shift between the components increases) and goes to zero at $y = \pm \dfrac{\Delta y}{2}$, where Δy is given by:

$$\cos\left(k\,\Delta\theta\,\frac{\Delta y}{2}\right) = -1 \tag{6}$$

that is, for:

$$k\,\Delta\theta\,\Delta y = 2\pi \tag{7}$$

The phases of the $(\mathbf{k}_2)$ and $(\mathbf{k}_3)$ waves are then in opposition with that of the $(\mathbf{k}_1)$ wave (Fig. 3). Using (4), we can rewrite (7) in a form analogous to that of relation (C-11) of Chapter I:

$$\Delta y \cdot \Delta k_y = 4\pi \tag{8}$$

Thus an angular dispersion of the wave vectors limits the lateral dimensions of the wave packets. Quantitatively, this limitation has the form of an uncertainty relation [formulas (7) and (8)].

3. Discussion

Consider a plane wave with wave vector $\mathbf{k}$ propagating along Ox. Any attempt to limit its extension perpendicular to Ox causes an angular dispersion to appear, that is, transforms it into a wave packet analogous to the ones we are studying here.

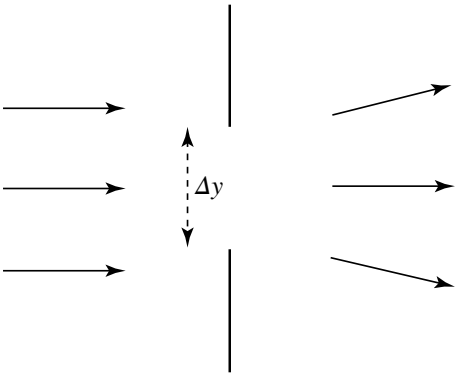


Figure 4: When the uncertainty Δy is decreased, the diffraction of the wave by the diaphragm increases the uncertainty Δk_y.

Imagine, for example, that we place in the path of the plane wave a screen pierced by a slit of width Δy. This will give rise to a diffracted wave (*cf.* Fig. 4). We know that the angular width of the diffraction pattern is given by:

$$2\,\Delta\theta \simeq 2\,\frac{\lambda}{\Delta y} \tag{9}$$

where $\lambda = \dfrac{2\pi}{|\mathbf{k}|}$ is the incident wavelength. This is indeed the same situation as above: formulas (7) and (9) are identical.

Complement F_I

The relationship between one- and three-dimensional problems

1 Three-dimensional wave packet 53
1-a Simple case 53
1-b General case 54
2 Justification of one-dimensional models 56

The space in which a classical or quantum particle moves is, of course, three-dimensional. This is why we wrote the Schrödinger equation (D-1) in Chapter I for a wave function $\psi(\mathbf{r})$ that depends on the three components x, y, z of $\mathbf{r}$. Nevertheless, we have repeatedly used in this chapter a one-dimensional model, in which only the x-variable is considered, without justifying this model in a very precise way. Therefore, this complement has two purposes: first (§ 1), to generalize to three dimensions the results given in § C of Chapter I; then (§ 2), to show how one can, in certain cases, rigorously justify the one-dimensional model.

1. Three-dimensional wave packet

1-a. Simple case

Let us begin by considering a very simple case, for which the following two hypotheses are satisfied:

– the wave packet is free [$V(\mathbf{r}) \equiv 0$]; it can therefore be written as in equation (C-6) of Chapter I:

$$\psi(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}} \int g(\mathbf{k})\, \mathrm{e}^{i[\mathbf{k}\cdot\mathbf{r}-\omega(\mathbf{k})t]}\, \mathrm{d}^3k \tag{1}$$

– moreover, the function $g(\mathbf{k})$ is of the form:

$$g(\mathbf{k}) = g_1(k_x) \times g_2(k_y) \times g_3(k_z) \tag{2}$$

Recall the expression for $\omega(\mathbf{k})$ in terms of $\mathbf{k}$:

$$\omega(\mathbf{k}) = \frac{\hbar\mathbf{k}^2}{2m} = \frac{\hbar}{2m}(k_x^2 + k_y^2 + k_z^2) \tag{3}$$

Substitute (2) and (3) into (1). It is possible to separate the three integrations with respect to k_x, k_y and k_z to obtain:

$$\psi(\mathbf{r},t) = \psi_1(x,t) \times \psi_2(y,t) \times \psi_3(z,t) \tag{4}$$

with:

$$\begin{cases} \psi_1(x,t) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{+\infty} g_1(k_x)\, \mathrm{e}^{i[k_x \cdot x - \omega(k_x)t]}\, \mathrm{d}k_x \\ \\ \omega(k_x) = \dfrac{\hbar k_x^2}{2m} \end{cases} \tag{5}$$

and analogous expressions for $\psi_2(y,t)$ and $\psi_3(z,t)$.

$\psi_1(x,t)$ indeed has the form of a one-dimensional wave packet. In this particular case, $\psi(\mathbf{r},t)$ is thus obtained simply by taking the product (4) of three one-dimensional wave packets, each evolving in a totally independent way.

1-b. General case

In the general case, where the potential $V(\mathbf{r})$ is arbitrary, formula (1) is no longer valid. It is then useful to introduce the three-dimensional Fourier transform $g(\mathbf{k},t)$ of the function $\psi(\mathbf{r},t)$ by writing:

$$\psi(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}} \int g(\mathbf{k},t)\, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}\, \mathrm{d}^3k \tag{6}$$

A priori, the t-dependence of $g(\mathbf{k},t)$, which brings in $V(\mathbf{r})$, is arbitrary. Moreover, there is no reason in general why we should be able to express $g(\mathbf{k},t)$ in the form of a product, as in (2). In order to generalize the results of § C-2 of Chapter I, we make the following hypothesis about its $\mathbf{k}$-dependence: $|g(\mathbf{k},t)|$ is (at a given time t) a function that has a very pronounced peak for values of $\mathbf{k}$ close to $\mathbf{k}_0$ and takes on a negligible value when the tip of $\mathbf{k}$ leaves a domain D_k centered at $\mathbf{k}_0$ and of dimensions Δk_x, Δk_y, Δk_z. As above, we set:

$$g(\mathbf{k},t) = |g(\mathbf{k},t)|\, \mathrm{e}^{i\alpha(\mathbf{k},t)} \tag{7}$$

so that the phase of the wave defined by the vector $\mathbf{k}$ can be written:

$$\xi(\mathbf{k},\mathbf{r},t) = \alpha(\mathbf{k},t) + k_x \cdot x + k_y \cdot y + k_z \cdot z \tag{8}$$

We can set forth an argument similar to that of § C-2 of Chapter I. First of all, the wave packet attains a maximum when all the waves, for which the tip of $\mathbf{k}$ is in D_k, are practically in phase, that is, when ξ varies very little within D_k. In general, $\xi(\mathbf{k},\mathbf{r},t)$ can be expanded about $\mathbf{k}_0$. Its variation between $\mathbf{k}_0$ and $\mathbf{k}$ is, to first order in $\delta\mathbf{k} = \mathbf{k} - \mathbf{k}_0$:

$$\begin{aligned} \delta\xi(\mathbf{k},\mathbf{r},t) \simeq \delta k_x \left[\frac{\partial}{\partial k_x}\, \xi(\mathbf{k},\mathbf{r},t)\right]_{\mathbf{k}=\mathbf{k}_0} + \delta k_y \left[\frac{\partial}{\partial k_y}\, \xi(\mathbf{k},\mathbf{r},t)\right]_{\mathbf{k}=\mathbf{k}_0} \\ + \delta k_z \left[\frac{\partial}{\partial k_z}\, \xi(\mathbf{k},\mathbf{r},t)\right]_{\mathbf{k}=\mathbf{k}_0} \end{aligned} \tag{9}$$

that is, more concisely[1], using (8):

$$\begin{aligned} \delta\xi(\mathbf{k},\mathbf{r},t) &\simeq \delta\mathbf{k} \cdot [\nabla_{\mathbf{k}}\, \xi(\mathbf{k},\mathbf{r},t)]_{\mathbf{k}=\mathbf{k}_0} \\ &\simeq \delta\mathbf{k} \cdot \left[\mathbf{r} + [\nabla_{\mathbf{k}}\, \alpha(\mathbf{k},t)]_{\mathbf{k}=\mathbf{k}_0}\right] \end{aligned} \tag{10}$$

[1] The symbol ∇ designates a "gradient": by definition, $\nabla f(x,y,z)$ is the vector whose coordinates are $\partial f/\partial x$, $\partial f/\partial y$, $\partial f/\partial z$. The index $\mathbf{k}$ in $\nabla_{\mathbf{k}}$ means that, as in (9), the differentiations must be performed with respect to the variables k_x, k_y and k_z.

We see from (10) that the variation of $\xi(\mathbf{k}, \mathbf{r}, t)$ within the domain D_k will be minimal for:

$$\mathbf{r} = \mathbf{r}_M(t) = -\left[\nabla_{\mathbf{k}}\, \alpha(\mathbf{k}, t)\right]_{\mathbf{k}=\mathbf{k}_0} \tag{11}$$

We have seen that, under these conditions, $|\psi(\mathbf{r}, t)|$ is maximum. Relation (11) therefore defines the position $\mathbf{r}_M(t)$ of the center of the wave packet and constitutes the generalization to three dimensions of equation (C-15) of Chapter I.

In what domain D_r, centered at $\mathbf{r}_M$ and of dimensions Δx, Δy, Δz, does the wave packet (6) take on non-negligible values? $|\psi(\mathbf{r}, t)|$ becomes much smaller than $|\psi(\mathbf{r}_M, t)|$ when the various $\mathbf{k}$ waves destroy each other by interference, that is, when the variation of $\xi(\mathbf{k}, \mathbf{r}, t)$ within the domain D_k is of the order of 2π (or roughly, of the order of 1 radian). Set $\delta\mathbf{r} = \mathbf{r} - \mathbf{r}_M$; if (11) is taken into account, relation (10) can be written:

$$\delta\xi(\mathbf{k}, \mathbf{r}, t) \simeq \delta\mathbf{k} \cdot \delta\mathbf{r} \tag{12}$$

The condition $\delta\xi(\mathbf{k}, \mathbf{r}, t) \gtrsim 1$ immediately gives us the relations which exist between the dimensions of D_r and those of D_k:

$$\begin{cases} \Delta x \;.\; \Delta k_x \gtrsim 1 \\ \Delta y \;.\; \Delta k_y \gtrsim 1 \\ \Delta z \;.\; \Delta k_z \gtrsim 1 \end{cases} \tag{13}$$

The Heisenberg relations then follow directly from the relation $\mathbf{p} = \hbar\mathbf{k}$:

$$\begin{aligned} &\Delta x \;.\; \Delta p_x \gtrsim \hbar \\ &\Delta y \;.\; \Delta p_y \gtrsim \hbar \\ &\Delta z \;.\; \Delta p_z \gtrsim \hbar \end{aligned} \tag{14}$$

These inequalities constitute the generalization to three dimensions of (C-23) of Chapter I.

Finally, note that the group velocity $\mathbf{V}_G$ of the wave packet can be obtained by differentiating (11) with respect to t:

$$\mathbf{V}_G = -\frac{\mathrm{d}}{\mathrm{d}t}\left[\nabla_{\mathbf{k}}\, \alpha(\mathbf{k}, t)\right]_{\mathbf{k}=\mathbf{k}_0} \tag{15}$$

In the special case of a free wave packet which does not, however, necessarily satisfy (2), we have:

$$\alpha(\mathbf{k}, t) = \alpha(\mathbf{k}, 0) - \omega(\mathbf{k})\, t \tag{16}$$

where $\omega(\mathbf{k})$ is given by (3). Formula (15) then gives:

$$\mathbf{V}_G = \left[\nabla_{\mathbf{k}}\, \omega(\mathbf{k})\right]_{\mathbf{k}=\mathbf{k}_0} = \frac{\hbar\mathbf{k}_0}{m} \tag{17}$$

which is the generalization of equation (C-31) of Chapter I.

2. Justification of one-dimensional models

When the potential is time-independent, we saw in § D-1 of Chapter I that it is possible to separate the time and space variables in the Schrödinger equation. This leads to the eigenvalue equation (D-8). We intend to show here how it is possible, in certain cases, to extend this method further and to separate as well the x, y, z variables in (D-8).

Assume that the potential energy $V(\mathbf{r})$ can be written:

$$V(\mathbf{r}) = V(x, y, z) = V_1(x) + V_2(y) + V_3(z) \tag{18}$$

and let us see if there exist solutions of the eigenvalue equation of the form:

$$\varphi(x, y, z) = \varphi_1(x) \times \varphi_2(y) \times \varphi_3(z) \tag{19}$$

An argument analogous to the one set forth in Chapter I (§ D-1-a) shows that this is possible if:

$$\left[-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2}{\mathrm{d}x^2} + V_1(x) \right] \varphi_1(x) = E_1 \, \varphi_1(x) \tag{20}$$

and if we have two other similar equations where x is replaced by y (or z), V_1 by V_2 (or V_3), and E_1 by E_2 (or E_3). In addition, it is also necessary that the relation:

$$E = E_1 + E_2 + E_3 \tag{21}$$

be satisfied.

Equation (20) is of the same type as (D-8), but in one dimension. The x, y and z variables are separated[2].

What happens, for example, if the potential energy $V(\mathbf{r})$ of a particle depends only on x? $V(\mathbf{r})$ can then be written in the form (18), where $V_1 = V$ and $V_2 = V_3 = 0$. Equations (20) in y and z correspond to the case already studied, in § C-1 of Chapter I, of the free particle in one dimension; their solutions are plane waves $\mathrm{e}^{ik_y \cdot y}$ and $\mathrm{e}^{ik_z \cdot z}$. All that remains is to solve equation (20), which amounts to considering a problem in only one dimension; nevertheless, the total energy of the particle in three dimensions is now:

$$E = E_1 + \frac{\hbar^2}{2m} \left[k_y^2 + k_z^2 \right] \tag{22}$$

The one-dimensional models studied in Chapter I thus actually correspond to a particle in three dimensions moving in a potential $V(\mathbf{r})$ that depends only on x. The solutions $\varphi_2(y)$ and $\varphi_3(z)$ are then very simple and correspond to particles which are "free along Oy" or along Oz. This is why we have concentrated all our attention on the study of the x-equation.

[2] It can be shown (*cf.* Chap. II, § F-4-a-β) that, when $V(\mathbf{r})$ has the form (18), all the solutions of the eigenvalue equation (D-8) are linear combinations of those we find here.

Complement G_I

One-dimensional Gaussian wave packet: spreading of the wave packet

1 Definition of a Gaussian wave packet 57
2 Calculation of Δx and Δp; uncertainty relation 58
3 Evolution of the wave packet 59
3-a Calculation of $\psi(x,t)$ 59
3-b Velocity of the wave packet 60
3-c Spreading of the wave packet 60

In this complement, we intend to study a particular (one-dimensional) free wave packet, for which the function $g(k)$ is Gaussian. The reason why this example is interesting lies in the fact that the calculations can be carried out exactly and to the very end. Thus, we can first verify, in this special case, the various properties of wave packets which we pointed out in § C of Chapter I. We shall then use these properties to study the variation in time of the width of this wave packet, which will reveal the phenomenon of spreading over time.

1. Definition of a Gaussian wave packet

Consider, in a one-dimensional model, a free particle [$V(x) \equiv 0$] whose wave function at time $t = 0$ is:

$$\psi(x,0) = \frac{\sqrt{a}}{(2\pi)^{3/4}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{a^2}{4}(k-k_0)^2}\, \mathrm{e}^{ikx}\, \mathrm{d}k \tag{1}$$

This wave packet is obtained by superposing plane waves e^{ikx} with the coefficients:

$$\frac{1}{\sqrt{2\pi}} g(k,0) = \frac{\sqrt{a}}{(2\pi)^{3/4}}\, \mathrm{e}^{-\frac{a^2}{4}(k-k_0)^2} \tag{2}$$

which correspond to a Gaussian function centered at $k = k_0$ (and multiplied by a numerical coefficient which normalizes the wave function). This is why the wave packet (1) is called Gaussian.

In the calculations that follow, we shall repeatedly come upon integrals of the type:

$$I(\alpha,\beta) = \int_{-\infty}^{+\infty} \mathrm{e}^{-\alpha^2(\xi+\beta)^2}\, \mathrm{d}\xi \tag{3}$$

where α and β are complex numbers [note that, for the integral (3) to converge, we must have $\mathrm{Re}\ \alpha^2 > 0$]. The method of residues enables us to show that this integral does not

depend on β:

$$I(\alpha, \beta) = I(\alpha, 0) \tag{4}$$

and that, when the condition $-\pi/4 < \text{Arg } \alpha < +\pi/4$ is fulfilled (which is always possible if Re $\alpha^2 > 0$), $I(\alpha, 0)$ is given by:

$$I(\alpha, 0) = \frac{1}{\alpha} I(1, 0) \tag{5}$$

Now all that remains is to evaluate $I(1, 0)$, which can be done classically, through a double integration in the xOy plane and a change into polar coordinates:

$$I(1, 0) = \int_{-\infty}^{+\infty} \mathrm{e}^{-\xi^2} \, \mathrm{d}\xi = \sqrt{\pi} \tag{6}$$

Thus we have:

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-\alpha^2(\xi+\beta)^2} \, \mathrm{d}\xi = \frac{\sqrt{\pi}}{\alpha} \tag{7}$$

with: $-\pi/4 < \text{Arg } \alpha < +\pi/4$.

Let us now calculate $\psi(x, 0)$. To do this, let us group, in the exponents of (1), the k-dependent terms into a perfect square, by writing them in the form:

$$-\frac{a^2}{4}(k - k_0)^2 + ikx = -\frac{a^2}{4}\left[k - k_0 - \frac{2ix}{a^2}\right]^2 + ik_0 x - \frac{x^2}{a^2} \tag{8}$$

We can then use (7), which yields:

$$\psi(x, 0) = \left(\frac{2}{\pi a^2}\right)^{1/4} \mathrm{e}^{ik_0 x} \, \mathrm{e}^{-x^2/a^2} \tag{9}$$

We find, as could be expected, that the Fourier transform of a Gaussian function is also a Gaussian (*cf.* Appendix I).

At time $t = 0$, the probability density of the particle is therefore given by:

$$|\psi(x, 0)|^2 = \sqrt{\frac{2}{\pi a^2}} \, \mathrm{e}^{-2x^2/a^2} \tag{10}$$

The curve representing $|\psi(x, 0)|^2$ is the familiar bell-shaped curve. The center of the wave packet [the maximum of $|\psi(x, 0)|^2$] is situated at the point $x = 0$. This is indeed what we could have found if we had applied the general formula (C-16) of Chapter I since, in this particular case, the function $g(k)$ is real.

2. Calculation of Δx and Δp; uncertainty relation

It is convenient, when one is studying a Gaussian function $f(x) = \mathrm{e}^{-x^2/b^2}$, to define its width Δx by:

$$\Delta x = \frac{b}{\sqrt{2}} \tag{11}$$

When x varies from 0 to $\pm\,\Delta x$, $f(x)$ is reduced by a factor of $1/\sqrt{e}$. This definition, which is, of course, arbitrary, has the advantage of coinciding with that of the "root-mean-square deviation" of the x variable (*cf.* Chap. III, § C-5).

With this convention, we can calculate the width Δx of the wave packet (10), which is equal to:

$$\Delta x = \frac{a}{2} \tag{12}$$

We can proceed in the same way to calculate the width Δk, since $|g(k,0)|^2$ is also a Gaussian function. This gives:

$$\Delta k = \frac{1}{a} \tag{13-a}$$

or:

$$\Delta p = \frac{\hbar}{a} \tag{13-b}$$

Thus we obtain:

$$\Delta x.\Delta p = \frac{\hbar}{2} \tag{14}$$

a result which is entirely compatible with the Heisenberg relation.

3. Evolution of the wave packet

3-a. Calculation of $\psi(x,t)$

In order to calculate the wave function $\psi(x,t)$ at time t, all we need to do is use the general formula (C-6) of Chapter I, which gives the wave function of a free particle. We obtain:

$$\psi(x,t) = \frac{\sqrt{a}}{(2\pi)^{3/4}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{a^2}{4}(k-k_0)^2}\, \mathrm{e}^{i[kx-\omega(k)t]}\, \mathrm{d}k \tag{15}$$

with $\omega(k) = \hbar k^2/2m$ (dispersion relation for a free particle). Let us show that, at time t, the wave packet still remains Gaussian. Expression (15) can be transformed by grouping, as above, all the k-dependent terms in the exponents into a perfect square. We can then use (7), and we find:

$$\psi(x,t) = \left(\frac{2a^2}{\pi}\right)^{1/4} \frac{\mathrm{e}^{i\varphi}}{\left(a^4 + \dfrac{4\hbar^2 t^2}{m^2}\right)^{1/4}}\, \mathrm{e}^{ik_0 x}\, \exp\left\{ -\frac{\left[x - \dfrac{\hbar k_0}{m} t\right]^2}{a^2 + \dfrac{2i\hbar t}{m}} \right\} \tag{16-a}$$

where φ is real and independent of x:

$$\varphi = -\,\theta - \frac{\hbar k_0^2}{2m}\, t \quad \text{with} \quad \tan 2\theta = \frac{2\hbar t}{ma^2} \tag{16-b}$$

Let us calculate the probability density $|\psi(x,t)|^2$ of the particle at time t. We obtain:

$$|\psi(x,t)|^2 = \sqrt{\frac{2}{\pi a^2}}\,\frac{1}{\sqrt{1+\dfrac{4\hbar^2t^2}{m^2a^4}}}\,\exp\left\{-\frac{2a^2\left(x-\dfrac{\hbar k_0}{m}t\right)^2}{a^4+\dfrac{4\hbar^2t^2}{m^2}}\right\} \tag{17}$$

Let us show that the norm of the wave packet, $\displaystyle\int_{-\infty}^{+\infty}|\psi(x,t)|^2\,\mathrm{d}x$, is not time-dependent (we shall see in Chapter III that this property results from the fact that the Hamiltonian H of the particle is Hermitian). We could, to this end, use (7) again in order to integrate expression (17) from $-\infty$ to $+\infty$. It is quicker to observe from expression (15) that the Fourier transform of $\psi(x,t)$ is given by:

$$g(k,t) = \mathrm{e}^{-iw(k)t}\,g(k,0) \tag{18}$$

$g(k,t)$ therefore obviously has the same norm as $g(k,0)$. Now the Parseval-Plancherel equality [*cf.* relation (45) of Appendix I] tells us that $\psi(x,t)$ and $g(k,t)$ have the same norm, as do $\psi(x,0)$ and $g(k,0)$. From this we deduce that $\psi(x,t)$ has the same norm as $\psi(x,0)$.

3-b. Velocity of the wave packet

We see in (17) that the probability density $|\psi(x,t)|^2$ is a Gaussian function, centered at $x = V_0t$, where the velocity V_0 is defined by:

$$V_0 = \frac{\hbar k_0}{m} \tag{19}$$

We could have expected this result, in view of the general expression (C-32) of Chapter I, which gives the group velocity V_G.

3-c. Spreading of the wave packet

Let us take up formula (17) again. The width $\Delta x(t)$ of the wave packet at time t, from definition (11), is equal to:

$$\Delta x(t) = \frac{a}{2}\sqrt{1+\frac{4\hbar^2t^2}{m^2a^4}} \tag{20}$$

We see (*cf.* Fig. 1) that the evolution of the wave packet is not purely a simple displacement at a velocity V_0. The wave packet also undergoes a deformation. When t increases from $-\infty$ to 0, the width of the wave packet decreases, reaching a minimum at $t = 0$. Then, as t continues to increase, $\Delta x(t)$ grows without bound (spreading of the wave packet).

It can be seen from (17) that the height of the wave packet also varies, but in opposition to the width, so the norm of $\psi(x,t)$ remains constant.

The properties of the function $g(k,t)$ are completely different. In fact [*cf.* formula (18)]:

$$|g(k,t)| = |g(k,0)| \tag{21}$$

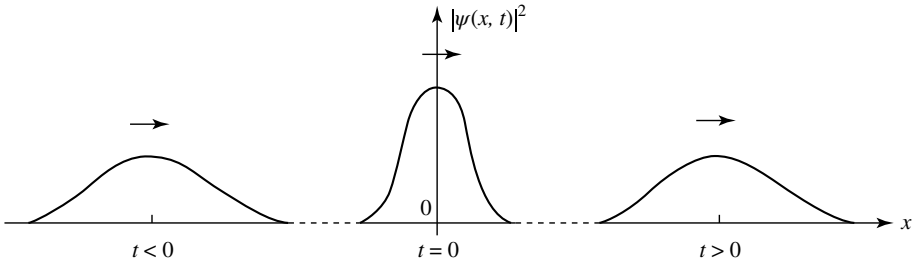


Figure 1: For negative t, the Gaussian wave packet decreases in width as it propagates. At time $t = 0$, it is a "minimum" wave packet: the product $\Delta x.\Delta p$ is equal to $\hbar/2$. Then, for $t > 0$, the wave packet spreads again as it propagates.

Therefore, the average momentum of the wave packet ($\hbar k_0$) and its momentum dispersion ($\hbar\Delta k$) do not vary in time. We shall see later (*cf.* Chap. III) that this arises from the fact that the momentum is a constant of the motion for a free particle. Physically, it is clear that since the free particle encounters no obstacle, the momentum distribution cannot change.

The existence of a momentum dispersion $\Delta p = \hbar\Delta k = \hbar/a$ means that the velocity of the particle is only known to within $\Delta v = \dfrac{\Delta p}{m} = \dfrac{\hbar}{ma}$. Imagine a group of classical particles starting at time $t = 0$ from the point $x = 0$, with a velocity dispersion equal to Δv. At time t, the dispersion of their positions will be $\delta x_{cl} = \Delta v\,|t| = \dfrac{\hbar\,|t|}{ma}$; this dispersion increases linearly with t, as shown in Figure 2. Let us draw on the same graph the curve which gives the evolution in time of $\Delta x(t)$; when t becomes infinite, $\Delta x(t)$ practically coincides with δx_{cl} [the branch of the hyperbola which represents $\Delta x(t)$ has for its asymptotes the straight lines which correspond to δx_{cl}]. Thus, we can say that, when t is very large, there exists a quasi-classical interpretation of the width Δx. On the other hand, when t approaches 0, $\Delta x(t)$ takes on values which differ more and more from δx_{cl}. The quantum particle must indeed constantly satisfy the Heisenberg relation $\Delta x \cdot \Delta p \geqslant \hbar/2$ which, since Δp is fixed, imposes a lower limit on Δx. This corresponds to what can be seen in Figure 2.

Comments:

(i) The spreading of a packet of free waves is a general phenomenon which is not limited to the special case studied here. It can be shown that, for an arbitrary free wave packet, the variation in time of its width has the shape shown in Figure 2 (*cf.* exercise 4 of Complement $\mathrm{L_{III}}$).

(ii) In Chapter I, a simple argument led us in (C-17) to $\Delta x \cdot \Delta k \simeq 1$, without making any particular hypothesis about $g(k)$. We simply assumed that $g(k)$ has a peak of width Δk whose shape is that of Figure 3 of Chapter I (which is indeed the case in this complement). Then how did we obtain $\Delta x \cdot \Delta k \gg 1$ (for example, for a Gaussian wave packet when t is large)?

Of course, this is only an apparent contradiction. In Chapter I, in order to find

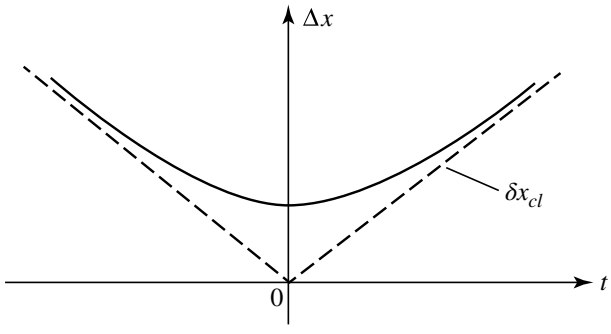


Figure 2: Variation in time of the width Δx of the wave packet of Figure 1. For large t, Δx approaches the dispersion δx_{cl} of the positions of a group of classical particles which left $x = 0$ at time $t = 0$ with a velocity dispersion $\Delta p/m$.

$\Delta x \cdot \Delta k \simeq 1$, we assumed in (C-13) that the argument $\alpha(k)$ of $g(k)$ could be approximated by a linear function in the domain Δk. Thus we implicitly assumed a supplementary hypothesis: that the nonlinear terms make a negligible contribution to the phase of $g(k)$ in the domain Δk. For example, for the terms which are of second order in $(k - k_0)$, it is necessary that:

$$\Delta k^2 \left[\frac{\mathrm{d}^2\alpha}{\mathrm{d}k^2}\right]_{k=k_0} \ll 2\pi \tag{22}$$

If, on the contrary, the phase $\alpha(k)$ cannot be approximated in the domain Δk by a linear function with an error much smaller than 2π, we find when we return to the argument of Chapter I that the wave packet is larger than was predicted by (C-17).

In the case of the Gaussian wave packet studied in the present complement, we have $\Delta k \simeq \frac{1}{a}$ and $\alpha(k) = -\frac{\hbar k^2}{2m}t$. Consequently, condition (22) can be written $\left(\frac{1}{a}\right)^2 \frac{\hbar t}{m} \ll 2\pi$. Indeed, we can verify from (20) that, as long as this condition is fulfilled, the product $\Delta x \cdot \Delta k$ is approximately equal to 1.

Complement H_I

Stationary states of a particle in one-dimensional square potentials

1 Behavior of a stationary wave function $\varphi(x)$ 63
1-a Regions of constant potential energy 63
1-b Behavior of $\varphi(x)$ at a potential energy discontinuity 64
1-c Outline of the calculation 65
2 Some simple cases 65
2-a Potential steps 65
2-b Potential barriers 68
2-c Bound states: square well potential 71

We saw in Chapter I (*cf.* § D-2) the interest in studying the motion of a particle in a "square potential" whose rapid spatial variations for certain values of x introduce purely quantum effects. The shape of the wave functions associated with the stationary states of the particle was predicted by considering an optical analogy which enabled us to understand very simply how these new physical effects appear.

In this complement, we outline the quantitative calculation of the stationary states of the particle. We shall give the results of this calculation for a certain number of simple cases, and discuss their physical implications. We limit ourselves to one-dimensional models (*cf.* Complement F_I).

1. Behavior of a stationary wave function $\varphi(x)$

1-a. Regions of constant potential energy

In the case of a square potential, $V(x)$ is a constant function $V(x) = V$ in certain regions of space. In such a region, equation (D-8) of Chapter I can be written:

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi(x) + \frac{2m}{\hbar^2}(E - V)\,\varphi(x) = 0 \tag{1}$$

We shall distinguish between several cases:

(i) $E > V$

Let us introduce the positive constant k, defined by

$$E - V = \frac{\hbar^2 k^2}{2m} \tag{2}$$

The solution of equation (1) can then be written:

$$\varphi(x) = A\,\mathrm{e}^{ikx} + A'\,\mathrm{e}^{-ikx} \tag{3}$$

where A and A' are complex constants.

(*ii*) $E < V$

This condition corresponds to regions of space which would be forbidden to the particle by the laws of classical mechanics. In this case, we introduce the positive constant ρ defined by:

$$V - E = \frac{\hbar^2 \rho^2}{2m} \tag{4}$$

and the solution of (1) can be written:

$$\varphi(x) = B\, \mathrm{e}^{\rho x} + B'\, \mathrm{e}^{-\rho x} \tag{5}$$

where B and B' are complex constants.

(*iii*) $E = V$. In this special case, $\varphi(x)$ is a linear function of x.

1-b. Behavior of $\varphi(x)$ at a potential energy discontinuity

How does the wave function behave at a point $x = x_1$, where the potential $V(x)$ is discontinuous? One might expect the wave function $\varphi(x)$ to behave strangely at this point, becoming itself discontinuous, for example. The aim of this section is to show that this is not the case: $\varphi(x)$ and $\mathrm{d}\varphi/\mathrm{d}x$ are continuous, and it is only the second derivative $\mathrm{d}^2\varphi/\mathrm{d}x^2$ that is discontinuous at $x = x_1$.

Without giving a rigorous proof, let us try to understand this property. To do this, recall that a square potential must be considered (*cf.* Chap. I, § D-2-a) as the limit, when $\varepsilon \longrightarrow 0$, of a potential $V_\varepsilon(x)$ equal to $V(x)$ outside the interval $[x_1 - \varepsilon, x_1 + \varepsilon]$, and varying continuously within this interval. Then consider the equation:

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi_\varepsilon(x) + \frac{2m}{\hbar^2}[E - V_\varepsilon(x)]\varphi_\varepsilon(x) = 0 \tag{6}$$

where $V_\varepsilon(x)$ is assumed to be bounded, independently of ε, within the interval $[x_1 - \varepsilon, x_1 + \varepsilon]$. Choose a solution $\varphi_\varepsilon(x)$ which, for $x < x_1 - \varepsilon$, coincides with a given solution of (1). The problem is to show that, when $\varepsilon \longrightarrow 0$, $\varphi_\varepsilon(x)$ tends towards a function $\varphi(x)$ which is continuous and differentiable at $x = x_1$. Let us grant that $\varphi_\varepsilon(x)$ remains bounded[1], whatever the value of ε, in the neighborhood of $x = x_1$. Physically, this means that the probability density remains finite. Integrating (6) between $x_1 - \eta$ and $x_1 + \eta$, we obtain:

$$\frac{\mathrm{d}\varphi_\varepsilon}{\mathrm{d}x}(x_1 + \eta) - \frac{\mathrm{d}\varphi_\varepsilon}{\mathrm{d}x}(x_1 - \eta) = \frac{2m}{\hbar^2}\int_{x_1-\eta}^{x_1+\eta} [V_\varepsilon(x) - E]\, \varphi_\varepsilon(x)\, \mathrm{d}x \tag{7}$$

At the limit where $\varepsilon \longrightarrow 0$, the function to be integrated on the right-hand side of this expression remains bounded, owing to our previous assumption. Consequently, if η tends towards zero, the integral also tends towards zero, and:

$$\frac{\mathrm{d}\varphi}{\mathrm{d}x}(x_1 + \eta) - \frac{\mathrm{d}\varphi}{\mathrm{d}x}(x_1 - \eta) \xrightarrow[\eta \to 0]{} 0 \tag{8}$$

Thus, at this limit, $\mathrm{d}\varphi/\mathrm{d}x$ is continuous at $x = x_1$, and so is $\varphi(x)$ (since it is the integral of a continuous function). On the other hand, $\mathrm{d}^2\varphi/\mathrm{d}x^2$ is discontinuous, and, as can be seen directly

[1] This point could be proved mathematically from the properties of the differential equation (1).

from (1), makes a jump at $x = x_1$, which is equal to $\frac{2m}{\hbar^2}\varphi(x_1)\sigma_V$ [where σ_V represents the change in $V(x)$ at $x = x_1$].

Comment:

It is essential, in the preceding argument, that $V_\varepsilon(x)$ remain bounded. In certain exercises of Complement K_I, for example, the case is considered for which $V(x) = \alpha\,\delta(x)$, an unbounded function whose integral remains finite. In this case, $\varphi(x)$ remains continuous, but $d\varphi/dx$ does not.

1-c. Outline of the calculation

The procedure for determining the stationary states in a "square potential" is therefore the following: in all regions where $V(x)$ is constant, write $\varphi(x)$ in whichever of the two forms (3) or (5) is applicable; then "match" these functions by requiring the continuity of $\varphi(x)$ and of $d\varphi/dx$ at the points where $V(x)$ is discontinuous.

2. Some simple cases

Let us now carry out the quantitative calculation of the stationary states, performed according to the method described above, for all the forms of the potential $V(x)$ considered in § D-2-c of Chapter I. This will ensure that the form of the solutions is indeed the one predicted by the optical analogy.

2-a. Potential steps

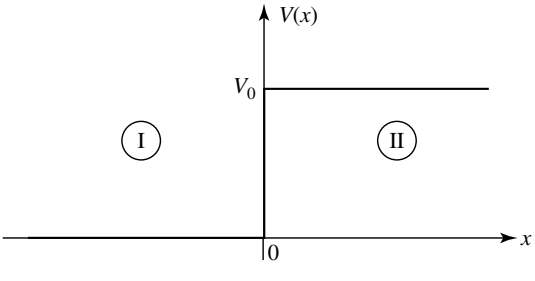


Figure 1: Potential step.

α. *Case where $E > V_0$; partial reflection*

Set:

$$\sqrt{\frac{2mE}{\hbar^2}} = k_1 \tag{9}$$

$$\sqrt{\frac{2m(E-V_0)}{\hbar^2}} = k_2 \tag{10}$$

The solution of (1) has the form (3) in the two regions I ($x < 0$) and II ($x > 0$):

$$\varphi_{\mathrm{I}}(x) = A_1 \, \mathrm{e}^{ik_1 x} + A'_1 \, \mathrm{e}^{-ik_1 x} \tag{11}$$

$$\varphi_{\mathrm{II}}(x) = A_2 \, \mathrm{e}^{ik_2 x} + A'_2 \, \mathrm{e}^{-ik_2 x} \tag{12}$$

Since equation (1) is homogeneous, the calculation method of § 1-c can only enable us to determine the ratios A'_1/A_1, A_2/A_1 and A'_2/A_1. In fact, the two matching conditions at $x = 0$ do not suffice for the determination of these three ratios. This is why we shall choose $A'_2 = 0$, which amounts to limiting ourselves to the case of an incident particle coming from $x = -\infty$. The matching conditions then give:

$$\frac{A'_1}{A_1} = \frac{k_1 - k_2}{k_1 + k_2} \tag{13}$$

$$\frac{A_2}{A_1} = \frac{2k_1}{k_1 + k_2} \tag{14}$$

$\varphi_{\mathrm{I}}(x)$ is the superposition of two waves. The first one (the term in A_1) corresponds to an incident particle, with momentum $p = \hbar k_1$, propagating from left to right. The second one (the term in A'_1) corresponds to a reflected particle, with momentum $-\hbar k_1$, propagating in the opposite direction. Since we have chosen $A'_2 = 0$, $\varphi_{\mathrm{II}}(x)$ consists of only one wave, which is associated with a transmitted particle. We shall see in Chapter III (*cf.* § D-1-c-β) how it is possible, using the concept of a probability current, to define the transmission coefficient T and the reflection coefficient R of the potential step (see also § 2 of Complement $\mathrm{B_{III}}$). These coefficients give the probability for the particle, arriving from $x = -\infty$, to pass the potential step at $x = 0$ or to turn back. Thus we find:

$$R = \left| \frac{A'_1}{A_1} \right|^2 \tag{15}$$

and, for[2] T:

$$T = \frac{k_2}{k_1} \left| \frac{A_2}{A_1} \right|^2 \tag{16}$$

Taking (13) and (14) into account, we then have:

$$R = 1 - \frac{4k_1 k_2}{(k_1 + k_2)^2} \tag{17}$$

$$T = \frac{4k_1 k_2}{(k_1 + k_2)^2} \tag{18}$$

[2]The physical origin of the factor k_2/k_1, appearing in T is discussed in § 2 of Complement $\mathrm{J_I}$.

It is easy to verify that $R+T=1$: it is certain that the particle will be either transmitted or reflected. Contrary to the predictions of classical mechanics, the incident particle has a non-zero probability of turning back. This point was explained in Chapter I, using the optical analogy and considering the reflection of a light wave from a plane interface (with $n_1 > n_2$). Furthermore, we know that in optics, no phase delay is created by such a reflection; equations (13) and (14) do indeed show that the ratios A_1'/A_1 and A_2/A_1 are real. Therefore, the quantum particle is not slowed down by its reflection or transmission (*cf.* Complement $\mathrm{J_I}$, § 2). Finally, using (9), (10) and (18) it is easy to verify that, if $E \gg V_0$, we have $T \simeq 1$: when the energy of the particle is sufficiently large compared to the height of the potential step, the particle clears this step as if it did not exist.

β. Case where $E < V_0$; total reflection

We then replace (10) and (12) by:

$$\sqrt{\frac{2m(V_0-E)}{\hbar^2}} = \rho_2 \tag{19}$$

$$\varphi_{\mathrm{II}}(x) = B_2\, \mathrm{e}^{\rho_2 x} + B_2'\, \mathrm{e}^{-\rho_2 x} \tag{20}$$

For the solution to remain bounded when $x \longrightarrow +\infty$, it is necessary that:

$$B_2 = 0 \tag{21}$$

The matching conditions at $x=0$ yield in this case:

$$\frac{A_1'}{A_1} = \frac{k_1 - i\rho_2}{k_1 + i\rho_2} \tag{22}$$

$$\frac{B_2'}{A_1} = \frac{2k_1}{k_1 + i\rho_2} \tag{23}$$

The reflection coefficient R is then equal to:

$$R = \left|\frac{A_1'}{A_1}\right|^2 = \left|\frac{k_1 - i\rho_2}{k_1 + i\rho_2}\right|^2 = 1 \tag{24}$$

As in classical mechanics, the particle is always reflected (total reflection). Nevertheless, there is an important difference, which has already been pointed out in Chapter I. Because of the existence of the evanescent wave $\mathrm{e}^{-\rho_2 x}$, the particle has a non-zero probability of presence in the region of space which, classically, would be forbidden to it. This probability decreases exponentially with x and becomes negligible when x is greater than the "range" $1/\rho_2$ of the evanescent wave. Note also that the coefficient A_1'/A_1 is complex. A certain phase shift appears upon reflection, which, physically, is due to the fact that the particle is delayed when it penetrates the $x > 0$ region (*cf.* Complement $\mathrm{J_I}$, § 1 and also $\mathrm{B_{III}}$, § 3). This phase shift is analogous to the one that appears when light is reflected from a metallic type of substance; however, there is no analogue in classical mechanics.

Comment:

When $V_0 \longrightarrow +\infty$, $\rho_2 \longrightarrow +\infty$, so that (22) and (23) yield:

$$\begin{cases} A_1' \longrightarrow -A_1 \\ B_2' \longrightarrow 0 \end{cases} \tag{25}$$

In the $x > 0$ region, the wave, whose range decreases without bound, tends towards zero. Since $(A_1 + A_1') \longrightarrow 0$, the wave function $\varphi(x)$ goes to zero at $x = 0$, so that it remains continuous at this point. On the other hand, its derivative, which changes abruptly from the value $2ikA_1$ to zero, is no longer continuous. This is due to the fact that since the potential jump is infinite at $x = 0$, the integral of (7) no longer tends towards zero when η tends towards 0.

2-b. Potential barriers

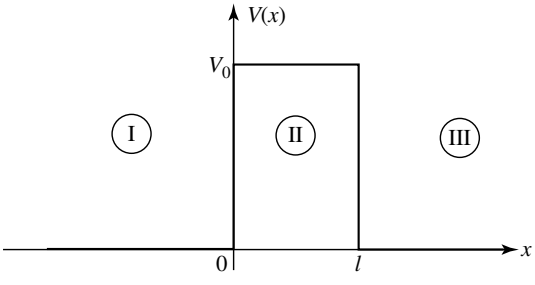


Figure 2: Square potential barrier.

α. Case where $E > V_0$; resonances[3]

Using notations (9) and (10), we find, in the three regions I ($x < 0$), II ($0 < x < l$) and III ($x > l$) shown in Fig. 2:

$$\varphi_{\mathrm{I}}(x) = A_1\, \mathrm{e}^{ik_1 x} + A_1'\, \mathrm{e}^{-ik_1 x} \tag{26-a}$$

$$\varphi_{\mathrm{II}}(x) = A_2\, \mathrm{e}^{ik_2 x} + A_2'\, \mathrm{e}^{-ik_2 x} \tag{26-b}$$

$$\varphi_{\mathrm{III}}(x) = A_3\, \mathrm{e}^{ik_1 x} + A_3'\, \mathrm{e}^{-ik_1 x} \tag{26-c}$$

Let us choose, as above, $A_3' = 0$ (incident particle coming from $x = -\infty$). The matching conditions at $x = l$ then give A_2 and A_2' in terms of A_3, and those at $x = 0$ give A_1 and A_1' in terms of A_2 and A_2' (and, consequently, in terms of A_3). Thus we find:

$$\begin{aligned} A_1 &= \left[\cos k_2 l - i\frac{k_1^2 + k_2^2}{2k_1 k_2} \sin k_2 l\right] \mathrm{e}^{ik_1 l}\, A_3 \\ A_1' &= i\frac{k_2^2 - k_1^2}{2k_1 k_2} \sin k_2 l\, \mathrm{e}^{ik_1 l}\, A_3 \end{aligned} \tag{27}$$

[3] V_0 can be either positive (the case of a potential barrier like the one shown in Figure 2) or negative (a potential well).

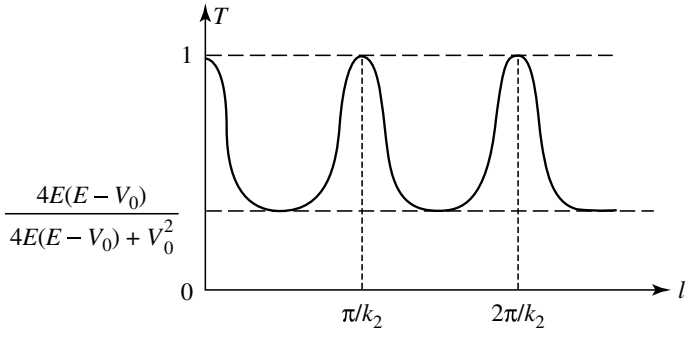


Figure 3: Variations of the transmission coefficient T of the barrier as a function of its width (the height V_0 of the barrier and the energy E of the particle are fixed). Resonances appear each time that l is an integral multiple of the half-wavelength π/k_2, in region II.

A_1'/A_1 and A_3/A_1 enable us to calculate the reflection coefficient R and the transmission coefficient T of the barrier:

$$R = \left|\frac{A_1'}{A_1}\right|^2 = \frac{(k_1^2 - k_2^2)^2 \sin^2 k_2 l}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2 k_2 l} \tag{28-a}$$

$$T = \left|\frac{A_3}{A_1}\right|^2 = \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2 k_2 l} \tag{28-b}$$

It is then easy to verify that $R + T = 1$. Taking (9) and (10) into account, we have:

$$T = \frac{4E(E-V_0)}{4E(E-V_0) + V_0^2 \sin^2\left[\sqrt{2m(E-V_0)}\, l/\hbar\right]} \tag{29}$$

The variations with respect to l of the transmission coefficient T are shown in Figure 3 (with E and V_0 fixed): T oscillates periodically between its minimum value, $\left[1 + \frac{V_0^2}{4E(E-V_0)}\right]^{-1}$, and its maximum value, which is 1. This function is the analogue of the one describing the transmission of a Fabry-Perot interferometer. As in optics, the resonances (obtained when $T = 1$, that is, when $k_2 l = n\pi$) correspond to the values of l which are integral multiples of the half-wavelength of the particle in region II. When $E > V_0$, the reflection of the particle at each of the potential discontinuities occurs without a phase shift of the wave function (*cf.* § 2-a-α). This is why the resonance condition $k_2 l = n\pi$ corresponds to the values of l for which a system of standing waves can exist in region II. On the other hand, far from the resonances, the various waves which are reflected at $x = 0$ and $x = l$ destroy each other by interference, so that the values of the wave function are small. A study of the propagation of a wave packet (analogous to the one in Complement J_I) would show that, if the resonance condition is satisfied, the wave packet spends a relatively long time in region II. In quantum mechanics this phenomenon is called *resonance scattering.*

β. Case where $E < V_0$; tunnel effect

We must now replace (26-b) by (20), ρ_2 still being given by (19). The matching conditions at $x = 0$ and $x = l$ enable us to calculate the transmission coefficient of the barrier. In fact, it is unnecessary to perform the calculations again: all we must do is replace, in the equations obtained in § α, the wave vector k_2 by $-i\rho_2$. We then have:

$$T = \left|\frac{A_3}{A_1}\right|^2 = \frac{4E(V_0 - E)}{4E(V_0 - E) + V_0^2 \sinh^2\left[\sqrt{2m(V_0 - E)}\, l/\hbar\right]} \tag{30}$$

with, of course, $R = 1 - T$. When $\rho_2 l \gg 1$, we have:

$$T \simeq \frac{16E(V_0 - E)}{V_0^2}\, \mathrm{e}^{-2\rho_2 l} \tag{31}$$

We have already seen in Chapter I why, contrary to the classical predictions, the particle has a non-zero probability of crossing the potential barrier. The wave function in region II is not zero, but has the behavior of an "evanescent wave" of range $1/\rho_2$. When $l \lesssim 1/\rho_2$, the particle has a considerable probability of crossing the barrier by the "tunnel effect". This effect has numerous physical applications: the inversion of the ammonia molecule (*cf.* Complement G_{IV}), the tunnel diode, the Josephson effect, the α-decay of certain nuclei, etc...

For an electron, the range of the evanescent wave is:

$$\left(\frac{1}{\rho_2}\right)_{el} \simeq \frac{1.96}{\sqrt{V_0 - E}}\ \text{Å} \tag{32}$$

where E and V_0 are expressed in electron-volts (this formula is easily obtained by replacing, in formula (8) of Complement A_I, $\lambda = 2\pi/k$ by $2\pi/\rho_2$). Now consider an electron of energy 1 eV which encounters a barrier for which $V_0 = 2$ eV and $l = 1$ Å. The range of the evanescent wave is then 1.96 Å, that is, of the order of l: the electron must then have a considerable probability of crossing the barrier. Indeed, formula (30) gives in this case:

$$T \simeq 0.78 \tag{33}$$

The quantum result is radically different from the classical result: the electron has approximately 8 chances out of 10 of crossing the barrier.

Let us now assume that the incident particle is a proton (whose mass is about 1 840 times that of the electron). The range $1/\rho_2$ then becomes:

$$\left(\frac{1}{\rho_2}\right)_{pr} \simeq \frac{1.96}{\sqrt{1\ 840(V_0 - E)}}\ \text{Å} \simeq \frac{4.6}{\sqrt{V_0 - E}}\ 10^{-2}\ \text{Å} \tag{34}$$

If we retain the same values: $E = 1$ eV, $V_0 = 2$ eV, $l = 1$ Å, we find a range $1/\rho_2$, much smaller than l. Formula (31) then gives:

$$T \simeq 4 \times 10^{-19} \tag{35}$$

Under these conditions, the probability of the proton's crossing the potential barrier is negligible. This is all the more true if we apply (31) to macroscopic objects, for which we find such small probabilities that they cannot possibly play any role in physical phenomena.

2-c. Bound states: square well potential

α. Well of finite depth

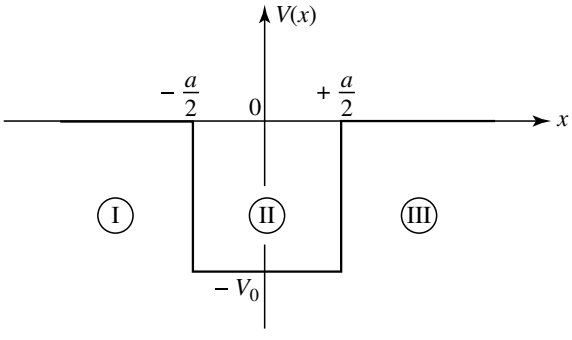


Figure 4: Square well potential.

We shall limit ourselves to studying the case $-V_0 < E < 0$ (the case $E > 0$ was included in the calculations of the preceding section 2-b-α).

In regions I $\left(x < -\frac{a}{2}\right)$, II $\left(-\frac{a}{2} \leqslant x \leqslant \frac{a}{2}\right)$, and III $\left(x > \frac{a}{2}\right)$ shown in Fig. 4, we have respectively:

$$\varphi_{\mathrm{I}}(x) = B_1\, \mathrm{e}^{\rho x} + B_1'\, \mathrm{e}^{-\rho x} \tag{36-a}$$

$$\varphi_{\mathrm{II}}(x) = A_2\, \mathrm{e}^{ikx} + A_2'\, \mathrm{e}^{-ikx} \tag{36-b}$$

$$\varphi_{\mathrm{III}}(x) = B_3\, \mathrm{e}^{\rho x} + B_3'\, \mathrm{e}^{-\rho x} \tag{36-c}$$

with

$$\rho = \sqrt{-\frac{2mE}{\hbar^2}} \tag{37}$$

$$k = \sqrt{\frac{2m(E+V_0)}{\hbar^2}} \tag{38}$$

Since $\varphi(x)$ must be bounded in region I, we must have:

$$B_1' = 0 \tag{39}$$

The matching conditions at $x = -\frac{a}{2}$ then give:

$$\begin{aligned} A_2 &= \mathrm{e}^{(-\rho+ik)a/2}\, \frac{\rho+ik}{2ik}\, B_1 \\ A_2' &= -\,\mathrm{e}^{-(\rho+ik)a/2}\, \frac{\rho-ik}{2ik}\, B_1 \end{aligned} \tag{40}$$

and those at $x = a/2$:

$$\frac{B_3}{B_1} = \frac{\mathrm{e}^{-\rho a}}{4ik\rho}\left[(\rho+ik)^2\ \mathrm{e}^{ika} - (\rho-ik)^2\ \mathrm{e}^{-ika}\right]$$
$$\frac{B'_3}{B_1} = \frac{\rho^2+k^2}{2k\rho}\sin ka \tag{41}$$

But $\varphi(x)$ must also be bounded in region III. Therefore, it is necessary that $B_3 = 0$, that is:

$$\left(\frac{\rho-ik}{\rho+ik}\right)^2 = \mathrm{e}^{2ika} \tag{42}$$

Since ρ and k depend on E, equation (42) can only be satisfied for certain values of E. Imposing a bound on $\varphi(x)$ in all regions of space thus entails the quantization of energy. More precisely, two cases are possible:

(i) if:

$$\frac{\rho-ik}{\rho+ik} = -\mathrm{e}^{ika} \tag{43}$$

we have:

$$\frac{\rho}{k} = \tan\left(\frac{ka}{2}\right) \tag{44}$$

Set:

$$k_0 = \sqrt{\frac{2mV_0}{\hbar^2}} = \sqrt{k^2+\rho^2} \tag{45}$$

We then obtain:

$$\frac{1}{\cos^2\left(\dfrac{ka}{2}\right)} = 1 + \tan^2\frac{ka}{2} = \frac{k^2+\rho^2}{k^2} = \left(\frac{k_0}{k}\right)^2 \tag{46}$$

Equation (43) is thus equivalent to the system of equations:

$$\begin{cases} \left|\cos\left(\dfrac{ka}{2}\right)\right| = \dfrac{k}{k_0} & \text{(47a)} \\ \tan\left(\dfrac{ka}{2}\right) > 0 & \text{(47b)} \end{cases}$$

The energy levels are determined by the intersection of a straight line, having a slope $1/k_0$, with sinusoidal arcs (long dashed lines in Figure 5). Thus we obtain a certain number of energy levels, whose wave functions are even. This becomes clear if we substitute (43) into (40) and (41); it is easy to verify that $B'_3 = B_1$ and that $A_2 = A'_2$, so that $\varphi(-x) = \varphi(x)$.

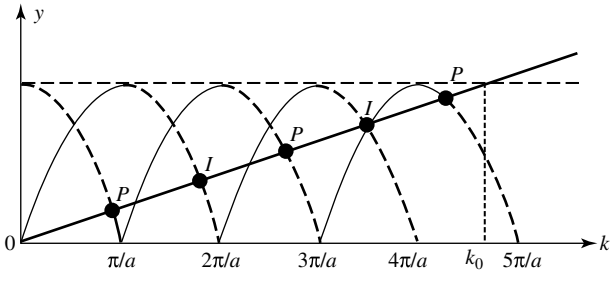


Figure 5: Graphic solution of equation (42), giving the energies of the bound states of a particle in a square well potential. In the case shown in the figure, there exist five bound states, three even (associated with the points P of the figure), and two odd (points I).

(*ii*) if:

$$\frac{\rho - ik}{\rho + ik} = \mathrm{e}^{ika} \tag{48}$$

a calculation of the same type leads to:

$$\begin{cases} \left|\sin\left(\dfrac{ka}{2}\right)\right| = \dfrac{k}{k_0} & \text{(49a)} \\ \tan\left(\dfrac{ka}{2}\right) < 0 & \text{(49b)} \end{cases}$$

The energy levels are then determined by the intersection of the same straight line as before with other sinusoidal arcs (*cf.* short dashed lines in Figure 5). The levels thus obtained fall between those found in (*i*). It can easily be shown that the corresponding wave functions are odd.

Comment:

If $k_0 \leqslant \dfrac{\pi}{a}$, that is, if:

$$V_0 \leqslant V_1 = \frac{\pi^2\hbar^2}{2ma^2} \tag{50}$$

Figure 5 shows that there exists only one bound state of the particle, and this state has an even wave function. Then, if $V_1 \leqslant V_0 < 4V_1$, a first odd level appears, and so on: when V_0 increases, there appear alternatively even and odd levels. If $V_0 \gg V_1$, the slope $1/k_0$ of the straight line of Figure 5 is very small: for the lowest energy levels, we practically have:

$$k = \frac{n\pi}{a} \tag{51}$$

where n is an integer, and consequently:

$$E = \frac{n^2\pi^2\hbar^2}{2ma^2} - V_0 \tag{52}$$

β. Infinitely deep well

Assume $V(x)$ to be zero for $0 < x < a$ and infinite everywhere else. Set:

$$k = \sqrt{\frac{2mE}{\hbar^2}} \tag{53}$$

According to the comment made at the end of § 2-a-β of this complement, $\varphi(x)$ must be zero outside the interval $[0, a]$, and continuous at $x = 0$, as well as at $x = a$. Now for $0 \leq x \leq a$:

$$\varphi(x) = A\, e^{ikx} + A'\, e^{-ikx} \tag{54}$$

Since $\varphi(0) = 0$, it can be deduced that $A' = -A$, which leads to:

$$\varphi(x) = 2iA \sin kx \tag{55}$$

Moreover, $\varphi(a) = 0$, so that:

$$k = \frac{n\pi}{a} \tag{56}$$

where n is an arbitrary positive integer. If we normalize function (55), taking (56) into account, we then obtain the stationary wave functions:

$$\varphi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) \tag{57}$$

with energies:

$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2} \tag{58}$$

The quantization of the energy levels is thus, in this case, particularly simple.

Comments:

(*i*) Relation (56) simply expresses the fact that the stationary states are determined by the condition that the width a of the well must contain an integral number of half-wavelengths, π/k. This is not the case when the well has a finite depth (*cf.* § 2-c-α); the difference between the two cases arises from the phase shift of the wave function that occurs upon reflection from a potential step (*cf.* § 2-a-β).

(*ii*) It can easily be verified from (51) and (52) that, if the depth V_0 of a finite well approaches infinity, we find the energy levels of an infinite well.

References and suggestions for further reading:

Eisberg and Resnick (1.3), Chap. 6; Ayant and Belorizky (1.10), Chap. 4; Messiah (1.17), Chap. III; Merzbacher (1.16), Chap. 6; Valentin (16.1), annex V.

Complement J_I

Behavior of a wave packet at a potential step

1 **Total reflection:** $E < V_0$ **75**
2 **Partial reflection:** $E > V_0$ **79**

In Complement H_I, we determined the stationary states of a particle in various "square" potentials. For certain cases (a step potential, for example), the stationary states obtained consist of unbounded plane waves (incident, reflected and transmitted). Of course, since they cannot be normalized, such wave functions cannot really represent a physical state of the particle. However, they can be linearly superposed to form normalizable wave packets. Moreover, since such a wave packet is expanded directly in terms of stationary wave functions, its time evolution is very simple to determine. All we need to do is multiply each of the coefficients of the expansion by an imaginary exponential $e^{-iEt/\hbar}$ with a well-defined frequency $\frac{E}{h}$ (chap. I, § D-1-b).

We intend, in this complement, to construct such wave packets and study their time evolution for the case where the potential presents a "step" of height V_0, as in Figure 1 of Complement H_I. In this way, we shall be able to describe precisely the quantum behavior of the particle when it arrives at the potential step by determining the motion and the deformation of its associated wave packet. This will also enable us to confirm various results obtained in H_I through the study of the stationary states alone (reflection and transmission coefficients, delay upon reflection, etc...).

We shall set:

$$\sqrt{\frac{2mE}{\hbar^2}} = k$$
$$\sqrt{\frac{2mV_0}{\hbar^2}} = K_0 \tag{1}$$

and, as in Complement H_I, we shall distinguish between two cases, corresponding to k smaller or greater than K_0.

1. Total reflection: $E < V_0$

In this case, the stationary wave functions are given by formulas (11) and (20) of Complement H_I (k_1 will be simply called k here), the coefficients A_1, A'_1, B_2 and B'_2 of these formulas being related by equations (21), (22) and (23) of H_I.

We are going to construct a wave packet from these stationary wave functions by linearly superposing them. We shall choose only values of k less than K_0, so that the waves forming the packet undergo total reflection. To ensure this, we shall choose a function $g(k)$ (which characterizes the wave packet) which is zero for $k > K_0$. We

are going to focus our attention on the negative region of the x-axis, to the left of the potential barrier. In Complement H_I, relation (22) shows that the coefficients A_1 and A_1' of expression (11) for a stationary wave in this region have the same modulus. Therefore, we can set:

$$\frac{A_1'(k)}{A_1(k)} = \mathrm{e}^{-2i\theta(k)} \tag{2}$$

with [*cf.* formula (19) of H_I]:

$$\tan\,\theta(k) = \frac{\sqrt{K_0^2 - k^2}}{k} \tag{3}$$

Finally, the wave packet which we are going to consider can be written, at time $t = 0$, for negative x:

$$\psi(x,0) = \frac{1}{\sqrt{2\pi}} \int_0^{K_0} \mathrm{d}k\, g(k)[\mathrm{e}^{ikx} + \mathrm{e}^{-2i\theta(k)}\mathrm{e}^{-ikx}] \tag{4}$$

As in § C of Chapter I, we assume that $|g(k)|$ has a pronounced peak of width Δk about the value $k = k_0 < K_0$.

In order to obtain the expression for the wave function $\psi(x,t)$ at any time t, we simply use the general relation (D-14) of Chapter I:

$$\begin{aligned}\psi(x,t) =& \frac{1}{\sqrt{2\pi}} \int_0^{K_0} \mathrm{d}k\, g(k)\, \mathrm{e}^{i[kx-\omega(k)t]} \\ &+ \frac{1}{\sqrt{2\pi}} \int_0^{K_0} \mathrm{d}k\, g(k)\, \mathrm{e}^{-i[kx+\omega(k)t+2\theta(k)]}\end{aligned} \tag{5}$$

where $\omega(k) = \hbar k^2/2m$. By construction, *this expression is valid only for negative* x. Its first term represents the incident wave packet; its second term, the reflected packet. For simplicity, we shall assume $g(k)$ to be real. The stationary phase condition (*cf.* Chap. I, § C-2) then enables us to calculate the position x_i of the center of the incident wave packet. If, at $k = k_0$, we set the derivative with respect to k of the argument of the first exponential equal to zero, we obtain:

$$x_i = t\left[\frac{\mathrm{d}\omega}{\mathrm{d}k}\right]_{k=k_0} = \frac{\hbar k_0}{m}\, t \tag{6}$$

In the same way, the position x_r of the center of the reflected packet is obtained by differentiating the argument of the second exponential. Differentiating equation (3), we find:

$$\begin{aligned}[1 + \tan^2\theta]\,\mathrm{d}\theta &= \left[1 + \frac{K_0^2 - k^2}{k^2}\right]\mathrm{d}\theta \\ &= -\frac{\mathrm{d}k}{k^2}\sqrt{K_0^2 - k^2} - \frac{\mathrm{d}k}{\sqrt{K_0^2 - k^2}}\end{aligned} \tag{7}$$

that is:

$$\frac{K_0^2}{k^2}\,\mathrm{d}\theta = -\frac{K_0^2}{k^2}\,\frac{1}{\sqrt{K_0^2 - k^2}}\,\mathrm{d}k \tag{8}$$

Thus we have:

$$x_r = -\left[t\frac{\mathrm{d}\omega}{\mathrm{d}k} + 2\frac{\mathrm{d}\theta}{\mathrm{d}k} \right]_{k=k_0} = -\frac{\hbar k_0}{m}\, t + \frac{2}{\sqrt{K_0^2 - k_0^2}} \tag{9}$$

Formulas (6) and (9) enable us to describe more precisely the motion of the particle, localized in a region of small width Δx centered at x_i or x_r.

First of all, let us consider what happens for negative t. The center x_i of the incident wave packet propagates from left to right with a constant velocity $\hbar k_0/m$. On the other hand, we see from formula (9) that x_r is positive, that is, situated outside the region $x < 0$ where expression (5) for the wave function is valid. This means that, for all negative values of x, the various waves of the second term of (5) interfere destructively: *for negative t, there is no reflected wave packet*, but only an incident wave packet like those we studied in § C of Chapter I.

The center of the incident wave packet arrives at the barrier at time $t = 0$. During a certain interval of time around $t = 0$, the wave packet is localized in the region $x \simeq 0$ where the barrier is, and its form is relatively complicated. But, *when t is sufficiently large*, we see from (6) and (9) that *it is the incident wave packet which has disappeared*, and we are left with only the reflected wave packet. It is now x_i which is positive, while x_r has become negative: the waves of the incident packet interfere destructively for all negative values of x, while those of the reflected packet interfere constructively for $x = x_r < 0$. The reflected wave packet propagates towards the left at a speed of $-\hbar k_0/m$, opposite to that of the incident packet, whose mirror image it is; its form is unchanged[1]. Moreover, formula (9) shows that the *reflection has introduced a delay τ*, given by:

$$\tau = -2\left[\frac{\mathrm{d}\theta/\mathrm{d}k}{\mathrm{d}\omega/\mathrm{d}k} \right]_{k=k_0} = \frac{2m}{\hbar k_0 \sqrt{K_0^2 - k_0^2}} \tag{10}$$

Contrary to what is predicted by classical mechanics, the particle is not instantaneously reflected. Note that the delay τ is related to the phase shift $2\theta(k)$ between the incident wave and the reflected wave for a given value of k. Nevertheless, it should be observed that the delay of the wave packet is not simply proportional to $\theta(k_0)$, as would be the case for an unbounded plane wave, but to the *derivative* $\mathrm{d}\theta/\mathrm{d}k$ evaluated at $k = k_0$. Physically, this delay is due to the fact that, for t close to zero, the probability of presence of the particle in the region $x > 0$, which is classically forbidden, is not zero [evanescent wave, see comment (i) below]. It can be said, metaphorically, that the particle spends a time of the order of τ in this region before retracing its steps. Formula (10) shows that the closer the average energy $\dfrac{\hbar^2 k_0^2}{2m}$ of the wave packet is to the height V_0 of the barrier, the longer the delay τ.

Comments:

(i) Here we have focussed on the behavior of the wave packet for $x < 0$, but it is also possible to study what happens for $x > 0$. In this region, the wave packet can be written:

[1] We assume Δk to be small enough for the spreading of the wave packet to be negligible during the time interval considered.

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_0^{K_0} \mathrm{d}k\, g(k) B_2'(k)\, \mathrm{e}^{-\rho(k)x} \mathrm{e}^{-i\omega(k)t} \tag{11}$$

where:

$$\rho(k) = \sqrt{K_0^2 - k^2} \tag{12}$$

$B_2'(k)$ is given by equation (23) of Complement H_I when we replace A_1 by 1, k_1 by k and ρ_2 by ρ. An argument analogous to the one in § C-2 of Chapter I then shows that the modulus $|\psi(x,t)|$ of expression (11) is maximum when the phase of the function to be integrated over k is stationary. Now, according to expressions (22) and (23) of H_I, the argument of B_2' is half that of A_1', which, according to (2), is equal to $-2\theta(k)$. Consequently, if we expand $\omega(k)$ and $\theta(k)$ in the neighborhood of $k = k_0$, we obtain, for the phase of the function to be integrated over k in (11):

$$\left\{ -\left[\frac{\mathrm{d}\theta}{\mathrm{d}k}\right]_{k=k_0} - \left[\frac{\mathrm{d}\omega}{\mathrm{d}k}\right]_{k=k_0} t \right\} (k-k_0) = -\frac{\hbar k_0}{m}(k-k_0)\left(t - \frac{\tau}{2}\right) \tag{13}$$

[we have used (10) and the fact that $g(k)$ is assumed real]. From this we can deduce that $|\psi(x,t)|$ is maximum in the $x > 0$ region for[2] $t = \frac{\tau}{2}$. The time at which the wave packet turns back is therefore $\tau/2$, which gives us the same delay τ upon reflection that we obtained above. We also see from expression (13) that, as soon as $\left| t - \frac{\tau}{2} \right|$ exceeds the time Δt defined by:

$$\frac{\hbar k_0}{m}\, \Delta k\, \Delta t \simeq 1 \tag{14}$$

where Δk is the width of $g(k)$, the waves go out of phase and expression (11) for $|\psi(x,t)|$ becomes negligible. Thus, the wave packet as a whole remains in the $x > 0$ region during an interval of time Δt of the order of:

$$\Delta t = \frac{1/\Delta k}{\hbar k_0/m} \tag{15}$$

which corresponds approximately to the time it takes, in the $x < 0$ region, to travel a distance comparable to its width $1/\Delta k$.

(*ii*) Since Δk is assumed to be much smaller than k_0 and K_0, the comparison of (10) and (15) shows that:

$$\Delta t \gg \tau \tag{16}$$

The delay upon reflection thus involves, for the reflected wave packet, a displacement which is much smaller than its width.

[2]Note that the phase (13) does not depend on x, contrary to what we found in Chapter I for a free wave packet. It follows that, in the $x > 0$ region, $|\psi(x,t)|$ does not have a pronounced peak that moves with respect to time.

2. Partial reflection: $E > V_0$

We now consider a function $g(k)$ of width Δk, centered at a value $k = k_0 > K_0$, which is zero for $k < K_0$. The wave packet is formed in this case by superposing, with coefficients $g(k)$, the stationary wave functions whose expressions are given by formulas (11) and (12) of Complement H_I. We shall choose $A'_2 = 0$ so as to have the particle arrive at the barrier from the negative region of the Ox axis, and we shall take $A_1 = 1$. The coefficients $A'_1(k)$ and $A_2(k)$ are obtained from formulas (13) and (14) of Complement H_I (in which A_1 is replaced by 1, k_1 by k, and k_2 by $\sqrt{k^2 - K_0^2}$).

In order to describe the wave packet by a single expression, valid for all values of x, we can use the Heaviside "step function" $\theta(x)$ defined by:

$$\begin{aligned} &\theta(x) = 0 \qquad \text{if } x < 0 \\ &\theta(x) = 1 \qquad \text{if } x > 0 \end{aligned} \tag{17}$$

The wave packet we are studying can then be written:

$$\begin{aligned} \psi(x,t) = {} & \theta(-x)\frac{1}{\sqrt{2\pi}}\int_{K_0}^{+\infty} \mathrm{d}k\ g(k)\ \mathrm{e}^{i[kx-\omega(k)t]} \\ & + \theta(-x)\frac{1}{\sqrt{2\pi}}\int_{K_0}^{+\infty} \mathrm{d}k\ g(k)A'_1(k)\ \mathrm{e}^{-i[kx+\omega(k)t]} \\ & + \theta(x)\frac{1}{\sqrt{2\pi}}\int_{K_0}^{+\infty} \mathrm{d}k\ g(k)A_2(k)\ \mathrm{e}^{i[\sqrt{k^2-K_0^2}\,x-\omega(k)t]} \end{aligned} \tag{18}$$

It is composed of three wave packets: incident, reflected and transmitted. As in § 1 above, the stationary phase condition gives the position of their respective centers x_i, x_r and x_t. Since $A'_1(k)$ and $A_2(k)$ are real, we find:

$$x_i = \frac{\hbar k_0}{m} t \tag{19-a}$$

$$x_r = -\frac{\hbar k_0}{m} t \tag{19-b}$$

$$x_t = \frac{\hbar\sqrt{k_0^2 - K_0^2}}{m}\, t \tag{19-c}$$

A discussion analogous to that of (6) and (9) leads to the following conclusions: *for negative t, only the incident wave packet exists; for sufficiently large positive t, only the reflected and transmitted wave packets exist* (Fig. 1). Note that there is no delay, either upon reflection or upon transmission (this is due to the fact that the coefficients $A'_1(k)$ and $A_2(k)$ are real).

The incident and reflected wave packets propagate with velocities of $\hbar k_0/m$ and $-\hbar k_0/m$ respectively. Let us assume Δk to be sufficiently small that, within the interval $\left[k_0 - \frac{\Delta k}{2}, k_0 + \frac{\Delta k}{2}\right]$, we can neglect the variation of $A'_1(k)$ compared to that of $g(k)$. We can then, in the second term of (18), replace $A'_1(k)$ by $A'_1(k_0)$ and take it outside the integral. It is then easy to see that the reflected wave packet has the same form as the incident wave packet, being its mirror image. Its amplitude is smaller, however,

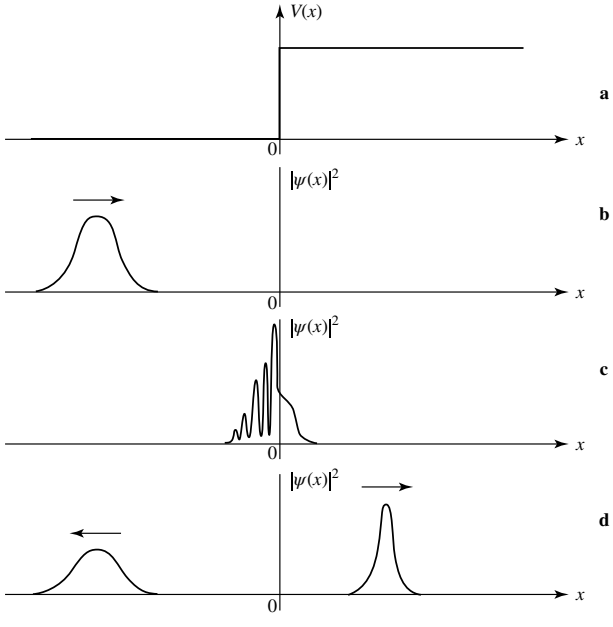


*Figure 1: Behavior of a wave packet at a potential step, in the case $E > V_0$. The potential is shown in figure **a**. In figure **b**, the wave packet is moving towards the step. Figure **c** shows the wave packet during the transitory period in which it splits in two. Interference between the incident and reflected waves are responsible for the oscillations of the wave packet in the $x < 0$ region. After a certain time (fig. **d**), we find two wave packets. The first one (the reflected wave packet) is returning towards the left; its amplitude is smaller than that of the incident wave packet, and its width is the same. The second one (the transmitted wave packet) propagates towards the right; its amplitude is slightly greater than that of the incident wave packet, but it is narrower.*

since, according to formula (13) of Complement H_I, $A_1'(k_0)$ is less than 1. The reflection coefficient R is, by definition, the ratio between the probabilities of finding the particle in the reflected wave packet and in the incident packet. Therefore, we have $R = |A_1'(k_0)|^2$, which indeed corresponds to equation (15) of Complement H_I [recall that we have chosen $A_1(k_0) = 1$].

The situation is different for the transmitted wave packet. We can still use the fact that Δk is very small in order to simplify its expression: we replace $A_2(k)$ by $A_2(k_0)$, and $\sqrt{k^2 - K_0^2}$ by the approximation:

$$\sqrt{k^2 - K_0^2} \simeq \sqrt{k_0^2 - K_0^2} + (k - k_0) \left[\frac{\mathrm{d}\sqrt{k^2 - K_0^2}}{\mathrm{d}k} \right]_{k=k_0}$$
$$\simeq q_0 + (k - k_0)\frac{k_0}{q_0} \tag{20}$$

with:

$$q_0 = \sqrt{k_0^2 - K_0^2} \tag{21}$$

The transmitted wave packet can then be written:

$$\psi_t(x,t) \simeq A_2(k_0)\, \mathrm{e}^{iq_0 x} \frac{1}{\sqrt{2\pi}} \int_{K_0}^{+\infty} \mathrm{d}k\, g(k)\, \mathrm{e}^{\,i\left[(k-k_0)\frac{k_0}{q_0}x - \omega(k)t\right]} \tag{22}$$

Let us compare this expression to the one for the incident wave packet:

$$\psi_i(x,t) = \mathrm{e}^{ik_0 x} \frac{1}{\sqrt{2\pi}} \int_{K_0}^{+\infty} \mathrm{d}k\, g(k)\, \mathrm{e}^{i[(k-k_0)x - \omega(k)t]} \tag{23}$$

We see that:

$$|\psi_t(x,t)| \simeq A_2(k_0) \left| \psi_i\!\left(\frac{k_0}{q_0}x, t\right) \right| \tag{24}$$

The transmitted wave packet thus has a slightly greater amplitude than that of the incident packet: according to formula (14) of Complement $\mathrm{H_I}$, $A_2(k_0)$ is greater than 1. However, its width is smaller, since, if $|\psi_i(x,t)|$ has a width Δx, formula (24) shows that the width of $|\psi_t(x,t)|$ is:

$$(\Delta x)_t = \frac{q_0}{k_0}\Delta x \tag{25}$$

The transmission coefficient (the ratio between the probabilities of finding the particle in the transmitted packet and in the incident packet) is thus the product of two factors:

$$T = \frac{q_0}{k_0}|A_2(k_0)|^2 \tag{26}$$

This indeed corresponds to formula (16) of Complement $\mathrm{H_I}$, since $A_1(k_0) = 1$. Finally, note that, taking into account the contraction of the transmitted wave packet along the Ox axis, we can find its velocity:

$$V_t = \frac{\hbar k_0}{m} \times \frac{q_0}{k_0} = \frac{\hbar q_0}{m} \tag{27}$$

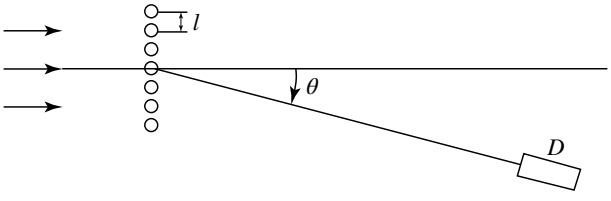
References and suggestions for further reading:

Schiff (1.18), Chap. 5, Figs. 16, 17, 18, 19; Eisberg and Resnick (1.3), § 6-3, Fig. 6-8; also see reference (1.32).

Complement K_I

Exercises

1. A beam of neutrons of mass M_n ($M_n \simeq 1{,}67\ 10^{-27}$ kg), of constant velocity and energy E, is incident on a linear chain of atomic nuclei, arranged in a regular fashion as shown in the figure (these nuclei could be, for example, those of a long linear molecule). We call l the distance between two consecutive nuclei, and d, their size ($d \ll l$). A neutron detector D is placed far away, in a direction which makes an angle of θ with the direction of the incident neutrons.



a) Describe qualitatively the phenomena observed at D when the energy E of the incident neutrons is varied.

b) The counting rate, as a function of E, presents a resonance about $E = E_1$. Knowing that there are no other resonances for $E < E_1$, show that one can determine l. Calculate l for $\theta = 30°$ and $E_1 = 1{,}3\ \ 10^{-20}$ joule.

c) At about what value of E must we begin to take the finite size of the nuclei into account?

2. Bound state of a particle in a "delta function potential"

Consider a particle whose Hamiltonian H [operator defined by formula (D-10) of Chapter I] is:

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \alpha\,\delta(x)$$

where α is a positive constant whose dimensions are to be found.

a) Integrate the eigenvalue equation of H between $-\varepsilon$ and $+\varepsilon$. Letting ε approach 0, show that the derivative of the eigenfunction $\varphi(x)$ presents a discontinuity at $x = 0$ and determine it in terms of α, m and $\varphi(0)$.

b) Assume that the energy E of the particle is negative (bound state). $\varphi(x)$ can then be written:

$$x < 0 \qquad \varphi(x) = A_1\,\mathrm{e}^{\rho x} + A_1'\,\mathrm{e}^{-\rho x}$$
$$x > 0 \qquad \varphi(x) = A_2\,\mathrm{e}^{\rho x} + A_2'\,\mathrm{e}^{-\rho x}$$

Express the constant ρ in terms of E and m. Using the results of the preceding question, calculate the matrix M defined by:

$$\begin{pmatrix} A_2 \\ A_2' \end{pmatrix} = M \begin{pmatrix} A_1 \\ A_1' \end{pmatrix}$$

Then, using the condition that $\varphi(x)$ must be square-integrable, find the possible values of the energy. Calculate the corresponding normalized wave functions.

c) Plot these wave functions on a graph. Give an order of magnitude for their width Δx.

d) What is the probability $\overline{\mathrm{d}\mathscr{P}}(p)$ that a measurement of the momentum of the particle in one of the normalized stationary states calculated above will give a result included between p and $p + \mathrm{d}p$? For what value of p is this probability maximum? In what domain, of dimension Δp, does it take on non-negligible values? Give an order of magnitude for the product $\Delta x \cdot \Delta p$.

3. Transmission of a "delta function" potential barrier

Consider a particle placed in the same potential as in the preceding exercise. The particle is now propagating from left to right along the Ox axis, with a positive energy E.

a) Show that a stationary state of the particle can be written:

$$\begin{cases} \text{if} \quad x < 0 \quad \varphi(x) = \mathrm{e}^{ikx} + A\,\mathrm{e}^{-ikx} \\ \text{if} \quad x > 0 \quad \varphi(x) = B\,\mathrm{e}^{ikx} \end{cases}$$

where k, A and B are constants which are to be calculated in terms of the energy E, of m and of α (watch out for the discontinuity in $\dfrac{\mathrm{d}\varphi}{\mathrm{d}x}$ at $x = 0$).

b) Set $-E_L = -m\alpha^2/2\hbar^2$ (bound state energy of the particle). Calculate, in terms of the dimensionless parameter E/E_L, the reflection coefficient R and the transmission coefficient T of the barrier. Study their variations with respect to E; what happens when $E \longrightarrow \infty$? How can this be interpreted? Show that, if the expression of T is extended for negative values of E, it diverges when $E \longrightarrow -E_L$, and discuss this result.

4. Return to exercise 2, using this time the Fourier transform.

a) Write the eigenvalue equation of H and the Fourier transform of this equation. Deduce directly from this the expression for $\overline{\varphi}(p)$, the Fourier transform of $\varphi(x)$, in terms of p, E, α and $\varphi(0)$. Then show that only one value of E, a negative one, is possible. Only the bound state of the particle, and not the ones in which it propagates, is found by this method; why? Then calculate $\varphi(x)$ and show that one can find in this way all the results of exercise 2.

b) The average kinetic energy of the particle can be written (*cf.* Chap. III):

$$E_k = \frac{1}{2m} \int_{-\infty}^{+\infty} p^2\, |\overline{\varphi}(p)|^2\, \mathrm{d}p$$

Show that, when $\overline{\varphi}(p)$ is a "sufficiently smooth" function, we also have:

$$E_k = -\frac{\hbar^2}{2m} \int_{-\infty}^{+\infty} \varphi^*(x)\, \frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2}\, \mathrm{d}x$$

These formulas enable us to obtain, in two different ways, the energy E_k for a particle in the bound state calculated in *a*). What result is obtained? Note that, in this case, $\varphi(x)$ is not "regular" at $x = 0$, where its derivative is discontinuous. It is then necessary to differentiate $\varphi(x)$ in the sense of distributions, which introduces a contribution of the point $x = 0$ to the average value we are looking for. Interpret this contribution physically: consider a square well, centered at $x = 0$, whose width a approaches 0 and whose depth V_0 approaches infinity (so that $aV_0 = \alpha$), and study the behavior of the wave function in this well.

5. Well consisting of two delta functions

Consider a particle of mass m whose potential energy is

$$V(x) = -\alpha\,\delta(x) - \alpha\,\delta(x - l) \qquad \alpha > 0$$

where l is a constant length.

a) Calculate the bound states of the particle, setting $E = -\dfrac{\hbar^2\rho^2}{2m}$. Show that the possible energies are given by the relation

$$\mathrm{e}^{-\rho l} = \pm\left(1 - \frac{2\rho}{\mu}\right)$$

where μ is defined by $\mu = \dfrac{2m\alpha}{\hbar^2}$. Give a graphic solution of this equation.

(*i*) *Ground state.* Show that this state is even (invariant with respect to reflection about the point $x = l/2$), and that its energy E_S is less than the energy $-E_L$ introduced in problem 3. Interpret this result physically. Represent graphically the corresponding wave function.

(*ii*) *Excited state.* Show that, when l is greater than a value to be specified, there exists an odd excited state, of energy E_A greater than $-E_L$. Find the corresponding wave function.

(*iii*) Explain how the preceding calculations enable us to construct a model which represents an ionized diatomic molecule (H_2^+, for example) whose nuclei are separated by a distance l. How do the energies of the two levels vary with respect to l? What happens at the limit where $l \longrightarrow 0$ and at the limit where $l \longrightarrow \infty$? If the repulsion of the two nuclei is taken into account, what is the total energy of the system? Show that the curve that gives the variation with respect to l of the energies thus obtained enables us to predict in certain cases the existence of bound states of H_2^+, and to determine the value of l at equilibrium. The calculation provides a very elementary model of the chemical bond.

b) Calculate the reflection and transmission coefficients of the system of two delta function barriers. Study their variations with respect to l. Do the resonances thus obtained occur when l is an integral multiple of the de Broglie wavelength of the particle? Why?

6. Consider a square well potential of width a and depth V_0 (in this exercise, we shall use systematically the notation of § 2-c-α of Complement H_I). We intend to study the properties of the bound state of a particle in this well when its width a approaches zero.

a) Show that there indeed exists only one bound state and calculate its energy E $\left(\text{we find } E \simeq -\dfrac{mV_0^2a^2}{2\hbar^2}, \text{ that is, an energy which varies with the square of the area } aV_0 \text{ of the well}\right)$.

b) Show that $\rho \longrightarrow 0$ and that $A_2 = A_2' \simeq B_1/2$. Deduce from this that, in the bound state, the probability of finding the particle outside the well approaches 1.

c) How can the preceding considerations be applied to a particle placed, as in exercise 2, in the potential $V(x) = -\alpha\,\delta(x)$?

7. Consider a particle placed in the potential

$$V(x) = 0 \quad \text{if} \quad x \geq a$$
$$V(x) = -V_0 \quad \text{if} \quad 0 \leq x < a,$$

with $V(x)$ infinite for negative x. Let $\varphi(x)$ be a wave function associated with a stationary state of the particle. Show that $\varphi(x)$ can be extended to give an odd wave function which corresponds to a stationary state for a square well of width $2a$ and depth V_0 (*cf.* Complement H_I, § 2-c-α). Discuss, with respect to a and V_0, the number of bound states of the particle. Is there always at least one such state, as for the symmetric square well?

8. Consider, in a two-dimensional problem, the oblique reflection of a particle from a potential step defined by:

$$V(x, y) = 0 \quad \text{if } x < 0$$
$$V(x, y) = V_0 \quad \text{if } x > 0$$

Study the motion of the center of the wave packet. In the case of total reflection, interpret physically the differences between the trajectory of this center and the classical trajectory (lateral shift upon reflection). Show that, when $V_0 \longrightarrow +\infty$, the quantum trajectory becomes asymptotic to the classical trajectory.

Chapter II

The mathematical tools of quantum mechanics

A Space of the one-particle wave function 88
A-1 Structure of the wave function space $\mathscr{F}$ 89
A-2 Discrete orthonormal bases in $\mathscr{F}$: $\{u_i(\mathbf{r})\}$ 91
A-3 Introduction of "bases" not belonging to $\mathscr{F}$ 94
B State space. Dirac notation 102
B-1 Introduction 102
B-2 "Ket" vectors and "bra" vectors 103
B-3 Linear operators 108
B-4 Hermitian conjugation 111
C Representations in state space 116
C-1 Introduction 116
C-2 Relations characteristic of an orthonormal basis 116
C-3 Representation of kets and bras 119
C-4 Representation of operators 121
C-5 Change of representations 124
D Eigenvalue equations. Observables 126
D-1 Eigenvalues and eigenvectors of an operator 126
D-2 Observables 130
D-3 Sets of commuting observables 133
E Two important examples of representations and observables 139
E-1 The $\{|\mathbf{r}\rangle\}$ and $\{|\mathbf{p}\rangle\}$ representations 139
E-2 The $\boldsymbol{R}$ and $\boldsymbol{P}$ operators 143
F Tensor product of state spaces 147
F-1 Introduction 147
F-2 Definition and properties of the tensor product 147
F-3 Eigenvalue equations in the product space 151
F-4 Applications 154

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

This chapter is intended to be a general survey of the basic mathematical tools used in quantum mechanics. We shall give a simple condensed presentation aimed at facilitating the study of subsequent chapters for readers unfamiliar with these tools. We make no attempt to be mathematically complete or rigorous. We feel it preferable to limit ourselves to a practical point of view, uniting in a single chapter the various concepts useful in quantum mechanics. In particular, we wish to stress the convenience of the Dirac notation for carrying out the various calculations to be performed.

In this spirit, we shall try to simplify the discussion as much as possible. Neither the general definitions nor the rigorous proofs which would be required by a mathematician are to be found here. For example, we shall sometimes speak of infinite-dimensional spaces and reason as if they had a finite number of dimensions. Moreover, many terms (square-integrable function, basis, etc...) will be employed with a meaning which, although commonly used in physics, is not exactly the one used in pure mathematics.

We begin in § A by studying the wave functions introduced in Chapter I. We show that these wave functions belong to an abstract vector space, which we call the "wave function space $\mathscr{F}$". This study will be carried out in detail as it introduces some basic concepts of the mathematical formalism of quantum mechanics: scalar products, linear operators, bases, etc... Starting in § B, we shall develop a more general formalism, characterizing the state of a system by a "state vector" belonging to a vector space: the "state space $\mathscr{E}$". Dirac notation, which greatly simplifies calculations in this formalism, is introduced. § C is intended to study the idea of a representation. § D is particularly recommended to the reader who is unfamiliar with the diagonalization of an operator: this operation will be constantly useful to us in what follows. In § E, we treat two important examples of representations. In particular, we show how the wave functions studied in § A are the "components" of state vectors in a particular representation. Finally, we introduce in § F the concept of a tensor product. This concept will be illustrated more concretely by a simple example in Complement $\mathrm{D_{IV}}$.

A. Space of the one-particle wave function

The probabilistic interpretation of the wave function $\psi(\mathbf{r}, t)$ of a particle was given in the preceding chapter: $|\psi(\mathbf{r}, t)|^2 \, \mathrm{d}^3r$ represents the probability of finding, at time t, the particle in a volume $\mathrm{d}^3r = \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$ about the point $\mathbf{r}$. As the total probability of finding the particle somewhere in space is equal to 1, we must have:

$$\int \mathrm{d}^3r \, |\psi(\mathbf{r}, t)|^2 = 1 \tag{A-1}$$

where the integration extends over all space.

Thus, we are led to studying the set of square-integrable functions. These are functions for which the integral (A-1) converges [1].

From a physical point of view, it is clear that the set L^2 is too wide in scope: given the meaning attributed to $|\psi(\mathbf{r}, t)|^2$, the wave functions that are actually used possess certain properties of regularity. We can only retain the functions $\psi(\mathbf{r}, t)$ which are everywhere defined, continuous, and infinitely differentiable (for example, to state that a function is really discontinuous at a given point in space has no physical meaning,

[1] This set is called L^2 by mathematicians and it has the structure of a Hilbert space

since no experiment enables us to have access to real phenomena on a very small scale, say of 10^{-30} m). It is also possible to confine ourselves to wave functions that have a bounded domain (which makes it certain that the particle can be found within a finite region of space, for example inside the laboratory). We shall not try to give a precise, general list of these supplementary conditions: we shall call $\mathscr{F}$ the set of wave functions composed of sufficiently regular functions of L^2 ($\mathscr{F}$ is a subspace of L^2).

A-1. Structure of the wave function space $\mathscr{F}$

A-1-a. $\mathscr{F}$ is a vector space

It can easily be shown that $\mathscr{F}$ satisfies all the criteria of a vector space. As an example, we demonstrate that if $\psi_1(\mathbf{r})$ and $\psi_2(\mathbf{r}) \in \mathscr{F}$, then[2]:

$$\psi(\mathbf{r}) = \lambda_1\psi_1(\mathbf{r}) + \lambda_2\psi_2(\mathbf{r}) \in \mathscr{F} \tag{A-2}$$

where λ_1 and λ_2 are two arbitrary complex numbers.

In order to show that $\psi(\mathbf{r})$ is square-integrable, expand $|\psi(\mathbf{r})|^2$:

$$|\psi(\mathbf{r})|^2 = |\lambda_1|^2|\psi_1(\mathbf{r})|^2 + |\lambda_2|^2|\psi_2(\mathbf{r})|^2 + \lambda_1^*\lambda_2\,\psi_1^*(\mathbf{r})\psi_2(\mathbf{r}) + \lambda_1\lambda_2^*\,\psi_1(\mathbf{r})\psi_2^*(\mathbf{r}) \tag{A-3}$$

The first two terms of the right-hand side of (A-3) are square integrable, since ψ_1 and ψ_2 belong to $\mathscr{F}$. The sum of the third and fourth terms is real, and its modulus has an upper limit:

$$2|\lambda_1||\lambda_2||\psi_1(\mathbf{r})||\psi_2(\mathbf{r})| \leq |\lambda_1||\lambda_2|\left[|\psi_1(\mathbf{r})|^2 + |\psi_2(\mathbf{r})|^2\right]$$

(this inequality is obtained by writing that $[\psi_1(\mathbf{r}) + \mathrm{e}^{i\alpha}\psi_2(\mathbf{r})]^2$ is necessarily positive, whatever the real value of α). The function $|\psi(\mathbf{r})|^2$ is therefore smaller that a sum of functions whose integrals converge, so that its integral also converges.

A-1-b. The scalar product

α. Definition

With each pair of elements of $\mathscr{F}$, $\varphi(\mathbf{r})$ and $\psi(\mathbf{r})$, taken in this order, we associate a *complex number*, denoted by (φ, ψ), which, by definition, is equal to:

$$\boxed{(\varphi, \psi) = \int \mathrm{d}^3r\ \varphi^*(\mathbf{r})\ \psi(\mathbf{r})} \tag{A-4}$$

(φ, ψ) is the *scalar product of* $\psi(\mathbf{r})$ *by* $\varphi(\mathbf{r})$ [this integral always converges if φ and ψ belong to $\mathscr{F}$].

β. Properties

They follow from definition (A-4):

$$(\varphi, \psi) = (\psi, \varphi)^* \tag{A-5}$$

$$(\varphi, \lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1(\varphi, \psi_1) + \lambda_2(\varphi, \psi_2) \tag{A-6}$$

$$(\lambda_1\varphi_1 + \lambda_2\varphi_2, \psi) = \lambda_1^*(\varphi_1, \psi) + \lambda_2^*(\varphi_2, \psi) \tag{A-7}$$

[2]The symbol $\in$ signifies: "belongs to".

The scalar product is *linear* with respect to the second function of the pair, *antilinear* with respect to the first one. If $(\varphi, \psi) = 0$, $\varphi(\mathbf{r})$ and $\psi(\mathbf{r})$ are said to be *orthogonal*.

$$\boxed{(\psi, \psi) = \int \mathrm{d}^3 r \; |\psi(\mathbf{r})|^2} \tag{A-8}$$

is a *real, positive* number, which is *zero if and only if* $\psi(\mathbf{r}) \equiv 0$.

$\sqrt{(\psi, \psi)}$ is called the norm of $\psi(\mathbf{r})$ [it can easily be verified that this number has all the properties of a norm]. The scalar product chosen above thus permits the definition of a norm in $\mathscr{F}$.

Let us finally mention the *Schwarz inequality* (*cf.* Complement $\mathrm{A_{II}}$):

$$|(\psi_1, \psi_2)| \leqslant \sqrt{(\psi_1, \psi_1)} \; \sqrt{(\psi_2, \psi_2)} \tag{A-9}$$

This becomes an equality if and only if the two functions ψ_1 and ψ_2 are proportional.

A-1-c. Linear operators

α. Definition

A linear operator A is, by definition, a mathematical entity which associates with every function $\psi(\mathbf{r}) \in \mathscr{F}$ another function $\psi'(\mathbf{r})$, the correspondence being linear:

$$\psi'(\mathbf{r}) = A\psi(\mathbf{r}) \tag{A-10-a}$$

$$A[\lambda_1 \psi_1(\mathbf{r}) + \lambda_2 \psi_2(\mathbf{r})] = \lambda_1 A\psi_1(\mathbf{r}) + \lambda_2 A\psi_2(\mathbf{r}) \tag{A-10-b}$$

Let us cite some simple examples of linear operators:

– the parity operator Π, whose definition is:

$$\Pi\psi(x, y, z) = \psi(-x, -y, -z) \tag{A-11}$$

– the operator that performs a multiplication by x, which we shall call X, and which is defined by:

$$X\psi(x, y, z) = x \; \psi(x, y, z) \tag{A-12}$$

– finally, the operator that we shall call D_x, which differentiates with respect to x, and whose definition is:

$$D_x\psi(x, y, z) = \frac{\partial \psi(x, y, z)}{\partial x} \tag{A-13}$$

[the two operators X and D_x, acting on a function $\psi(\mathbf{r}) \in \mathscr{F}$, can transform it into a function which is no longer necessarily square-integrable].

β. Product of operators

Let A and B be two linear operators. Their product AB is defined by:

$$\boxed{(AB)\psi(\mathbf{r}) = A\,[B\psi(\mathbf{r})]} \tag{A-14}$$

B is first allowed to act on $\psi(\mathbf{r})$, which gives $\varphi(\mathbf{r}) = B\psi(\mathbf{r})$, then A operates on the new function $\varphi(\mathbf{r})$.

In general, $AB \neq BA$. We call the *commutator* of A and B the operator written $[A, B]$ and defined by

$$\boxed{[A, B] = AB \ - \ BA} \tag{A-15}$$

We shall calculate, as an example, the commutator $[X, D_x]$. In order to do this, we shall take an arbitrary function $\psi(\mathbf{r})$:

$$\begin{aligned}[X, D_x] \ \psi(\mathbf{r}) &= \left(x\frac{\partial}{\partial x} - \frac{\partial}{\partial x}x\right)\psi(\mathbf{r}) \\ &= x\frac{\partial}{\partial x}\,\psi(\mathbf{r}) - \frac{\partial}{\partial x}\,[x\psi(\mathbf{r})] \\ &= x\frac{\partial}{\partial x}\,\psi(\mathbf{r}) \ - \ \psi(\mathbf{r}) \ - \ x\frac{\partial}{\partial x}\,\psi(\mathbf{r}) = \ -\,\psi(\mathbf{r})\end{aligned} \tag{A-16}$$

Since this is true for all $\psi(\mathbf{r})$, it can be deduced that:

$$[X, D_x] = \ -\,1 \tag{A-17}$$

A-2. Discrete orthonormal bases in $\mathscr{F}$: $\{u_i(\mathbf{r})\}$

A-2-a. Definition

Consider a countable set of functions of $\mathscr{F}$, labeled by a discrete index i ($i = 1, 2, ..., n, ...$):

$$u_1(\mathbf{r}) \in \mathscr{F}, \ \ u_2(\mathbf{r}) \in \mathscr{F}, \ \ ..., \ \ u_i(\mathbf{r}) \in \mathscr{F}, \ \ ...$$

– The set $\{u_i(\mathbf{r})\}$ is *orthonormal* if:

$$(u_i, u_j) = \int \mathrm{d}^3r \ \ u_i^*(\mathbf{r}) \ u_j(\mathbf{r}) = \delta_{ij} \tag{A-18}$$

where δ_{ij}, the Kronecker delta function, is equal to 1 for $i = j$ and to 0 for $i \neq j$.

– It constitutes a basis[3] if every function $\psi(\mathbf{r}) \in \mathscr{F}$ can be expanded in one and only one way in terms of the $\{\, u_i(\mathbf{r}) \,\}$:

$$\boxed{\psi(\mathbf{r}) = \sum_i \, c_i \, u_i(\mathbf{r})} \tag{A-19}$$

A-2-b. Components of a wave function in the $\{\, u_i(\mathbf{r}) \,\}$ basis

Multiply the two sides of (A-19) by $u_j^*(\mathbf{r})$ and integrate over all space. From (A-6) and (A-18)[4]:

[3] When the set $\{u_i(\mathbf{r})\}$ constitutes a basis, it is sometimes said to be a complete set of functions. It must be noted that the word complete is used with a meaning different from the one it usually has in mathematics.

[4] To be totally rigorous, one should make certain that one can interchange $\sum_i$ and $\int \mathrm{d}^3r$. We shall systematically ignore this kind of problem.

$$(u_j, \psi) = \left(u_j, \sum_i c_i \, u_i \right) = \sum_i c_i (u_j, u_i)$$
$$= \sum_i c_i \, \delta_{ij} = c_j \tag{A-20}$$

that is:

$$\boxed{c_i = (u_i, \psi) = \int \mathrm{d}^3 r \;\; u_i^*(\mathbf{r}) \, \psi(\mathbf{r})} \tag{A-21}$$

The component c_i of $\psi(\mathbf{r})$ on $u_i(\mathbf{r})$ is therefore equal to the scalar product of $\psi(\mathbf{r})$ by $u_i(\mathbf{r})$. Once the $\{ u_i(\mathbf{r}) \}$ basis has been chosen, it is equivalent to specify $\psi(\mathbf{r})$ or the set of its components c_i with respect to the basis functions. The set of numbers c_i is said to *represent* $\psi(\mathbf{r})$ in the $\{ u_i(\mathbf{r}) \}$ basis.

Comments:

(i) Note the analogy with an orthonormal basis $\{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \}$ of the ordinary three-dimensional space, R^3. The fact that $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ are orthogonal and unitary can indeed be expressed by:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \qquad (i, j = 1, 2, 3) \tag{A-22}$$

Any vector $\mathbf{V}$ of R^3 can be expanded in this basis:

$$\mathbf{V} = \sum_{i=1}^{3} v_i \, \mathbf{e}_i \tag{A-23}$$

with

$$v_i = \mathbf{e}_i \cdot \mathbf{V} \tag{A-24}$$

Formulas (A-18), (A-19) and (A-21) thus generalize, as it were, the well-known formulas, (A-22), (A-23) and (A-24). However, it must be noted that the v_i are real numbers, while the c_i are complex numbers.

(*ii*) The same function $\psi(\mathbf{r})$ obviously has different components in two different bases. We shall study the problem of a change in basis later.

(*iii*) We can also, in the $\{ u_i(\mathbf{r}) \}$ basis, represent a linear operator A by a set of numbers which can be arranged in the form of a matrix. We shall take up this question again in § C, after we have introduced Dirac notation.

A-2-c. Expression for the scalar product in terms of the components

Let $\varphi(\mathbf{r})$ and $\psi(\mathbf{r})$ be two wave functions which can be expanded as follows:

$$\varphi(\mathbf{r}) = \sum_i b_i \, u_i(\mathbf{r})$$
$$\psi(\mathbf{r}) = \sum_j c_j \, u_j(\mathbf{r}) \tag{A-25}$$

Their scalar product can be calculated by using (A-6), (A-7) and (A-18):

$$(\varphi,\psi) = \left(\sum_i b_i u_i, \sum_j c_j u_j\right) = \sum_{i,j} b_i^* c_j (u_i, u_j)$$
$$= \sum_{i,j} b_i^* c_j \, \delta_{ij}$$

that is:

$$\boxed{(\varphi,\psi) = \sum_i b_i^* c_i} \tag{A-26}$$

In particular:

$$\boxed{(\psi,\psi) = \sum_i |c_i|^2} \tag{A-27}$$

The scalar product of two wave functions (or the square of the norm of a wave function) can thus be very simply expressed in terms of the components of these functions in the $\{\, u_i(\mathbf{r}) \,\}$ basis.

Comment:

Let $\mathbf{V}$ and $\mathbf{W}$ be two vectors of R^3, with components v_i and w_j. The analytic expression of their scalar product is well-known:

$$\mathbf{V} \cdot \mathbf{W} = \sum_{i=1}^{3} v_i \, w_i \tag{A-28}$$

Formula (A-26) can therefore be considered to be a generalization of (A-28).

A-2-d. Closure relation

Relation (A-18), called the orthonormalization relation, expresses the fact that the functions of the set $\{\, u_i(\mathbf{r}) \,\}$ are normalized to 1 and orthogonal with respect to each other. We are now going to establish another relation, called the closure relation, which expresses the fact that this set constitutes a basis.

If $\{\, u_i(\mathbf{r}) \,\}$ is a basis of $\mathscr{F}$, there exists an expansion such as (A-19) for every function $\psi(\mathbf{r}) \in \mathscr{F}$. Substitute into (A-19) expression (A-21) for the various components c_i [the name of the integration variable must be changed, since $\mathbf{r}$ already appears in (A-19)]:

$$\psi(\mathbf{r}) = \sum_i c_i \, u_i(\mathbf{r}) = \sum_i (u_i, \psi) \, u_i(\mathbf{r})$$
$$= \sum_i \left[\int \mathrm{d}^3 r' \, u_i^*(\mathbf{r}') \, \psi(\mathbf{r}')\right] u_i(\mathbf{r}) \tag{A-29}$$

Interchanging $\sum_i$ and $\int \mathrm{d}^3r'$, we obtain:

$$\psi(\mathbf{r}) = \int \mathrm{d}^3r' \ \psi(\mathbf{r}') \left[\sum_i u_i(\mathbf{r})\, u_i^*(\mathbf{r}') \right] \tag{A-30}$$

$\sum_i u_i(\mathbf{r})\, u_i^*(\mathbf{r}')$ is therefore a function $F(\mathbf{r}, \mathbf{r}')$ of $\mathbf{r}$ and of $\mathbf{r}'$ such that, for every function $\psi(\mathbf{r})$, we have:

$$\psi(\mathbf{r}) = \int \mathrm{d}^3r' \ \psi(\mathbf{r}')\, F(\mathbf{r}, \mathbf{r}') \tag{A-31}$$

Equation (A-31) is characteristic of the function $\delta(\mathbf{r} - \mathbf{r}')$ (*cf.* Appendix II). From this it can be deduced that:

$$\boxed{\sum_i u_i(\mathbf{r})\, u_i^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')} \tag{A-32}$$

Reciprocally, if an orthonormal set $\{ u_i(\mathbf{r}) \}$ satisfies the closure relation (A-32), it constitutes a basis. Any function $\psi(\mathbf{r})$ can indeed be written in the form:

$$\psi(\mathbf{r}) = \int \mathrm{d}^3r' \ \psi(\mathbf{r}')\, \delta(\mathbf{r} - \mathbf{r}') \tag{A-33}$$

Substituting (A-32) for $\delta(\mathbf{r}-\mathbf{r}')$ into this expression, we obtain formula (A-30). To return to (A-29), all we must do is again interchange summation and integration. This equation then expresses the fact that $\psi(\mathbf{r})$ can always be expanded in terms of the $u_i(\mathbf{r})$ and gives the coefficients of this expansion.

Comment:

We shall re-examine the closure relation using the Dirac notation in § C, and we shall see that it can be given a simple geometric interpretation.

A-3. Introduction of "bases" not belonging to $\mathscr{F}$

The $\{ u_i(\mathbf{r}) \}$ bases studied above are composed of square-integrable functions. It can also be convenient to introduce "bases" of functions not belonging to either $\mathscr{F}$ or L^2, but in terms of which any wave function $\psi(\mathbf{r})$ can nevertheless be expanded. We are going to give examples of such bases and we shall show how it is possible to extend to them the important formulas established in the preceding section.

A-3-a. Plane waves

For simplicity, we treat the one-dimensional case. We shall therefore study square-integrable functions $\psi(x)$ which depend only on the x variable. In Chapter I we saw the

advantage of using the Fourier transform $\overline{\psi}(p)$ of $\psi(x)$:

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}p \;\; \overline{\psi}(p)\, \mathrm{e}^{ipx/\hbar} \tag{A-34-a}$$

$$\overline{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}x \;\; \psi(x)\, \mathrm{e}^{-ipx/\hbar} \tag{A-34-b}$$

Consider the function $v_p(x)$, defined by:

$$\boxed{v_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\, \mathrm{e}^{ipx/\hbar}} \tag{A-35}$$

$v_p(x)$ is a plane wave, with the wave vector $p/\hbar$. The integral over the whole x axis of $|v_p(x)|^2 = \dfrac{1}{2\pi\hbar}$ diverges. Therefore $v_p(x) \notin \mathscr{F}_x$. We shall designate by $\{\, v_p(x) \,\}$ the set of all plane waves, that is, of all functions $v_p(x)$ corresponding to the various values of p. The number p, which varies continuously between $-\infty$ and $+\infty$, will be considered as a *continuous index* which permits us to label the various functions of the set $\{\, v_p(x) \,\}$ [recall that the index i used for the set $\{\, u_i(\mathbf{r}) \,\}$ considered above was discrete].

Formulas (A-34) can be rewritten using (A-35):

$$\psi(x) = \int_{-\infty}^{+\infty} \mathrm{d}p \;\; \overline{\psi}(p)\; v_p(x) \tag{A-36}$$

$$\overline{\psi}(p) = (v_p, \psi) = \int_{-\infty}^{+\infty} \mathrm{d}x \;\; v_p^*(x) \;\; \psi(x) \tag{A-37}$$

These two formulas can be compared to (A-19) and (A-21). Relation (A-36) expresses the idea that every function $\psi(x) \in \mathscr{F}_x$ can be expanded in one and only one way in terms of the $v_p(x)$, that is, the plane waves. Since the index p varies continuously and not discretely, the summation $\sum_i$ appearing in (A-19) must be replaced by an integration over p. Relation (A-37), like (A-21), gives the component $\overline{\psi}(p)$ of $\psi(x)$ on $v_p(x)$ in the form of a scalar product[5] (v_p, ψ). The set of these components, which correspond to the various possible values of p, constitutes a function of p, $\overline{\psi}(p)$, the Fourier transform of $\psi(x)$.

Thus, $\overline{\psi}(p)$ is the analogue of c_i. These two complex numbers, which depend either on p or on i, represent the *components of the same function $\psi(x)$ in two different bases*: $\{\, v_p(x) \,\}$ *and* $\{\, u_i(x) \,\}$.

This point also appears clearly if we calculate the square of the norm of $\psi(x)$. According to Parseval's relation [Appendix I, formula (45)], we have:

$$\boxed{(\psi, \psi) = \int_{-\infty}^{+\infty} \mathrm{d}p\; |\overline{\psi}(p)|^2} \tag{A-38}$$

[5] We have only defined the scalar product for two square-integrable functions, but this definition can easily be extended to cases like this one, provided that the corresponding integral converges.

a formula which resembles (A-27), if we replace c_i by $\overline{\psi}(p)$ and $\sum_i$ by $\int \mathrm{d}p$.

Let us show that the $v_p(x)$ satisfy a closure relation. Using the formula [*cf.* Appendix II, equation (34)]:

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}k\ \mathrm{e}^{iku} = \delta(u) \tag{A-39}$$

we find:

$$\boxed{\int_{-\infty}^{+\infty} \mathrm{d}p\ v_p(x)\ v_p^*(x') = \frac{1}{2\pi}\int \frac{\mathrm{d}p}{\hbar}\ \mathrm{e}^{i\frac{p}{\hbar}(x-x')} = \delta(x-x')} \tag{A-40}$$

This formula is the analogue of (A-32) with, again, the substitution of $\int \mathrm{d}p$ for $\sum_i$.

Finally, let us calculate the scalar product $(v_p, v_{p'})$ in order to see if there exists an equivalent of the orthonormalization relation. Again using (A-39), we obtain:

$$(v_p, v_{p'}) = \int_{-\infty}^{+\infty} \mathrm{d}x\ v_p^*(x)\ v_{p'}(x)$$

that is:

$$\boxed{(v_p, v_{p'}) = \frac{1}{2\pi}\int \frac{\mathrm{d}x}{\hbar}\ \mathrm{e}^{i\frac{x}{\hbar}(p'-p)} = \delta(p-p')} \tag{A-41}$$

Compare (A-41) and (A-18). Instead of having two discrete indices i and j and a Kronecker delta δ_{ij}, we now have two continuous indices p and p' and a delta function of the difference between the indices, $\delta(p-p')$. Note that if we set $p = p'$, the scalar product (v_p, v_p) diverges; again we see that $v_p(x) \notin \mathscr{F}_x$. Although this constitutes a misuse of the term, we shall call (A-41) an "orthonormalization" relation. It is also sometimes said that the $v_p(x)$ are "orthonormalized in the Dirac sense".

The generalization to three dimensions presents no difficulties. We consider the plane waves:

$$v_{\mathbf{p}}(\mathbf{r}) = \left(\frac{1}{2\pi\hbar}\right)^{3/2} \mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \tag{A-42}$$

The functions of the $\{\, v_p(\mathbf{r}) \,\}$ basis now depend on the three continuous indices p_x, p_y, p_z, condensed into the notation $\mathbf{p}$. It is then easy to show that:

$$\psi(\mathbf{r}) = \int \mathrm{d}^3p\ \ \overline{\psi}(\mathbf{p})\ \ v_{\mathbf{p}}(\mathbf{r}) \tag{A-43}$$

$$\overline{\psi}(\mathbf{p}) = (v_{\mathbf{p}}, \psi) = \int \mathrm{d}^3r\ \ v_{\mathbf{p}}^*(\mathbf{r})\ \psi(\mathbf{r}) \tag{A-44}$$

$$(\varphi, \psi) = \int \mathrm{d}^3p\ \ \overline{\varphi}^*(\mathbf{p})\ \overline{\psi}(\mathbf{p}) \tag{A-45}$$

$$\int \mathrm{d}^3p\ \ v_{\mathbf{p}}(\mathbf{r})\ v_{\mathbf{p}}^*(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \tag{A-46}$$

$$(v_{\mathbf{p}}, v_{\mathbf{p}'}) = \delta(\mathbf{p} - \mathbf{p}') \tag{A-47}$$

They represent the generalizations of (A-36), (A-37), (A-38), (A-40) and (A-41).

Thus the $v_p(\mathbf{r})$ can be considered to constitute a "continuous basis". All the formulas established above for the discrete basis $\{ u_i(\mathbf{r}) \}$ can be extended to this continuous basis, using the correspondence rules summarized in table (II-1).

$$\boxed{\begin{array}{c} i \leftrightarrow \mathbf{p} \\ \sum_i \leftrightarrow \int \mathrm{d}^3 p \\ \delta_{ij} \leftrightarrow \delta(\mathbf{p} - \mathbf{p}') \end{array}}$$

Table (II-1)

A-3-b. "Delta functions"

In the same way, let us introduce a set of functions of $\mathbf{r}$, $\{ \xi_{\mathbf{r}_0}(\mathbf{r}) \}$, labeled by the continuous index $\mathbf{r}_0$ (condensed notation for x_0, y_0, z_0) and defined by:

$$\boxed{\xi_{\mathbf{r}_0}(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)} \tag{A-48}$$

$\{ \xi_{\mathbf{r}_0}(\mathbf{r}) \}$ represents the set of delta functions centered at the various points $\mathbf{r}_0$ of space; $\xi_{\mathbf{r}_0}(\mathbf{r})$ is obviously not square-integrable: $\xi_{\mathbf{r}_0}(\mathbf{r}) \notin \mathscr{F}$.

Then consider the following relations, which are valid for every function $\psi(\mathbf{r}) \in \mathscr{F}$:

$$\psi(\mathbf{r}) = \int \mathrm{d}^3 r_0 \ \psi(\mathbf{r}_0)\, \delta(\mathbf{r} - \mathbf{r}_0) \tag{A-49}$$

$$\psi(\mathbf{r}_0) = \int \mathrm{d}^3 r \ \delta(\mathbf{r}_0 - \mathbf{r})\, \psi(\mathbf{r}) \tag{A-50}$$

They can be rewritten, using (A-48), in the form:

$$\psi(\mathbf{r}) = \int \mathrm{d}^3 r_0 \ \psi(\mathbf{r}_0)\, \xi_{\mathbf{r}_0}(\mathbf{r}) \tag{A-51}$$

$$\psi(\mathbf{r}_0) = (\xi_{\mathbf{r}_0}, \psi) = \int \mathrm{d}^3 r \ \xi^*_{\mathbf{r}_0}(\mathbf{r})\, \psi(\mathbf{r}) \tag{A-52}$$

(A-51) expresses the fact that every function $\psi(\mathbf{r}) \in \mathscr{F}$ can be expanded in one and only one way in terms of the $\xi_{\mathbf{r}_0}(\mathbf{r})$. (A-52) shows that the component of $\psi(\mathbf{r})$ on the function $\xi_{\mathbf{r}_0}(\mathbf{r})$ (we are dealing here with real basis functions) is precisely the value $\psi(\mathbf{r}_0)$ of $\psi(\mathbf{r})$ at the point $\mathbf{r}_0$. (A-51) and (A-52) are analogous to (A-19) and (A-21): we simply replace the discrete index i by the continuous index $\mathbf{r}_0$, and $\sum_i$ by $\int \mathrm{d}^3 r_0$.

$\psi(\mathbf{r}_0)$ *is therefore the equivalent of* c_i: these two complex numbers, which depend either on $\mathbf{r}_0$ or on i, represent the components of the same function $\psi(\mathbf{r})$ in two different bases: $\{ \xi_{\mathbf{r}_0}(\mathbf{r}) \}$ and $\{ u_i(\mathbf{r}) \}$.

Formula (A-26) becomes here:

$$\boxed{(\varphi, \psi) = \int \mathrm{d}^3 r_0 \;\; \varphi^*(\mathbf{r}_0)\, \psi(\mathbf{r}_0)} \tag{A-53}$$

We see that the application of (A-26) to the case of the continuous basis $\{ \xi_{\mathbf{r}_0}(\mathbf{r}) \}$ results in the definition (A-4) of the scalar product.

Finally, note that the $\xi_{\mathbf{r}_0}(\mathbf{r})$ satisfy "orthonormalization" and closure relations of the same type as those for the $v_{\mathbf{p}}(\mathbf{r})$. Thus we have [formula (28) of Appendix II]:

$$\boxed{\int \mathrm{d}^3 r_0 \;\; \xi_{\mathbf{r}_0}(\mathbf{r})\, \xi^*_{\mathbf{r}_0}(\mathbf{r}') = \int \mathrm{d}^3 r_0 \;\; \delta(\mathbf{r} - \mathbf{r}_0)\, \delta(\mathbf{r}' - \mathbf{r}_0) = \delta(\mathbf{r} - \mathbf{r}')} \tag{A-54}$$

and:

$$\boxed{(\xi_{\mathbf{r}_0}, \xi_{\mathbf{r}'_0}) = \int \mathrm{d}^3 r \;\; \delta(\mathbf{r} - \mathbf{r}_0)\, \delta(\mathbf{r} - \mathbf{r}'_0) = \delta(\mathbf{r}_0 - \mathbf{r}'_0)} \tag{A-55}$$

All the formulas established for the discrete basis $\{ u_i(\mathbf{r}) \}$ can be generalized for the continuous basis $\{ \xi_{\mathbf{r}_0}(\mathbf{r}) \}$, using the correspondence rules summarized in Table (II-2).

$$\boxed{\begin{array}{c} i \leftrightarrow \mathbf{r}_0 \\ \displaystyle\sum_i \leftrightarrow \int \mathrm{d}^3 r_0 \\ \delta_{ij} \leftrightarrow \delta(\mathbf{r}_0 - \mathbf{r}'_0) \end{array}}$$

Table (II-2)

Important comment:

The usefulness of the continuous bases that we have just introduced is revealed more clearly in what follows. However, we must not lose sight of the following point: *a physical state must always correspond to a square-integrable wave function.* In no case can $v_{\mathbf{p}}(\mathbf{r})$ or $\xi_{\mathbf{r}_0}(\mathbf{r})$ represent the state of a particle. These functions are nothing more than intermediaries, very useful in calculations involving operations on the wave functions $\psi(\mathbf{r})$ which are used to describe a physical state.

An analogous situation is encountered in classical optics, where the plane monochromatic wave is a mathematically very useful, but physically unrealizable, idealization. Even the most selective filters always permit the passage of a frequency band $\Delta\nu$, which may be very small but is never exactly zero.

The same holds true for the functions $\xi_{\mathbf{r}_0}(\mathbf{r})$. We can imagine a square-integrable wave function, localized about $\mathbf{r}_0$, for example:

$$\xi^{(\varepsilon)}_{\mathbf{r}_0}(\mathbf{r}) = \delta^{(\varepsilon)}(\mathbf{r} - \mathbf{r}_0) = \delta^{(\varepsilon)}(x - x_0)\delta^{(\varepsilon)}(y - y_0)\delta^{(\varepsilon)}(z - z_0)$$

where the $\delta^{(\varepsilon)}$ are functions which have a peak of width ε and amplitude $\dfrac{1}{\varepsilon}$, centered at x_0, y_0 or z_0, such that $\displaystyle\int_{-\infty}^{+\infty} \delta^{(\varepsilon)}(x - x_0)\, \mathrm{d}x = 1$ (see § 1-b of Appendix II

for examples of such functions). When $\varepsilon \longrightarrow 0$, $\xi_{\mathbf{r}_0}^{(\varepsilon)}(\mathbf{r}) \longrightarrow \xi_{\mathbf{r}_0}(\mathbf{r})$, which is no longer square-integrable. However, it is impossible to have a physical state that corresponds to this limit: as localized as the physical state of a particle may be, ε is never exactly zero.

A-3-c. Generalization: continuous "orthonormal" bases

α. *Definition*

Generalizing the results obtained in the two preceding paragraphs, we shall call a *continuous* "*orthonormal*" basis, a set of functions of $\mathbf{r}$, $\{ w_\alpha(\mathbf{r}) \}$, labeled by a continuous index α, which satisfy the two following relations, called *orthonormalization* and *closure* relations:

$$(w_\alpha, w_{\alpha'}) = \int \mathrm{d}^3 r \;\; w_\alpha^*(\mathbf{r})\; w_{\alpha'}(\mathbf{r}) = \delta(\alpha - \alpha') \tag{A-56}$$

$$\int \mathrm{d}\alpha \;\; w_\alpha(\mathbf{r})\; w_\alpha^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \tag{A-57}$$

Comments:

(i) If $\alpha = \alpha'$, (w_α, w_α) diverges. Therefore, $w_\alpha(\mathbf{r}) \notin \mathscr{F}$.

(*ii*) α can represent several indices, as is the case for $\mathbf{r}_0$ and $\mathbf{p}$ in the above examples.

(*iii*) It is possible to imagine a basis that includes both functions $u_i(\mathbf{r})$, labeled by a discrete index, and functions $w_\alpha(\mathbf{r})$, labeled by a continuous index. In this case, the set of $u_i(\mathbf{r})$ does not form a basis; the set of $w_\alpha(\mathbf{r})$ must be added to it.

Let us cite an example of this situation. Consider the case of the square well studied in § D-2-c of Chapter I (see also Complement $\mathrm{H_I}$). As we shall see later, the set of stationary states of a particle in a time-independent potential constitutes a basis. For $E < 0$, we have discrete energy levels, to which correspond square-integrable wave functions labeled by a discrete index. But these are not the only possible stationary states. Equation (D-17) of Chapter I is also satisfied, for all $E > 0$, by solutions which are bounded but which extend over all space and are thus not square-integrable.

In the case of a "mixed" (discrete and continuous) basis, $\{ u_i(\mathbf{r}), w_\alpha(\mathbf{r}) \}$, the orthonormalization relations are:

$$\begin{aligned} &(u_i, u_j) = \delta_{ij} \\ &(w_\alpha, w_{\alpha'}) = \delta(\alpha - \alpha') \\ &(u_i, w_\alpha) = 0 \end{aligned} \tag{A-58}$$

And the closure relation becomes:

$$\sum_i u_i(\mathbf{r})\; u_i^*(\mathbf{r}') \;+ \int \mathrm{d}\alpha \;\; w_\alpha(\mathbf{r}) \;\; w_\alpha^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \tag{A-59}$$

β. Components of a wave function $\psi(\mathbf{r})$

We can always write:

$$\psi(\mathbf{r}) = \int \mathrm{d}^3 r' \;\; \psi(\mathbf{r}') \;\; \delta(\mathbf{r} - \mathbf{r}') \tag{A-60}$$

Using the expression for $\delta(\mathbf{r} - \mathbf{r}')$ given by (A-57), and assuming that we can reverse the order of $\int \mathrm{d}^3 r'$ and $\int \mathrm{d}\alpha$, we obtain:

$$\psi(\mathbf{r}) = \int \mathrm{d}\alpha \left[\int \mathrm{d}^3 r' \;\; w_\alpha^*(\mathbf{r}') \;\; \psi(\mathbf{r}') \right] w_\alpha(\mathbf{r}) \tag{A-61}$$

that is:

$$\boxed{\psi(\mathbf{r}) = \int \mathrm{d}\alpha \;\; c(\alpha) \;\; w_\alpha(\mathbf{r})} \tag{A-62}$$

with:

$$\boxed{c(\alpha) = (w_\alpha, \psi) = \int \mathrm{d}^3 r' \;\; w_\alpha^*(\mathbf{r}') \;\; \psi(\mathbf{r}')} \tag{A-63}$$

(A-62) expresses the fact that every wave function $\psi(\mathbf{r})$ has a unique expansion in terms of the $w_\alpha(\mathbf{r})$. The component $c(\alpha)$ of $\psi(\mathbf{r})$ on $w_\alpha(\mathbf{r})$ is equal, according to (A-63), to the scalar product (w_α, ψ).

γ. Expression for the scalar product and the norm in terms of the components

Let $\varphi(\mathbf{r})$ and $\psi(\mathbf{r})$ be two square-integrable functions whose components in terms of the $w_\alpha(\mathbf{r})$ are known:

$$\varphi(\mathbf{r}) = \int \mathrm{d}\alpha \;\; b(\alpha) \;\; w_\alpha(\mathbf{r}) \tag{A-64}$$

$$\psi(\mathbf{r}) = \int \mathrm{d}\alpha' \;\; c(\alpha') \;\; w_{\alpha'}(\mathbf{r}) \tag{A-65}$$

Calculate their scalar product:

$$\begin{aligned} (\varphi, \psi) &= \int \mathrm{d}^3 r \;\; \varphi^*(\mathbf{r}) \;\; \psi(\mathbf{r}) \\ &= \int \mathrm{d}\alpha \int \mathrm{d}\alpha' \;\; b^*(\alpha) \;\; c(\alpha') \int \mathrm{d}^3 r \;\; w_\alpha^*(\mathbf{r}) \;\; w_{\alpha'}(\mathbf{r}) \end{aligned} \tag{A-66}$$

The last integral is given by (A-56):

$$(\varphi, \psi) = \int \mathrm{d}\alpha \int \mathrm{d}\alpha' \;\; b^*(\alpha) \;\; c(\alpha') \;\; \delta(\alpha - \alpha')$$

that is:

$$\boxed{(\varphi, \psi) = \int \mathrm{d}\alpha \ b^*(\alpha) \ c(\alpha)} \tag{A-67}$$

In particular:

$$\boxed{(\psi, \psi) = \int \mathrm{d}\alpha \ |c(\alpha)|^2} \tag{A-68}$$

All the formulas of § A-2 can thus be generalized, using the correspondence rules of table (II-3).

$$\boxed{\begin{array}{c} i \leftrightarrow \alpha \\ \sum_i \leftrightarrow \int \mathrm{d}\alpha \\ \delta_{ij} \leftrightarrow \delta(\alpha - \alpha') \end{array}}$$

Table (II-3)

The most important formulas established in this section are assembled in table (II-4). Actually, it is not necessary to remember them in this form: we shall see that the introduction of Dirac notation enables us to rederive them very simply.

Table (II-4)

	Discrete basis $\{ u_i(\mathbf{r}) \}$	Continuous basis $\{ w_\alpha(\mathbf{r}) \}$
Ortho-normalization relation	$(u_i, u_j) = \delta_{ij}$	$(w_\alpha, w_{\alpha'}) = \delta(\alpha - \alpha')$
Closure relation	$\sum_i u_i(\mathbf{r})\, u_i^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')$	$\int \mathrm{d}\alpha\ w_\alpha(\mathbf{r})\, w_\alpha^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')$
Expansion of a wave function $\psi(\mathbf{r})$	$\psi(\mathbf{r}) = \sum_i c_i\, u_i(\mathbf{r})$	$\psi(\mathbf{r}) = \int \mathrm{d}\alpha\ c(\alpha)\ w_\alpha(\mathbf{r})$
Expression for the components of $\psi(\mathbf{r})$	$c_i = (u_i, \psi) = \int \mathrm{d}^3 r\, u_i^*(\mathbf{r})\, \psi(\mathbf{r})$	$c(\alpha) = (w_\alpha, \psi) = \int \mathrm{d}^3 r\, w_\alpha^*(\mathbf{r})\, \psi(\mathbf{r})$
Scalar product	$(\varphi, \psi) = \sum_i b_i^* c_i$	$(\varphi, \psi) = \int \mathrm{d}\alpha\ b^*(\alpha)\ c(\alpha)$
Square of the norm	$(\psi, \psi) = \sum_i \lvert c_i \rvert^2$	$(\psi, \psi) = \int \mathrm{d}\alpha\ \lvert c(\alpha) \rvert^2$

B. State space. Dirac notation

B-1. Introduction

In Chapter I, we stated the following postulate: the quantum state of a particle is defined, at a given instant, by a wave function $\psi(\mathbf{r})$. The probabilistic interpretation of this wave function requires that it be square-integrable. This requirement led us to study the $\mathscr{F}$-space (§ A). We then found, in particular, that the same function $\psi(\mathbf{r})$ can be represented by several distinct sets of components, each one corresponding to the choice of a basis [table (II-5)]. This result can be interpreted in the following manner: $\{ c_i \}$, or $\overline{\psi}(\mathbf{p})$, or $c(\alpha)$, characterizes the state of a particle just as well as the wave function $\psi(\mathbf{r})$ [if the basis being used has been specified previously]. Furthermore, $\psi(\mathbf{r})$ itself appears, in table (II-5), on the same footing as $\{ c_i \}$, $\overline{\psi}(\mathbf{p})$ and $c(\alpha)$: the value $\psi(\mathbf{r}_0)$ which the wave function takes on at a point $\mathbf{r}_0$ of space can be considered as its component with respect to a specific function $\xi_{\mathbf{r}_0}(\mathbf{r})$ of a particular basis (the δ function basis).

Basis	*Components of* $\psi(\mathbf{r})$
$u_i(\mathbf{r})$	$c_i, i = 1, 2, ..., n, ...$
$v_{\mathbf{p}}(\mathbf{r})$	$\overline{\psi}(\mathbf{p})$
$\xi_{\mathbf{r}_0}(\mathbf{r})$	$\psi(\mathbf{r}_0)$
$w_\alpha(\mathbf{r})$	$c(\alpha)$

Table (II-5)

We thus find ourselves in a situation which is analogous to the one encountered in ordinary space, R^3: the position of a point in space can be described by a set of three numbers, which are its coordinates with respect to a system of axes defined in advance. If one changes axes, another set of coordinates corresponds to the same point. But the geometrical vector concept and vector calculation enable us to avoid referring to a system of axes; this considerably simplifies both formulas and reasoning.

We are going to use a similar approach here: each quantum state of a particle will be characterized by a *state vector*, belonging to an abstract space, $\mathscr{E}_{\mathbf{r}}$, called the *state space* of a particle. The fact that the space $\mathscr{F}$ is a subspace of L^2 means that $\mathscr{E}_{\mathbf{r}}$ is a subspace of a Hilbert space. We are going to define the notation and the rules of vector calculation in $\mathscr{E}_{\mathbf{r}}$.

Actually, the introduction of state vectors and the state space does more than merely simplify the formalism. It also permits a generalization of the formalism. Indeed, there exist physical systems whose quantum description cannot be given by a wave function: we shall see in Chapters IV and IX that this is the case when the spin degrees of freedom are taken into account, even for a single particle. Consequently, the first postulate that we shall set forth in Chapter III will be the following: *the quantum state of any physical system is characterized by a state vector, belonging to a space* $\mathscr{E}$ *which is the state space of the system.*

Therefore, in the rest of this chapter, we are going to develop a vector calculus in $\mathscr{E}$. The concepts which we are going to introduce and the results which we shall obtain are valid for whatever physical system we might consider. Nevertheless, to illustrate these concepts and results, we shall apply them to the simple case of a (spinless) particle, since this is the case we have previously considered.

We shall begin, in this paragraph, by defining the *Dirac notation*, which will prove to be very useful in the formal manipulations which we shall have to perform.

B-2. "Ket" vectors and "bra" vectors

B-2-a. Elements of $\mathscr{E}$: kets

α. Notation

Any element, or vector, of space $\mathscr{E}$ is called a *ket vector*, or, more simply, a *ket*. It is represented by the symbol $|\ \rangle$, inside which is placed a distinctive sign which enables us to distinguish the corresponding ket from all others, for example: $|\,\psi\,\rangle$.

In particular, since the concept of a wave function is now familiar to us, we shall define the space $\mathscr{E}_{\mathbf{r}}$ of the states of a particle by associating with every square-integrable

function $\psi(\mathbf{r})$ a ket vector $|\,\psi\,\rangle$ of $\mathscr{E}_{\mathbf{r}}$:

$$\psi(\mathbf{r}) \in \mathscr{F} \iff |\,\psi\,\rangle \in \mathscr{E}_{\mathbf{r}} \tag{B-1}$$

Afterwards, we shall transpose into $\mathscr{E}_{\mathbf{r}}$ the different operations that we introduced for $\mathscr{F}$. Although $\mathscr{F}$ and $\mathscr{E}_{\mathbf{r}}$ are isomorphic, we shall carefully distinguish between them in order to avoid confusion and to reserve the possibilities of generalization mentioned above in § B-1. We stress the fact that an $\mathbf{r}$-dependence no longer appears in $|\,\psi\,\rangle$; only the letter ψ appears, to remind us with which function it is associated. $\psi(\mathbf{r})$ will be interpreted (§ E) as the set of the components of the ket $|\,\psi\,\rangle$ in a particular basis, $\mathbf{r}$ playing the role of an index [*cf.* § A-3-b and table (II-5)]. Consequently, the procedure which we are adopting here consists in initially characterizing a vector by its components in a privileged coordinate system, which will later be treated on the same footing as all other coordinate systems.

We shall designate by $\mathscr{E}_x$ the state space of a (spinless) particle in only one dimension, that is, the abstract space constructed as in (B-1), but using wave functions that depend only on the x variable.

β. Scalar product

With each pair of kets $|\,\varphi\,\rangle$ and $|\,\psi\,\rangle$, taken in this order, we associate a complex number, which is their scalar product, $(|\,\varphi\,\rangle, |\,\psi\,\rangle)$, and which satisfies the various properties described by equations (A-5), (A-6) and (A-7). We shall later rewrite these formulas in Dirac notation after we have introduced the concept of a "bra".

In $\mathscr{E}_{\mathbf{r}}$, the scalar product of two kets will coincide with the scalar product defined above for the associated wave functions.

B-2-b. Elements of the dual space $\mathscr{E}^*$ of $\mathscr{E}$: bras

α. Definition of the dual space $\mathscr{E}^*$

Recall, first of all, the definition of a *linear functional* defined on the kets $|\,\psi\,\rangle$ of $\mathscr{E}$. A linear functional χ is a linear operation which associates a complex number with every ket $|\,\psi\,\rangle$:

$$\begin{aligned} &|\,\psi\,\rangle \in \mathscr{E} \xrightarrow{\chi} \text{ number } \chi(|\,\psi\,\rangle) \\ &\chi(\lambda_1|\,\psi_1\,\rangle + \lambda_2|\,\psi_2\,\rangle) = \lambda_1\chi(|\,\psi_1\,\rangle) + \lambda_2\chi(|\,\psi_2\,\rangle) \end{aligned} \tag{B-2}$$

Linear functional and linear operator must not be confused. In both cases, one is dealing with linear operations, but the former associates each ket with a complex number, while the latter associates another ket.

It can be shown that the set of linear functionals defined on the kets $|\,\psi\,\rangle \in \mathscr{E}$ constitutes a vector space, which is called the *dual space* of $\mathscr{E}$ and which will be symbolized by $\mathscr{E}^*$.

β. Bra notation for the vectors of $\mathscr{E}^*$

Any element, or vector, of the space $\mathscr{E}^*$ is called a *bra vector*, or, more simply, a *bra*. It is symbolized by $\langle\ |$. For example, the bra $\langle\,\chi\,|$ designates the linear functional

χ and we shall henceforth use the notation $\langle\,\chi\,|\,\psi\,\rangle$ to denote the *number* obtained by causing the linear functional $\langle\,\chi\,|\in\mathscr{E}^*$ to act on the ket $|\,\psi\,\rangle\in\mathscr{E}$:

$$\boxed{\chi(|\,\psi\,\rangle)=\langle\,\chi\,|\,\psi\,\rangle} \tag{B-3}$$

The origin of this terminology is the word "bracket", used to denote the symbol $\langle\ |\ \rangle$. Hence the name "bra" for the left-hand side $\langle\ |$, and the name "ket" for the right-hand side $|\ \rangle$ of this symbol.

B-2-c. Correspondence between kets and bras

α. To every ket corresponds a bra

The existence of a scalar product in $\mathscr{E}$ will now enable us to show that we can associate, with every ket $|\,\varphi\,\rangle\in\mathscr{E}$, an element of $\mathscr{E}^*$, that is, a bra, which will be denoted by $\langle\,\varphi\,|$.

The ket $|\,\varphi\,\rangle$ does indeed enable us to define a linear functional: the one that associates (in a linear way), with each ket $|\,\psi\,\rangle\in\mathscr{E}$, a complex number equal to the scalar product $(|\,\varphi\,\rangle,|\,\psi\,\rangle)$ of $|\,\psi\,\rangle$ by $|\,\varphi\,\rangle$. Let $\langle\,\varphi\,|$ be this linear functional; it is thus defined by the relation:

$$\boxed{\langle\,\varphi\,|\,\psi\,\rangle=(|\,\varphi\,\rangle,|\,\psi\,\rangle)} \tag{B-4}$$

β. This correspondance is antilinear

In the space $\mathscr{E}$, the scalar product is antilinear with respect to the first vector. In the notation of (B-4), this is expressed by:

$$\begin{aligned}(\lambda_1|\,\varphi_1\,\rangle+\lambda_2|\,\varphi_2\,\rangle,|\,\psi\,\rangle)&=\lambda_1^*\,(|\,\varphi_1\,\rangle,|\,\psi\,\rangle)+\lambda_2^*\,(|\,\varphi_2\,\rangle,|\,\psi\,\rangle)\\&=\lambda_1^*\,\langle\,\varphi_1\,|\,\psi\,\rangle+\lambda_2^*\,\langle\,\varphi_2\,|\,\psi\,\rangle\\&=(\lambda_1^*\,\langle\,\varphi_1\,|+\lambda_2^*\,\langle\,\varphi_2\,|)\,|\,\psi\,\rangle\end{aligned} \tag{B-5}$$

It appears from (B-5) that the bra associated with the ket $\lambda_1|\,\varphi_1\,\rangle+\lambda_2|\,\varphi_2\,\rangle$ is the bra $\lambda_1^*\,\langle\,\varphi_1\,|+\lambda_2^*\,\langle\,\varphi_2\,|$:

$$\boxed{\lambda_1|\,\varphi_1\,\rangle+\lambda_2|\,\varphi_2\,\rangle\Longrightarrow\lambda_1^*\,\langle\,\varphi_1\,|+\lambda_2^*\,\langle\,\varphi_2\,|} \tag{B-6}$$

The ket $\Longrightarrow$ bra correspondence is therefore antilinear.

Comment:

If λ is a complex number, and $|\,\psi\,\rangle$ a ket, $\lambda\,|\,\psi\,\rangle$ is a ket ($\mathscr{E}$ is a vector space). We are sometimes led to write it as $|\,\lambda\,\psi\,\rangle$:

$$|\,\lambda\psi\,\rangle=\lambda\,|\,\psi\,\rangle \tag{B-7}$$

One must then be careful to remember that $\langle\,\lambda\,\psi\,|$ *represents the bra associated with the ket* $|\,\lambda\,\psi\,\rangle$. Since the correspondence between a bra and a ket is antilinear, we have:

$$\langle\,\lambda\,\psi\,|=\lambda^*\langle\,\psi\,| \tag{B-8}$$

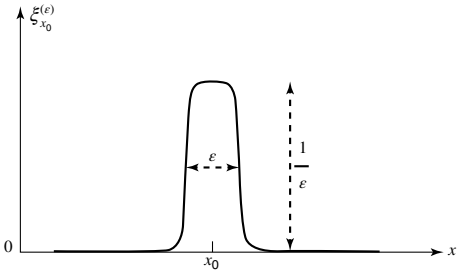


Figure 1: $\xi^{(\varepsilon)}_{x_0}(x)$ is a function having a peak at $x = x_0$ (of width ε and amplitude $1/\varepsilon$), whose integral between $-\infty$ and $+\infty$ is equal to 1.

γ. *Dirac notation for the scalar product*

We now have at our disposal two distinct notations for designating the scalar product of $|\,\psi\,\rangle$ by $|\,\varphi\,\rangle$: $(|\,\varphi\,\rangle, |\,\psi\,\rangle)$ or $\langle\,\varphi\,|\,\psi\,\rangle$, $\langle\,\varphi\,|$ being the bra associated with the ket $|\,\varphi\,\rangle$. Henceforth we shall use only the (Dirac) notation: $\langle\,\varphi\,|\,\psi\,\rangle$. Table (II-6) summarizes, in Dirac notation, the properties of the scalar product, already given in § A-1-b.

$$\langle\,\varphi\,|\,\psi\,\rangle = \langle\,\psi\,|\,\varphi\,\rangle^* \tag{B-9}$$

$$\langle\,\varphi\,|\,\lambda_1\psi_1 + \lambda_2\psi_2\,\rangle = \lambda_1\,\langle\,\varphi\,|\,\psi_1\,\rangle + \lambda_2\,\langle\,\varphi\,|\,\psi_2\,\rangle \tag{B-10}$$

$$\langle\,\lambda_1\varphi_1 + \lambda_2\varphi_2\,|\,\psi\,\rangle = \lambda_1^*\,\langle\,\varphi_1\,|\,\psi\,\rangle + \lambda_2^*\,\langle\,\varphi_2\,|\,\psi\,\rangle \tag{B-11}$$

$$\langle\,\psi\,|\,\psi\,\rangle \text{ real, positive; zero if and only if } |\,\psi\,\rangle = 0 \tag{B-12}$$

Table (II-6)

δ. *Is there a ket to correspond to every bra?*

Although to every ket there corresponds a bra, we shall see, in two examples chosen in $\mathscr{F}$, that it is possible to find bras that have no corresponding kets. We shall later show why this difficulty does not hinder us in quantum mechanics.

(i) Counter-examples chosen in $\mathscr{F}$

For simplicity, we shall reason in one dimension.

Let $\xi^{(\varepsilon)}_{x_0}(x)$ be a sufficiently regular real function, such that $\displaystyle\int_{-\infty}^{+\infty} \mathrm{d}x\ \xi^{(\varepsilon)}_{x_0}(x) = 1$, and having the form of a peak of width ε and amplitude $1/\varepsilon$, centered at $x = x_0$ [see Fig. 1; $\xi^{(\varepsilon)}_{x_0}(x)$ is, for example, one of the functions considered in § 1-b of Appendix II]. If $\varepsilon \neq 0$, $\xi^{(\varepsilon)}_{x_0}(x) \in \mathscr{F}_x$ (the square of its norm is of the order of $1/\varepsilon$). Denote by $|\,\xi^{(\varepsilon)}_{x_0}\,\rangle$ the corresponding ket:

$$\xi^{(\varepsilon)}_{x_0}(x) \Longleftrightarrow |\,\xi^{(\varepsilon)}_{x_0}\,\rangle \tag{B-13}$$

If $\varepsilon \neq 0$, $|\,\xi_{x_0}^{(\varepsilon)}\,\rangle \in \mathscr{E}_x$. Let $\langle\,\xi_{x_0}^{(\varepsilon)}\,|$ be the bra associated with this ket; for every $|\,\psi\,\rangle \in \mathscr{E}_x$, we have:

$$\langle\,\xi_{x_0}^{(\varepsilon)}\,|\,\psi\,\rangle = (\xi_{x_0}^{(\varepsilon)}, \psi) = \int_{-\infty}^{+\infty} \mathrm{d}x\ \xi_{x_0}^{(\varepsilon)}(x)\ \psi(x) \tag{B-14}$$

Now let ε approach zero. On the one hand:

$$\lim_{\varepsilon \to 0} \xi_{x_0}^{(\varepsilon)}(x) = \xi_{x_0}(x) \notin \mathscr{F}_x \tag{B-15}$$

[the square of the norm of $\xi_{x_0}^{(\varepsilon)}(x)$, which is of the order of $1/\varepsilon$, diverges when $\varepsilon \longrightarrow 0$]; therefore:

$$\lim_{\varepsilon \to 0} |\,\xi_{x_0}^{(\varepsilon)}\,\rangle \notin \mathscr{E}_x \tag{B-16}$$

On the other hand, when $\varepsilon \longrightarrow 0$, integral (B-14) approaches a perfectly well-defined limit, $\psi(x_0)$ [since, for sufficiently small ε, $\psi(x)$ can be replaced in (B-14) by $\psi(x_0)$ and removed from the integral]. Consequently, $\langle\,\xi_{x_0}^{(\varepsilon)}\,|$ approaches a bra which we shall denote by $\langle\,\xi_{x_0}\,|$: $\langle\,\xi_{x_0}\,|$ is the linear functional which associates, with every ket $|\,\psi\,\rangle$ of $\mathscr{E}_x$, the value $\psi(x_0)$ taken on by the associated wave function at the point x_0:

$$\begin{aligned} &\lim_{\varepsilon \to 0} \langle\,\xi_{x_0}^{(\varepsilon)}\,| = \langle\,\xi_{x_0}\,| \in \mathscr{E}_x^* \\ &\text{If } |\,\psi\,\rangle \in \mathscr{E}_x\ ,\ \langle\,\xi_{x_0}\,|\,\psi\,\rangle = \psi(x_0) \end{aligned} \tag{B-17}$$

Thus we see that *the bra* $\langle\,\xi_{x_0}\,|$ *exists, but no ket corresponds to it.*

In the same way, let us consider a plane wave which is truncated outside an interval of width L:

$$v_{p_0}^{(L)}(x) = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{ip_0 x/\hbar} \ \text{ if } \ -\frac{L}{2} \leqslant x \leqslant +\frac{L}{2} \tag{B-18}$$

with the function $v_{p_0}^{(L)}(x)$ going rapidly to zero outside this interval (while remaining continuous and differentiable). We shall denote by $|\,v_{p_0}^{(L)}\,\rangle$ the ket associated with $v_{p_0}^{(L)}(x)$:

$$v_{p_0}^{(L)}(x) \in \mathscr{F}_x \Longleftrightarrow |\,v_{p_0}^{(L)}\,\rangle \in \mathscr{E}_x \tag{B-19}$$

The square of the norm of $|\,v_{p_0}^{(L)}\,\rangle$, which is practically equal to $L/2\pi\hbar$, diverges if $L \longrightarrow \infty$. Therefore:

$$\lim_{L \to \infty} |\,v_{p_0}^{(L)}\,\rangle \notin \mathscr{E}_x \tag{B-20}$$

Now let us consider the bra $\langle\,v_{p_0}^{(L)}\,|$ associated with $|\,v_{p_0}^{(L)}\,\rangle$. For every $|\,\psi\,\rangle \in \mathscr{E}_x$, we have:

$$\langle\,v_{p_0}^{(L)}\,|\,\psi\,\rangle = (v_{p_0}^{(L)}, \psi) \simeq \frac{1}{\sqrt{2\pi\hbar}} \int_{-L/2}^{+L/2} \mathrm{d}x\ \mathrm{e}^{-ip_0 x/\hbar}\ \psi(x) \tag{B-21}$$

When $L \longrightarrow \infty$, $\langle\,v_{p_0}^{(L)}\,|\,\psi\,\rangle$ has a limit: the value $\overline{\psi}(p_0)$ of the Fourier transform $\overline{\psi}(p)$ of $\psi(x)$ for $p = p_0$. Therefore, when $L \longrightarrow \infty$, $\langle\,v_{p_0}^{(L)}\,|$ tends towards a perfectly well-defined bra $\langle\,v_{p_0}\,|$:

$$\begin{aligned} &\lim_{L \to \infty} \langle\,v_{p_0}^{(L)}\,| = \langle\,v_{p_0}\,| \in \mathscr{E}_x^* \\ &\text{If } |\,\psi\,\rangle \in \mathscr{E}_x\ ,\ \langle\,v_{p_0}\,|\,\psi\,\rangle = \overline{\psi}(p_0) \end{aligned} \tag{B-22}$$

Here again, *no ket corresponds to the bra* $\langle v_{p_0} |$.

(*ii*) Physical resolution of the preceding difficulties

This dissymmetry of the correspondence between kets and bras is related, as the preceding examples show, to the existence of "continuous bases" for $\mathscr{F}_x$. Since the functions constituting these "bases" do not belong to $\mathscr{F}_x$, we cannot associate a ket of $\mathscr{E}_x$ with them. However, their scalar product with an arbitrary function of $\mathscr{F}_x$ is defined, and this permits us to associate with them a linear functional in $\mathscr{E}_x$, that is, a bra belonging to $\mathscr{E}_x^*$. The reason for using such "continuous bases" lies in their usefulness in certain practical calculations. The same reason (which will become more apparent in what follows) leads us here to reestablish the symmetry between kets and bras by introducing "generalized kets", defined using functions that are not square-integrable, but whose scalar product with every function of $\mathscr{F}_x$ exists. In what follows, we shall work with "kets" such as $|\xi_{x_0}\rangle$ or $|v_{p_0}\rangle$, associated with $\xi_{x_0}(x)$ or $v_{p_0}(x)$. It must not be forgotten that these generalized "kets" cannot, strictly speaking, represent physical states. They are merely intermediaries, useful in calculations involving certain operations to be performed on the true kets of the space $\mathscr{E}_x$, which actually characterize realizable quantum states.

This method poses a certain number of mathematical problems, which can be avoided by adopting the following physical point of view: $|\xi_{x_0}\rangle$ (or $|v_{p_0}\rangle$) actually denotes $|\xi_{x_0}^{(\varepsilon)}\rangle$ (or $|v_{p_0}^{(L)}\rangle$) where ε is very small (or L is very large) compared to all the other lengths in the problem we are considering. In all the intermediary calculations where $|\xi_{x_0}^{(\varepsilon)}\rangle$ (or $|v_{p_0}^{(L)}\rangle$) appears, the limit $\varepsilon = 0$ (or $L \longrightarrow \infty$) is never attained, so that one is always working in $\mathscr{E}_x$. The physical result obtained at the end of the calculation depends very little on the value of ε, as long as ε is sufficiently small with respect to all the other lengths: it is then possible to neglect ε, that is, to set $\varepsilon = 0$, in the final result (the procedure to be used for L is analogous).

The objection could be raised that, unlike $\{\xi_{x_0}(x)\}$ and $\{v_{p_0}(x)\}$, $\{\xi_{x_0}^{(\varepsilon)}(x)\}$ and $\{v_{p_0}^{(L)}(x)\}$ are not orthonormal bases, insofar as they do not rigorously satisfy the closure relation. In fact, they fulfill it approximately. For example, the expression $\int \mathrm{d}x_0\, \xi_{x_0}^{(\varepsilon)}(x)\, \xi_{x_0}^{(\varepsilon)}(x')$ is a function of $(x-x')$ which can serve as an excellent approximation for $\delta(x-x')$. Its graphical representation is practically a triangle of base 2ε and height $\dfrac{1}{\varepsilon}$, centered at $x-x'=0$ (Appendix II, § 1-c-*iv*). If ε is negligible compared to all the other lengths in the problem, the difference between this expression and $\delta(x-x')$ is physically inappreciable.

In general, the dual space $\mathscr{E}^*$ and the state space $\mathscr{E}$ are not isomorphic, except, of course, if $\mathscr{E}$ is finite-dimensional[6]: although to each ket $|\psi\rangle$ of $\mathscr{E}$ there corresponds a bra $\langle\psi|$ in $\mathscr{E}^*$, the converse is not true. Nevertheless, we shall agree to use, in addition to vectors belonging to $\mathscr{E}$ (whose norm is finite), *generalized kets* with infinite norms but whose scalar product with every ket of $\mathscr{E}$ is finite. Thus, to each bra $\langle\varphi|$ of $\mathscr{E}^*$, there will correspond a ket. But generalized kets do not represent physical states of the system.

B-3. Linear operators

B-3-a. Definitions

They are the same as those of § A-1-c.

A linear operator A associates with every ket $|\psi\rangle \in \mathscr{E}$ another ket $|\psi\rangle' \in \mathscr{E}$, the

[6] It is true that the Hilbert space L^2 and its dual space are isomorphic; however, we have taken for the wave function space $\mathscr{F}$ a subspace of L^2, which explains why $\mathscr{F}^*$ is "larger" than $\mathscr{F}$.

correspondence being linear:

$$|\,\psi'\,\rangle = A\,|\,\psi\,\rangle \tag{B-23}$$

$$A(\lambda_1\,|\,\psi_1\,\rangle + \lambda_2\,|\,\psi_2\,\rangle) = \lambda_1\,A\,|\,\psi_1\,\rangle + \lambda_2\,A\,|\,\psi_2\,\rangle \tag{B-24}$$

The product of two linear operators A and B, written AB, is defined in the following way:

$$(AB)\,|\,\psi\,\rangle = A(B\,|\,\psi\,\rangle) \tag{B-25}$$

B first acts on $|\,\psi\,\rangle$ to give the ket $B\,|\,\psi\,\rangle$; A then acts on the ket $B\,|\,\psi\,\rangle$. In general, $AB \neq BA$. The commutator $[A, B]$ of A and B is, by definition:

$$[A, B] = AB - BA \tag{B-26}$$

Let $|\,\varphi\,\rangle$ and $|\,\psi\,\rangle$ be two kets. We call the *matrix element* of A between $|\,\varphi\,\rangle$ and $|\,\psi\,\rangle$, the scalar product:

$$\langle\,\varphi\,|\,(A\,|\,\psi\,\rangle) \tag{B-27}$$

Consequently, this is *a number* which depends linearly on $|\,\psi\,\rangle$ and antilinearly on $|\,\varphi\,\rangle$.

B-3-b. Examples of linear operators: projectors

α. *Important comment about Dirac notation*

We have begun to sense, in the preceding, the simplicity and convenience of the Dirac formalism. For example, $\langle\,\varphi\,|$ denotes a linear functional (a bra), and $\langle\,\psi_1\,|\,\psi_2\,\rangle$, the scalar product of two kets $|\,\psi_1\,\rangle$ and $|\,\psi_2\,\rangle$. The number associated by the linear functional $\langle\,\varphi\,|$ with an arbitrary ket $|\,\psi\,\rangle$ is then written simply by juxtaposing the symbols $\langle\,\varphi\,|$ and $|\,\psi\,\rangle$: $\langle\,\varphi\,|\,\psi\,\rangle$. This is the scalar product of $|\,\psi\,\rangle$ by the ket $|\,\varphi\,\rangle$ corresponding to $\langle\,\varphi\,|$ (which is why it is useful to have a one-to-one correspondence between kets and bras).

Now assume that we write $\langle\,\varphi\,|$ and $|\,\psi\,\rangle$ in the opposite order:

$$|\,\psi\,\rangle\,\langle\,\varphi\,| \tag{B-28}$$

We shall see that if we abide by the rule of juxtaposition of symbols, this expression represents an operator. Choose an arbitrary ket $|\,\chi\,\rangle$ and consider:

$$|\,\psi\,\rangle\,\langle\,\varphi\,|\,\chi\,\rangle \tag{B-29}$$

We already know that $\langle\,\varphi\,|\,\chi\,\rangle$ is a complex number; consequently,(B-29) is a ket, obtained by multiplying $|\,\psi\,\rangle$ by the scalar $\langle\,\varphi\,|\,\chi\,\rangle$. $|\,\psi\,\rangle\,\langle\,\varphi\,|$, applied to an arbitrary ket, gives another ket: it is an operator.

Thus we see that *the order of the symbols is of critical importance.* Only complex numbers can be moved about with impunity, because of the linearity of the space $\mathscr{E}$ and of the operators which we shall use. Indeed, if λ is a number:

$$\begin{cases} |\,\psi\,\rangle\,\lambda = \lambda\,|\,\psi\,\rangle \\ \langle\,\psi\,|\,\lambda = \lambda\,\langle\,\psi\,| \\ A\lambda\,|\,\psi\,\rangle = \lambda A\,|\,\psi\,\rangle \qquad \text{(where } A \text{ is a linear operator)} \\ \langle\,\varphi\,|\,\lambda\,|\,\psi\,\rangle = \lambda\,\langle\,\varphi\,|\,\psi\,\rangle = \langle\,\varphi\,|\,\psi\,\rangle\,\lambda \end{cases}$$

(B-30)

But, for kets, bras and operators, the order must always be carefully respected in writing the formulas: this is the price that must be paid for the simplicity of the Dirac formalism.

β. The projector P_ψ onto a ket $|\psi\rangle$

Let $|\psi\rangle$ be a ket which is normalized to one:

$$\langle\psi|\psi\rangle = 1 \tag{B-31}$$

Consider the operator P_ψ, defined by:

$$P_\psi = |\psi\rangle\langle\psi| \tag{B-32}$$

and apply it to an arbitrary ket $|\varphi\rangle$:

$$P_\psi|\varphi\rangle = |\psi\rangle\langle\psi|\varphi\rangle \tag{B-33}$$

P_ψ, acting on an arbitrary ket $|\varphi\rangle$, gives a ket proportional to $|\psi\rangle$. The coefficient of proportionality $\langle\psi|\varphi\rangle$ is the scalar product of $|\varphi\rangle$ by $|\psi\rangle$.

The "geometrical" significance of P_ψ is therefore clear: it is the "orthogonal projection" operator onto the ket $|\psi\rangle$.

This interpretation is confirmed by the fact that $P_\psi^2 = P_\psi$ (projecting twice in succession onto a given vector is equivalent to projecting a single time). To see this, we write:

$$P_\psi^2 = P_\psi P_\psi = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| \tag{B-34}$$

In this expression, $\langle\psi|\psi\rangle$ is a number, which is equal to 1 [formula (B-31)]. Therefore:

$$P_\psi^2 = |\psi\rangle\langle\psi| = P_\psi \tag{B-35}$$

γ. Projector onto a subspace

Let $|\varphi_1\rangle$, $|\varphi_2\rangle$, ..., $|\varphi_q\rangle$, be q normalized vectors which are orthogonal to each other:

$$\langle\varphi_i|\varphi_j\rangle = \delta_{ij} \quad ; \quad i,j = 1,\,2,\,...,\,q \tag{B-36}$$

We denote by $\mathscr{E}_q$ the subspace of $\mathscr{E}$ spanned by these q vectors.

Let P_q be the linear operator defined by:

$$P_q = \sum_{i=1}^{q} |\varphi_i\rangle\langle\varphi_i| \tag{B-37}$$

Calculating P_q^2:

$$P_q^2 = \sum_{i=1}^{q}\sum_{j=1}^{q} |\varphi_i\rangle\langle\varphi_i|\varphi_j\rangle\langle\varphi_j| \tag{B-38}$$

we get, using (B-36):

$$P_q^2 = \sum_{i=1}^{q}\sum_{j=1}^{q} |\,\varphi_i\,\rangle\,\langle\,\varphi_j\,|\,\delta_{ij} = \sum_{i=1}^{q} |\,\varphi_i\,\rangle\,\langle\,\varphi_i\,| = P_q \tag{B-39}$$

P_q is therefore a projector. It is easy to see that P_q projects onto the subspace $\mathscr{E}_q$, since for any $|\,\psi\,\rangle \in \mathscr{E}$:

$$P_q\,|\,\psi\,\rangle = \sum_{i=1}^{q} |\,\varphi_i\,\rangle\,\langle\,\varphi_i\,|\,\psi\,\rangle \tag{B-40}$$

P_q acting on $|\,\psi\,\rangle$ gives the linear superposition of the projections of $|\,\psi\,\rangle$ onto the various $|\,\varphi_i\,\rangle$, that is, the projection of $|\,\psi\,\rangle$ onto the subspace $\mathscr{E}_q$.

B-4. Hermitian conjugation

B-4-a. Action of a linear operator on a bra

Until now, we have only defined the action of a linear operator A on kets. We are now going to see that it is also possible to define the action of A on bras.

Let $\langle\,\varphi\,|$ be a well-defined bra, and consider the set of all kets $|\,\psi\,\rangle$. With each of these kets can be associated the complex number $\langle\,\varphi\,|\,(A\,|\,\psi\,\rangle)$, already defined above as the matrix element of A between $|\,\varphi\,\rangle$ and $|\,\psi\,\rangle$. Since A is linear and the scalar product depends linearly on the ket, the number $\langle\,\varphi\,|\,(A\,|\,\psi\,\rangle)$ depends linearly on $|\,\psi\,\rangle$. Thus, for fixed $\langle\,\varphi\,|$ and A, we can associate with every ket $|\,\psi\,\rangle$ a number which depends linearly on $|\,\psi\,\rangle$. The specification of $\langle\,\varphi\,|$ and A therefore defines a new linear functional on the kets of $\mathscr{E}$, that is, a new bra belonging to $\mathscr{E}^*$. We shall denote this new bra by $\langle\,\varphi\,|\,A$. The relation which defines $\langle\,\varphi\,|\,A$ thus be written:

$$\boxed{(\langle\,\varphi\,|\,A)\,|\,\psi\,\rangle = \langle\,\varphi\,|\,(A\,|\,\psi\,\rangle)} \tag{B-41}$$

The operator A associates with every bra $\langle\,\varphi\,|$ a new bra $\langle\,\varphi\,|\,A$. Let us show that the correspondence is linear. In order to do this, consider a linear combination of bras $\langle\,\varphi_1\,|$ and $\langle\,\varphi_2\,|$:

$$\langle\,\chi\,| = \lambda_1\,\langle\,\varphi_1\,| + \lambda_2\,\langle\,\varphi_2\,| \tag{B-42}$$

(which means that $\langle\,\chi\,|\,\psi\,\rangle = \lambda_1\,\langle\,\varphi_1\,|\,\psi\,\rangle + \lambda_2\,\langle\,\varphi_2\,|\,\psi\,\rangle$). From (B-41), we have:

$$\begin{aligned}(\langle\,\chi\,|\,A)\,|\,\psi\,\rangle &= \langle\,\chi\,|\,(A\,|\,\psi\,\rangle)\\ &= \lambda_1\,\langle\,\varphi_1\,|\,(A\,|\,\psi\,\rangle) + \lambda_2\,\langle\,\varphi_2\,|\,(A\,|\,\psi\,\rangle)\\ &= \lambda_1\,(\langle\,\varphi_1\,|\,A)\,|\,\psi\,\rangle + \lambda_2\,(\langle\,\varphi_2\,|\,A)\,|\,\psi\,\rangle)\end{aligned} \tag{B-43}$$

Since $|\,\psi\,\rangle$ is arbitrary, it follows that:

$$\begin{aligned}\langle\,\chi\,|\,A &= (\lambda_1\,\langle\,\varphi_1\,| + \lambda_2\,\langle\,\varphi_2\,|)\,A\\ &= \lambda_1\,\langle\,\varphi_1\,|\,A + \lambda_2\,\langle\,\varphi_2\,|\,A\end{aligned} \tag{B-44}$$

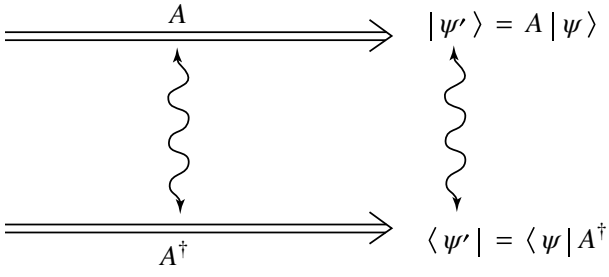


Figure 2: Definition of the adjoint operator $A^\dagger$ of an operator A using the correspondence between kets and bras.

Equation (B-41) therefore defines a linear operation on bras. The bra $\langle\varphi|A$ is the bra which results from the action of the linear operator A on the bra $\langle\varphi|$.

Commments:

(i) From definition (B-41) of $\langle\varphi|A$, we see that the place of the parenthesis in the symbol defining the matrix element of A between $|\varphi\rangle$ and $|\psi\rangle$ is of no importance. Therefore, we shall henceforth designate this matrix element by the notation $\langle\varphi|A|\psi\rangle$:

$$\langle\varphi|A|\psi\rangle = (\langle\varphi|A)|\psi\rangle = \langle\varphi|(A|\psi\rangle) \tag{B-45}$$

(*ii*) The relative order of $\langle\varphi|$ and A is very important in the notation $\langle\varphi|A$ (*cf.* § 3-b-a above). One must write $\langle\varphi|A$ and not $A\langle\varphi|$: $\langle\varphi|A$ acting on a ket $|\psi\rangle$ gives a number $\langle\varphi|A|\psi\rangle$; $\langle\varphi|A$ is therefore indeed a bra. On the other hand, $A\langle\varphi|$, acting on a ket $|\psi\rangle$, would give $A\langle\varphi|\psi\rangle$, that is, an operator (the operator A multiplied by the number $\langle\varphi|\psi\rangle$). We have not defined any mathematical object of this sort: $A\langle\varphi|$ therefore has no meaning.

B-4-b. The adjoint operator $A^\dagger$ of a linear operator A

We are now going to see that the correspondence between kets and bras, studied in § B-2-c, enables us to associate with every linear operator A another linear operator $A^\dagger$, called the adjoint operator (or Hermitian conjugate) of A.

Let $|\psi\rangle$ then be an arbitrary ket of $\mathscr{E}$. The operator A associates with it another ket $|\psi'\rangle = A|\psi\rangle$ of $\mathscr{E}$ (Fig. 2).

To the ket $|\psi\rangle$ corresponds a bra $\langle\psi|$; in the same way, to $|\psi'\rangle$ corresponds $\langle\psi'|$. This correspondence between kets and bras thus permits us to define the action of the operator $A^\dagger$ on the bras: the operator $A^\dagger$ associates with the bra $\langle\psi|$ corresponding to the ket $\langle\psi|$, the bra $\langle\psi'|$ corresponding to the ket $|\psi'\rangle = A|\psi\rangle$. We write: $\langle\psi'| = \langle\psi|A^\dagger$.

Let us show that the relation $\langle\psi'| = \langle\psi|A^\dagger$ is linear. We know that, to the bra $\lambda_1\langle\psi_1| + \lambda_2\langle\psi_2|$, corresponds the ket $\lambda_1^*|\psi_1\rangle + \lambda_2^*|\psi_2\rangle$ (the correspondence between a

bra and a ket is antilinear). The operator A transforms $\lambda_1^* \,|\, \psi_1 \,\rangle + \lambda_2^* \,|\, \psi_2 \,\rangle$ into $\lambda_1^* \, A \,|\, \psi_1 \,\rangle + \lambda_2^* \, A \,|\, \psi_2 \,\rangle = \lambda_1^* \,|\, \psi_1' \,\rangle + \lambda_2^* \,|\, \psi_2' \,\rangle$. Finally, to this ket corresponds the bra: $\lambda_1 \,\langle\, \psi_1' \,| + \lambda_2 \,\langle\, \psi_2' \,| = \lambda_1 \,\langle\, \psi_1 \,| \, A^\dagger + \lambda_2 \,\langle\, \psi_2 \,| \, A^\dagger$. From this we conclude that:

$$(\lambda_1 \,\langle\, \psi_1 \,| + \lambda_2 \,\langle\, \psi_2 \,|) \, A^\dagger = \lambda_1 \,\langle\, \psi_1 \,| \, A^\dagger + \lambda_2 \,\langle\, \psi_2 \,| \, A^\dagger \tag{B-46}$$

$A^\dagger$ is therefore a linear operator, defined by the formula:

$$\boxed{|\, \psi' \,\rangle = A \,|\, \psi \,\rangle \Longleftrightarrow \langle\, \psi' \,| = \langle\, \psi \,| \, A^\dagger} \tag{B-47}$$

From (B-47), it is easy to deduce another important relationship satisfied by the operator $A^\dagger$. Using the properties of the scalar product, one can always write:

$$\langle\, \psi' \,|\, \varphi \,\rangle = \langle\, \varphi \,|\, \psi' \,\rangle^* \tag{B-48}$$

where $|\, \varphi \,\rangle$ is an arbitrary ket of $\mathscr{E}$. Using expressions (B-47) for $|\, \psi' \,\rangle$ and $\langle\, \psi' \,|$, we obtain:

$$\boxed{\langle\, \psi \,| \, A^\dagger \,|\, \varphi \,\rangle = \langle\, \varphi \,| \, A \,|\, \psi \,\rangle^*} \tag{B-49}$$

a relation which is valid for all $|\, \varphi \,\rangle$ and $|\, \psi \,\rangle$.

Comment about notation:

We have already mentioned a notation which can lead to confusion: $|\, \lambda \, \psi \,\rangle$ and $\langle\, \lambda \, \psi \,|$, where λ is a scalar [formulas (B-7) and (B-8)]. The same problem arises with the expressions $|\, A \, \psi \,\rangle$ and $\langle\, A \, \psi \,|$, where A is a linear operator. $|\, A \, \psi \,\rangle$ is another way of designating the ket $A \,|\, \psi \,\rangle$:

$$|\, A \, \psi \,\rangle = A \,|\, \psi \,\rangle \tag{B-50}$$

$\langle\, A \, \psi \,|$ is the bra associated with the ket $|\, A \, \psi \,\rangle$. Using (B-50) and (B-57), we see that:

$$\langle\, A \, \psi \,| = \langle\, \psi \,| \, A^\dagger \tag{B-51}$$

When a linear operator A is taken outside the bra symbol, it must be replaced by its adjoint $A^\dagger$ (and placed to the right of the bra).

B-4-c. Correspondence between an operator and its adjoint

By using (B-47) or (B-49), it is easy to show that:

$$(A^\dagger)^\dagger = A \tag{B-52}$$

$$(\lambda A)^\dagger = \lambda^* \, A^\dagger \qquad \text{(where } \lambda \text{ is a number)} \tag{B-53}$$

$$(A + B)^\dagger = A^\dagger + B^\dagger \tag{B-54}$$

Now let us calculate $(AB)^\dagger$. To do this, consider the ket $|\,\varphi\,\rangle = AB\,|\,\psi\,\rangle$. Write it in the form $|\,\varphi\,\rangle = A\,|\,\chi\,\rangle$, setting $|\,\chi\,\rangle = B\,|\,\psi\,\rangle$. Then:

$$\langle\,\varphi\,| = \langle\,\psi\,|\,(AB)^\dagger = \langle\,\chi\,|\,A^\dagger = \langle\,\psi\,|\,B^\dagger A^\dagger$$

since $\langle\,\chi\,| = \langle\,\psi\,|\,B^\dagger$. From this, we deduce that:

$$\boxed{(AB)^\dagger = B^\dagger A^\dagger} \tag{B-55}$$

Note that *the order changes when one takes the adjoint of a product of operators.*

Comment:

Since $(A^\dagger)^\dagger = A$, we can write, using (B-51):

$$\langle\,A^\dagger\varphi\,| = \langle\,\varphi\,|\,(A^\dagger)^\dagger = \langle\,\varphi\,|\,A$$

Thus the left-hand side of (B-41) can be rewritten in the form $\langle\,A^\dagger\,\varphi\,|\,\psi\,\rangle$. In the same way, the right-hand side of this same equation can be put, with the notation of (B-50), into the form $\langle\,\varphi\,|\,A\,\psi\,\rangle$. From this results the following relation, sometimes used to define the adjoint operator $A^\dagger$ of A:

$$\langle\,A^\dagger\varphi\,|\,\psi\,\rangle = \langle\,\varphi\,|\,A\,\psi\,\rangle \tag{B-56}$$

B-4-d. Hermitian conjugation in Dirac notation

In the preceding section, we introduced the concept of an adjoint operator by using the correspondence between kets and bras. A ket $|\,\psi\,\rangle$ and its corresponding bra $\langle\,\psi\,|$ are said to be "Hermitian conjugates" of each other. The operation of Hermitian conjugation is represented by the wavy arrows in Figure 2; we see that it associates $A^\dagger$ with A. This is the reason why $A^\dagger$ is also called the Hermitian conjugate operator of A.

The operation of Hermitian conjugation changes the order of the objects to which it is applied. Thus we see in Figure 2 that $A\,|\,\psi\,\rangle$ becomes $\langle\,\psi\,|\,A^\dagger$. The ket $|\,\psi\,\rangle$ is changed into $\langle\,\psi\,|$, the operator A into $A^\dagger$, and the order is reversed. In the same way, we saw in (B-55) that the Hermitian conjugate of a product of two operators is equal to the product of the Hermitian conjugates taken in the opposite order. Finally, let us show that:

$$(|\,u\,\rangle\,\langle\,v\,|)^\dagger = |\,v\,\rangle\,\langle\,u\,| \tag{B-57}$$

($|\,u\,\rangle$ is replaced by $\langle\,u\,|$, and $\langle\,v\,|$ by $|\,v\,\rangle$, and the order is changed). Applying relation (B-49) to the operator $|\,u\,\rangle\langle\,v\,|$, we find:

$$\langle\,\psi\,|\,(|\,u\,\rangle\,\langle\,v\,|)^\dagger\,|\,\varphi\,\rangle = [\langle\,\varphi\,|\,(|\,u\,\rangle\,\langle\,v\,|)\,|\,\psi\,\rangle]^* \tag{B-58}$$

Now, if we use property (B-9) of the scalar product:

$$\begin{aligned}[\langle\,\varphi\,|\,(|\,u\,\rangle\,\langle\,v\,|)\,|\,\psi\,\rangle]^* &= \langle\,\varphi\,|\,u\,\rangle^*\,\langle\,v\,|\,\psi\,\rangle^* = \langle\,\psi\,|\,v\,\rangle\,\langle\,u\,|\,\varphi\,\rangle \\ &= \langle\,\psi\,|\,(|\,v\,\rangle\,\langle\,u\,|)\,|\,\varphi\,\rangle \end{aligned} \tag{B-59}$$

By comparing (B-58) and (B-59), we can derive (B-57).

The result of the operation of Hermitian conjugation on a constant remains to be found. We see from (B-6) and (B-53) that this operation simply transforms λ into λ^* (complex conjugation). This is in agreement with the fact that $\langle \varphi | \psi \rangle^* = \langle \psi | \varphi \rangle$.

To summarize, the Hermitian conjugate of a ket is a bra, and vice versa; that of an operator is its adjoint; that of a number, its complex conjugate. In Dirac notation, the operation of Hermitian conjugation is very simple to perform; it suffices to apply the following rule:

RULE

To obtain the Hermitian conjugate (or the adjoint) of any expression composed of constants, kets, bras and operators, one must:

– *Replace* { the constants by their complex conjugates; the kets by the bras associated with them; the bras by the kets associated with them; the operators by their adjoints

– *Reverse the order* of the factors (the position of the constants, nevertheless, is of no importance).

EXAMPLES

$\lambda \langle u | A | v \rangle | w \rangle \langle \psi |$ is an operator (λ and $\langle u | A | v \rangle$ are numbers). The adjoint of this operator is obtained by using the preceding rule: $| \psi \rangle \langle w | \langle v | A^\dagger | u \rangle \lambda^*$, which can also be written $\lambda^* \langle v | A^\dagger | u \rangle | \psi \rangle \langle w |$, changing the position of the numbers λ^* and $\langle v | A^\dagger | u \rangle$.

In the same way, $\lambda | u \rangle \langle v | w \rangle$ is a ket (λ and $\langle v | w \rangle$ are constants). The conjugate bra is $\langle w | v \rangle \langle u | \lambda^*$, which can also be written $\lambda^* \langle w | v \rangle \langle u |$.

B-4-e. Hermitian operators

An operator A is said to be Hermitian if it is equal to its adjoint, that is, if:

$$A = A^\dagger \tag{B-60}$$

Combining (B-60) and (B-49), we see that a Hermitian operator satisfies the relation:

$$\langle \psi | A | \varphi \rangle = \langle \varphi | A | \psi \rangle^* \tag{B-61}$$

which is valid for all $| \varphi \rangle$ and $| \psi \rangle$.

Finally, for a Hermitian operator, (B-56) becomes

$$\langle A \varphi | \psi \rangle = \langle \varphi | A \psi \rangle \tag{B-62}$$

We shall treat Hermitian operators in more detail later, when we consider the problem of eigenvalues and eigenvectors. Moreover, we shall see in Chapter III that Hermitian operators play a fundamental role in quantum mechanics.

If formula (B-57) is applied to the case where $|u\rangle = |v\rangle = |\psi\rangle$, we see that the projector $P_\psi = |\psi\rangle\langle\psi|$ is Hermitian:

$$P_\psi^\dagger = |\psi\rangle\langle\psi| = P_\psi \tag{B-63}$$

Comment:

The product of two Hermitian operators A and B is Hermitian only if $[A, B] = 0$. Indeed, if $A = A^\dagger$ and $B = B^\dagger$, it can be shown using (B-55) that $(AB)^\dagger = B^\dagger A^\dagger = BA$, which is equal to AB only if $[A, B] = 0$.

C. Representations in state space

C-1. Introduction

C-1-a. Definition of a representation

Choosing a representation means choosing an orthonormal basis, either discrete or continuous, in the state space $\mathscr{E}$. Vectors and operators are then represented in this basis by *numbers*: components for the vectors, matrix elements for the operators. The vectorial calculus introduced in § B then becomes a matrix calculus with these numbers. The choice of a representation is, in theory, arbitrary. Actually, it obviously depends on the particular problem being studied: in each case, one chooses the representation that leads to the simplest calculations.

C-1-b. Aim of section C

Using the Dirac notation, and for any arbitrary $\mathscr{E}$ space, we are going to treat again all the concepts introduced in §§ A-2 and A-3 for discrete and continuous bases of $\mathscr{F}$.

We shall write the two characteristic relations of a basis in Dirac notation: the orthonormalization and closure relations. Then we shall show how, using these two relations, it is possible to solve all specific problems involving a representation and the transformation from one representation to another.

C-2. Relations characteristic of an orthonormal basis

C-2-a. Orthonormalization relation

A set of kets, discrete $\{|u_i\rangle\}$ or continuous $\{|w_\alpha\rangle\}$, is said to be *orthonormal* if the kets of this set satisfy the orthonormalization relation:

$$\boxed{\langle u_i | u_j \rangle = \delta_{ij}} \tag{C-1}$$

or

$$\boxed{\langle w_\alpha | w_{\alpha'} \rangle = \delta(\alpha - \alpha')} \tag{C-2}$$

It can be seen that, for a continuous set, $\langle w_\alpha | w_\alpha \rangle$ does not exist: the $| w_\alpha \rangle$ have an infinite norm and therefore do not belong to $\mathscr{E}$. Nevertheless, the vectors of $\mathscr{E}$ can be expanded on the $| w_\alpha \rangle$. It is useful, consequently, to accept the $| w_\alpha \rangle$ as generalized kets (see the discussions in §§ A-3 and B-2-c).

C-2-b. Closure relation

A discrete set, $\{ | u_i \rangle \}$, or a continuous one, $\{ | w_\alpha \rangle \}$, constitutes a *basis* if every ket $| \psi \rangle$ belonging to $\mathscr{E}$ has a unique expansion on the $| u_i \rangle$ or the $| w_\alpha \rangle$:

$$| \psi \rangle = \sum_i c_i | u_i \rangle \tag{C-3}$$

$$| \psi \rangle = \int \mathrm{d}\alpha \; c(\alpha) | w_\alpha \rangle \tag{C-4}$$

Let us assume, moreover, that the basis is orthonormal. Then perform the scalar multiplication on both sides of (C-3) with $\langle u_j |$, and on both sides of (C-4) with $\langle w_{\alpha'} |$. We obtain, using (C-1) or (C-2), expressions for the components c_j or $c(\alpha')$:

$$\langle u_j | \psi \rangle \; = c_j \tag{C-5}$$

$$\langle w_{\alpha'} | \psi \rangle = c(\alpha') \tag{C-6}$$

Then replace in (C-3) c_i by $\langle u_i | \psi \rangle$, and in (C-4) $c(\alpha)$ by $\langle w_\alpha | \psi \rangle$:

$$\begin{aligned} | \psi \rangle &= \sum_i c_i | u_i \rangle = \sum_i \langle u_i | \psi \rangle | u_i \rangle \\ &= \sum_i | u_i \rangle \langle u_i | \psi \rangle = \left(\sum_i | u_i \rangle \langle u_i | \right) | \psi \rangle \end{aligned} \tag{C-7}$$

$$\begin{aligned} | \psi \rangle &= \int \mathrm{d}\alpha \; c(\alpha) | w_\alpha \rangle = \int \mathrm{d}\alpha \; \langle w_\alpha | \psi \rangle | w_\alpha \rangle \\ &= \int \mathrm{d}\alpha \; | w_\alpha \rangle \langle w_\alpha | \psi \rangle = \left(\int \mathrm{d}\alpha \; | w_\alpha \rangle \langle w_\alpha | \right) | \psi \rangle \end{aligned} \tag{C-8}$$

[since, in (C-7), we can place the number $\langle u_i | \psi \rangle$ after the ket $| u_i \rangle$; in the same way, in (C-8), we can place the number $\langle w_\alpha | \psi \rangle$ after the ket $| w_\alpha \rangle$].

Thus, we see two operators appear, $\sum_i | u_i \rangle \langle u_i |$ and $\int \mathrm{d}\alpha \; | w_\alpha \rangle \langle w_\alpha |$. They act on every ket $| \psi \rangle$ belonging to $\mathscr{E}$ to give the same ket $| \psi \rangle$. Since $| \psi \rangle$ is arbitrary, it follows that:

$$P_{\{u_i\}} = \sum_i | u_i \rangle \langle u_i | = \mathbb{1} \tag{C-9}$$

$$P_{\{w_\alpha\}} = \int \mathrm{d}\alpha \; | w_\alpha \rangle \langle w_\alpha | = \mathbb{1} \tag{C-10}$$

where $\mathbb{1}$ denotes the identity operator in $\mathscr{E}$. Relation (C-9), or (C-10), is called the *closure relation*. Conversely, let us show that relations (C-9) and (C-10) express the fact

that the sets $\{ |u_i\rangle \}$ and $\{ |w_\alpha\rangle \}$ constitute bases. For every $|\psi\rangle$ belonging to $\mathscr{E}$, we can write:

$$|\psi\rangle = \mathbb{1}\,|\psi\rangle = P_{\{u_i\}}\,|\psi\rangle = \sum_i |u_i\rangle\langle u_i|\psi\rangle$$
$$= \sum_i c_i\,|u_i\rangle \tag{C-11}$$

with:

$$c_i = \langle u_i|\psi\rangle \tag{C-12}$$

In the same way:

$$|\psi\rangle = \mathbb{1}\,|\psi\rangle = P_{\{w_\alpha\}}\,|\psi\rangle = \int \mathrm{d}\alpha\,|w_\alpha\rangle\langle w_\alpha|\psi\rangle$$
$$= \int \mathrm{d}\alpha\ c(\alpha)\,|w_\alpha\rangle \tag{C-13}$$

with:

$$c(\alpha) = \langle w_\alpha|\psi\rangle \tag{C-14}$$

Thus, every ket $|\psi\rangle$ has a unique expansion on the $|u_i\rangle$ or on the $|w_\alpha\rangle$. Each of these two sets therefore forms a basis, a discrete one or a continuous one. We also see that relation (C-9), or (C-10), spares us the need of memorizing expressions (C-12) and (C-14) for the components c_i and $c(\alpha)$.

Comments:

(*i*) We shall see later (§ E) that, in the case of the $\mathscr{F}$-space, relations (A-32) and (A-57) can easily be deduced from (C-9) and (C-10).

(*ii*) Geometrical interpretation of the closure relation.

From the discussion of § B-3-b, we see that $\sum_i |u_i\rangle\langle u_i|$ is a projector: the projector onto the subspace $\mathscr{E}'$ spanned by $|u_1\rangle, |u_2\rangle, ..., |u_i\rangle, ...$ If the $|u_i\rangle$ form a basis, every ket of $\mathscr{E}$ can be expanded on the $|u_i\rangle$; the subspace $\mathscr{E}'$ is then identical with the $\mathscr{E}$-space itself. Consequently, it is reasonable for $\sum_i |u_i\rangle\langle u_i|$ to be equal to the identity operator: projecting onto $\mathscr{E}$ a ket which belongs to $\mathscr{E}$ does not modify this ket. The same argument can be applied to $\int \mathrm{d}\alpha\ |w_\alpha\rangle\langle w_\alpha|$.

We can now find an equivalent of the closure relation for the three-dimensional space of ordinary geometry, R^3. If $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ are three orthonormal vectors of this space, and P_1, P_2 and P_3 are the projectors onto these three vectors, the fact that $\{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \}$ forms a basis in R^3 is expressed by the relation

$$P_1 + P_2 + P_3 = \mathbb{1} \tag{C-15}$$

On the other hand, $\{ \mathbf{e}_1, \mathbf{e}_2 \}$ constitutes an orthonormal set but not a basis of R^3. This is expressed by the fact that the projector $P_1 + P_2$ (which projects onto the plane spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$) is not equal to $\mathbb{1}$; for example: $(P_1 + P_2)\,\mathbf{e}_3 = 0$.

Table (II-7) summarizes the only fundamental formulas required for any calculation in the $\{\,|\,u_i\,\rangle\,\}$ or $\{\,|\,w_\alpha\,\rangle\,\}$ representation.

$\{\,\|\,u_i\,\rangle\,\}$ representation	$\{\,\|\,w_\alpha\,\rangle\,\}$ representation
$\langle\,u_i\,\|\,u_j\,\rangle = \delta_{ij}$ $P_{\{u_i\}} = \sum_i \|\,u_i\,\rangle\,\langle\,u_i\,\| = \mathbb{1}$	$\langle\,w_\alpha\,\|\,w_{\alpha'}\,\rangle = \delta(\alpha-\alpha')$ $P_{\{w_\alpha\}} = \int \mathrm{d}\alpha\ \|\,w_\alpha\,\rangle\,\langle\,w_\alpha\,\| = \mathbb{1}$

Table (II-7)

C-3. Representation of kets and bras

C-3-a. Representation of kets

In the $\{\,|\,u_i\,\rangle\,\}$ basis, the ket $|\,\psi\,\rangle$ is represented by the set of its components, that is, by the set of numbers $c_i = \langle\,u_i\,|\,\psi\,\rangle$. These numbers can be arranged vertically to form a one-column matrix (with, in general, a countable infinity of rows):

$$\begin{pmatrix} \langle\,u_1\,|\,\psi\,\rangle \\ \langle\,u_2\,|\,\psi\,\rangle \\ \vdots \\ \langle\,u_i\,|\,\psi\,\rangle \\ \vdots \end{pmatrix} \tag{C-16}$$

In the continuous $\{\,|\,w_\alpha\,\rangle\,\}$ basis, the ket $|\,\psi\,\rangle$ is represented by a continuous infinity of numbers, $c(\alpha) = \langle\,w_\alpha\,|\,\psi\,\rangle$, that is, by a function of α. It is then possible to draw a vertical axis, along which are placed the various possible values of α. To each of these values corresponds a number, $\langle\,w_\alpha\,|\,\psi\,\rangle$:

$$\alpha \downarrow \quad \begin{pmatrix} \vdots \\ \vdots \\ \langle w_\alpha\,|\psi\rangle \\ \vdots \\ \vdots \end{pmatrix} \tag{C-17}$$

C-3-b. Representations of bras

Let $\langle\,\varphi\,|$ be an arbitrary bra. In the $\{\,|\,u_i\,\rangle\,\}$ basis, we can write:

$$\langle\,\varphi\,| = \langle\,\varphi\,|\,\mathbb{1} = \langle\,\varphi\,|\,P_{\{u_i\}} = \sum_i \langle\,\varphi\,|\,u_i\,\rangle\,\langle\,u_i\,| \tag{C-18}$$

$\langle\,\varphi\,|$ has a unique expansion on the bras $\langle\,u_i\,|$. The components of $\langle\,\varphi\,|$, $\langle\,\varphi\,|\,u_i\,\rangle$, are the complex conjugates of the components $b_i = \langle\,u_i\,|\,\varphi\,\rangle$ of the ket $|\,\varphi\,\rangle$ associated with $\langle\,\varphi\,|$.

In the same way, we obtain, in the $\{\,|\,w_\alpha\,\rangle\,\}$ basis:

$$\langle\,\varphi\,| = \langle\,\varphi\,|\,\mathbb{1} = \langle\,\varphi\,|\,P_{\{w_\alpha\}} = \int \mathrm{d}\alpha\,\langle\,\varphi\,|\,w_\alpha\,\rangle\,\langle\,w_\alpha\,| \tag{C-19}$$

The components of $\langle\,\varphi\,|$, $\langle\,\varphi\,|\,w_\alpha\,\rangle$, are the complex conjugates of the components $b(\alpha) = \langle\,w_\alpha\,|\,\varphi\,\rangle$ of the ket $|\,\varphi\,\rangle$ associated with $\langle\,\varphi\,|$.

We have agreed to arrange the components of a ket vertically. Before describing how to arrange the components of a bra, let us show how the closure relation enables us to find simply the expression for the scalar product of two kets in terms of their components. We know that we can always place $\mathbb{1}$ between $\langle\,\varphi\,|$ and $|\,\psi\,\rangle$ in the expression for the scalar product:

$$\begin{aligned}\langle\,\varphi\,|\,\psi\,\rangle &= \langle\,\varphi\,|\,\mathbb{1}\,|\,\psi\,\rangle = \langle\,\varphi\,|\,P_{\{u_i\}}\,|\,\psi\,\rangle \\ &= \sum_i \langle\,\varphi\,|\,u_i\,\rangle\,\langle\,u_i\,|\,\psi\,\rangle = \sum_i b_i^* c_i\end{aligned} \tag{C-20}$$

In the same way:

$$\begin{aligned}\langle\,\varphi\,|\,\psi\,\rangle &= \langle\,\varphi\,|\,\mathbb{1}\,|\,\psi\,\rangle = \langle\,\varphi\,|\,P_{\{w_\alpha\}}\,|\,\psi\,\rangle \\ &= \int \mathrm{d}\alpha\,\langle\,\varphi\,|\,w_\alpha\,\rangle\,\langle\,w_\alpha\,|\,\psi\,\rangle = \int \mathrm{d}\alpha\;\, b^*(\alpha)\,c(\alpha)\end{aligned} \tag{C-21}$$

Let us arrange the components $\langle\,\varphi\,|\,u_i\,\rangle$ of the bra $\langle\,\varphi\,|$ horizontally, to form a row matrix (having one row and an infinite number of columns):

$$\big(\langle\,\varphi\,|\,u_1\,\rangle\;\langle\,\varphi\,|\,u_2\,\rangle\;\cdots\;\langle\,\varphi\,|\,u_i\,\rangle\;\cdots\big) \tag{C-22}$$

Using this convention, $\langle\,\varphi\,|\,\psi\,\rangle$ is the matrix product of the column matrix which represents $|\,\psi\,\rangle$ and the row matrix which represents $\langle\,\varphi\,|$. The result is a matrix having one row and one column, that is, a *number*:

In the $\{\,|\,w_\alpha\,\rangle\,\}$ basis, $\langle\,\varphi\,|$ has a continuous infinity of components $\langle\,\varphi\,|\,w_\alpha\,\rangle$. The various values of α are placed along a horizontal axis. To each of these values corresponds a component $\langle\,\varphi\,|\,w_\alpha\,\rangle$ of bra $\langle\,\varphi\,|$:

$$\begin{array}{c}\xrightarrow{\qquad\qquad\alpha\qquad\qquad} \\ \big(\,\ldots\;\ldots\;\langle\varphi\,|w_\alpha\rangle\;\ldots\;\ldots\,\big)\end{array} \tag{C-23}$$

Comment:

In a given representation, the matrices which represent a ket $|\,\psi\,\rangle$ and the associated bra $\langle\,\psi\,|$ are Hermitian conjugates of each other (in the matrix sense): one passes from one matrix to the other by interchanging rows and columns and taking the complex conjugate of each element.

C-4. Representation of operators

C-4-a. Representation of A by a "square" matrix

Given a linear operator A, we can, in a $\{\,|\,u_i\,\rangle\,\}$ or $\{\,|\,w_\alpha\,\rangle\,\}$ basis, associate with it a series of numbers defined by:

$$\boxed{A_{ij} = \langle\, u_i \,|\, A \,|\, u_j \,\rangle} \tag{C-24}$$

$$\text{or} \quad \boxed{A(\alpha, \alpha') = \langle\, w_\alpha \,|\, A \,|\, w_{\alpha'} \,\rangle} \tag{C-25}$$

These numbers depend on two indices and can therefore be arranged in a "square" matrix having a countable or continuous infinity of rows and columns. The usual convention is to have the first index fix the rows and the second, the columns. Thus, in the $\{\,|\,u_i\,\rangle\,\}$ basis, the operator A is represented by the matrix:

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1j} & \cdots \\ A_{21} & A_{22} & \cdots & A_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_{i1} & A_{i2} & \cdots & A_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix} \tag{C-26}$$

We see that the jth column is made up of the components, in the $\{\,|\,u_i\,\rangle\,\}$ basis, of the transform $A\,|\,u_j\,\rangle$ of the basis vector $|\,u_j\,\rangle$.

For a continuous basis, we draw two perpendicular axes. To a point that has α' for its abscissa and α for its ordinate, there corresponds the number $A(\alpha, \alpha')$:

$$\begin{array}{c} \xrightarrow{\qquad\qquad\alpha'\qquad\qquad} \\ \\ \alpha \downarrow \end{array} \quad \begin{pmatrix} & \vdots & \\ & \vdots & \\ \cdots & A(\alpha, \alpha') & \cdots \\ & & \end{pmatrix} \tag{C-27}$$

Let us use the closure relation to calculate the matrix which represents the operator AB in the $\{\,|\,u_i\,\rangle\,\}$ basis:

$$\begin{aligned} \langle\, u_i \,|\, AB \,|\, u_j \,\rangle &= \langle\, u_i \,|\, A\mathbb{1}B \,|\, u_j \,\rangle \\ &= \langle\, u_i \,|\, AP_{\{u_k\}}B \,|\, u_j \,\rangle \\ &= \sum_k \langle\, u_i \,|\, A \,|\, u_k \,\rangle \langle\, u_k \,|\, B \,|\, u_j \,\rangle \end{aligned} \tag{C-28}$$

The convention chosen above for the arrangement of the elements A_{ij} [or $A(\alpha, \alpha')$] is therefore consistent with the one relating to the product of two matrices: (C-28) expresses the fact that the matrix representing the operator AB is the product of the matrices associated with A and B.

C-4-b. Matrix representation of the ket $|\psi'\rangle = A|\psi\rangle$

The problem is the following: knowing the components of $|\psi\rangle$ and the matrix elements of A in a given representation, how can we calculate the components of $|\psi'\rangle = A|\psi\rangle$ in the same representation?

In the $\{|u_i\rangle\}$ basis, the coordinate c'_i of $|\psi'\rangle$ are given by:

$$c'_i = \langle u_i|\psi'\rangle = \langle u_i|A|\psi\rangle \tag{C-29}$$

If we simply insert the closure relation between A and $|\psi\rangle$, we obtain:

$$\begin{aligned} c'_i &= \langle u_i|A\mathbb{1}|\psi\rangle = \langle u_i|AP_{\{u_j\}}|\psi\rangle \\ &= \sum_j \langle u_i|A|u_j\rangle\langle u_j|\psi\rangle \\ &= \sum_j A_{ij}\, c_j \end{aligned} \tag{C-30}$$

For the $\{|w_\alpha\rangle\}$ basis, we obtain, in the same way:

$$\begin{aligned} c'(\alpha) &= \langle w_\alpha|\psi'\rangle = \langle w_\alpha|A|\psi\rangle \\ &= \int \mathrm{d}\alpha'\, \langle w_\alpha|A|w_{\alpha'}\rangle\langle w_{\alpha'}|\psi\rangle \\ &= \int \mathrm{d}\alpha'\, A(\alpha,\alpha')\; c(\alpha') \end{aligned} \tag{C-31}$$

The matrix expression for $|\psi'\rangle = A|\psi\rangle$ is therefore very simple. We see, for example from (C-30), that the column matrix representing $|\psi'\rangle$ is equal to the product of the column matrix representing $|\psi\rangle$ and the square matrix representing A:

$$\begin{pmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_i \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1j} & \cdots \\ A_{21} & A_{22} & \cdots & A_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_{i1} & A_{i2} & \cdots & A_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \\ \vdots & \vdots & & \vdots & \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \\ c_j \\ \vdots \end{pmatrix} \tag{C-32}$$

C-4-c. Expression for the number $\langle\varphi|A|\psi\rangle$

By inserting the closure relation between $\langle\varphi|$ and A and again between A and $|\psi\rangle$, we obtain:
– for the $\{|u_i\rangle\}$ basis:

$$\begin{aligned} \langle\varphi|A|\psi\rangle &= \langle\varphi|P_{\{u_i\}}AP_{\{u_j\}}|\psi\rangle \\ &= \sum_{i,j} \langle\varphi|u_i\rangle\langle u_i|A|u_j\rangle\langle u_j|\psi\rangle \\ &= \sum_{i,j} b_i^*\, A_{ij}\; c_j \end{aligned} \tag{C-33}$$

– for the $\{\,|\,w_\alpha\,\rangle\,\}$ basis:

$$\begin{aligned}\langle\,\varphi\,|\,A\,|\,\psi\,\rangle &= \langle\,\varphi\,|\,P_{\{w_\alpha\}}AP_{\{w_{\alpha'}\}}\,|\,\psi\,\rangle \\ &= \iint \mathrm{d}\alpha\,\mathrm{d}\alpha'\;\langle\,\varphi\,|\,w_\alpha\,\rangle\,\langle\,w_\alpha\,|\,A\,|\,w_{\alpha'}\,\rangle\,\langle\,w_{\alpha'}\,|\,\psi\,\rangle \\ &= \iint \mathrm{d}\alpha\,\mathrm{d}\alpha'\;\; b^*(\alpha)\;A(\alpha,\alpha')\;\;c(\alpha') \end{aligned} \tag{C-34}$$

In the matrix formalism, the interpretation of these formulas is as follows: $\langle\,\varphi\,|\,A\,|\,\psi\,\rangle$ is a number, that is, a matrix with one row and one column, obtained by multiplying the column matrix representing $|\,\psi\,\rangle$ first by the square matrix representing A and then by the row matrix representing $\langle\,\varphi\,|$. For example, in the $\{\,|\,u_i\,\rangle\,\}$ basis:

$$\langle\,\varphi\,|\,A\,|\,\psi\,\rangle = \begin{pmatrix} b_1^* & b_2^* & \cdots & b_i^* & \cdots \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1j} & \cdots \\ A_{21} & A_{22} & \cdots & A_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_{i1} & A_{i2} & \cdots & A_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \\ \vdots & \vdots & & \vdots & \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \\ c_j \\ \vdots \end{pmatrix} \tag{C-35}$$

Comments:

(*i*) It can be shown in the same way that the bra $\langle\,\varphi\,|\,A$ is represented by a row matrix, the product of the square matrix representing A by the row matrix representing $\langle\,\varphi\,|$ [the first two matrices of the right-hand side of (C-35)]. Again we see the importance of the order of the symbols: the expression $A\,\langle\,\varphi\,|$ would lead to a matrix operation which is undefined (the product of a row matrix by a square matrix).

(*ii*) From a matrix point of view, equation (B-41) which defines $\langle\,\varphi\,|\,A$ merely expresses the associativity of the product of the three matrices that appear in (C-35).

(*iii*) Using the preceding conventions, we express $|\,\psi\,\rangle\,\langle\,\psi\,|$ by a square matrix:

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_i \\ \vdots \end{pmatrix} \begin{pmatrix} c_1^* & c_2^* & \cdots & c_j^* & \cdots \end{pmatrix} = \begin{pmatrix} c_1c_1^* & c_1c_2^* & \cdots & c_1c_j^* & \cdots \\ c_2c_1^* & c_2c_2^* & \cdots & c_2c_j^* & \cdots \\ \vdots & \vdots & & \vdots & \\ c_ic_1^* & c_ic_2^* & \cdots & c_ic_j^* & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix} \tag{C-36}$$

This is indeed an operator, while $\langle\,\psi\,|\,\psi\,\rangle$, the product of a column matrix by a row matrix, is a number.

C-4-d. Matrix representation of the adjoint $A^\dagger$ of A

Using (B-49), we obtain easily:

$$(A^\dagger)_{ij} = \langle\,u_i\,|\,A^\dagger\,|\,u_j\,\rangle = \langle\,u_j\,|\,A\,|\,u_i\,\rangle^* = A^*_{ji} \tag{C-37}$$

or

$$A^\dagger(\alpha,\alpha') = \langle\, w_\alpha \,|\, A^\dagger \,|\, w_{\alpha'} \,\rangle = \langle\, w_{\alpha'} \,|\, A \,|\, w_\alpha \,\rangle^* = A^*(\alpha',\alpha) \tag{C-38}$$

Therefore, the matrices representing A and $A^\dagger$ in a given representation are Hermitian conjugates of each other, in the matrix sense: *one passes from one to the other by interchanging rows and columns and then taking the complex conjugate.*

If A is Hermitian, $A^\dagger = A$, and we can then replace $(A^\dagger)_{ij}$ by A_{ij} in (C-37), and $A^\dagger(\alpha,\alpha')$ by $A(\alpha,\alpha')$ in (C-38):

$$A_{ij} = A^*_{ji} \tag{C-39}$$

$$A(\alpha,\alpha') = A^*(\alpha',\alpha) \tag{C-40}$$

A Hermitian operator is therefore represented by a Hermitian matrix, that is, one in which *any two elements which are symmetric with respect to the principal diagonal are complex conjugates of each other.* In particular, for $i = j$ or $\alpha = \alpha'$, (C-39) and (C-40) become:

$$A_{ii} = A^*_{ii} \tag{C-41}$$

$$A(\alpha,\alpha) = A^*(\alpha,\alpha) \tag{C-42}$$

The diagonal elements of a Hermitian matrix are therefore always real numbers.

C-5. Change of representations

C-5-a. Outline of the problem

In a given representation, a ket (or a bra, or an operator) is represented by a matrix. If we change representations, that is, bases, the same ket (or bra, or operator) will be represented by a different matrix. How are these two matrices related?

For the sake of simplicity, we shall assume here that we are going from one discrete orthonormal basis $\{\,|\,u_i\,\rangle\,\}$ to another discrete orthonormal basis $\{\,|\,t_k\,\rangle\,\}$. In § E, we shall study an example of changing from one continuous basis to another continuous basis.

The change of basis is defined by specifying the components $\langle\, u_i \,|\, t_k \,\rangle$ of each of the kets of the new basis in terms of each of the kets of the old one. We shall set:

$$S_{ik} = \langle\, u_i \,|\, t_k \,\rangle \tag{C-43}$$

S is the matrix of the basis change (transformation matrix). If $S^\dagger$ denotes its Hermitian conjugate:

$$(S^\dagger)_{ki} = (S_{ik})^* = \langle\, t_k \,|\, u_i \,\rangle \tag{C-44}$$

The following calculations can be performed very easily, and without memorization, by using the two closure relations:

$$P_{\{u_i\}} = \sum_i |\, u_i \,\rangle\, \langle\, u_i \,| = \mathbb{1} \tag{C-45}$$

$$P_{\{t_k\}} = \sum_k |\, t_k \,\rangle\, \langle\, t_k \,| = \mathbb{1} \tag{C-46}$$

and the two orthonormalization relations:

$$\langle u_i \,|\, u_j \rangle = \delta_{ij} \tag{C-47}$$
$$\langle t_k \,|\, t_l \rangle \ = \delta_{kl} \tag{C-48}$$

Comment:

The transformation matrix, S, is unitary (Complement C_{II}). That is, it satisfies:

$$S^\dagger S = S S^\dagger = I \tag{C-49}$$

where I is the unit matrix. Indeed, we see that:

$$\begin{aligned}(S^\dagger S)_{kl} &= \sum_i S^\dagger_{ki} S_{il} = \sum_i \langle t_k \,|\, u_i \rangle \langle u_i \,|\, t_l \rangle \\ &= \langle t_k \,|\, t_l \rangle = \delta_{kl}\end{aligned} \tag{C-50}$$

In the same way:

$$\begin{aligned}(S S^\dagger)_{ij} &= \sum_k S_{ik} S^\dagger_{kj} = \sum_k \langle u_i \,|\, t_k \rangle \langle t_k \,|\, u_j \rangle \\ &= \langle u_i \,|\, u_j \rangle = \delta_{ij}\end{aligned} \tag{C-51}$$

C-5-b. Transformation of the components of a ket

To obtain the components $\langle t_k \,|\, \psi \rangle$ of a ket $|\, \psi \rangle$ in the new basis from its components $\langle u_i \,|\, \psi \rangle$ in the old basis, one simply inserts (C-45) between $\langle t_k \,|$ and $|\, \psi \rangle$:

$$\begin{aligned}\langle t_k \,|\, \psi \rangle &= \langle t_k \,|\, \mathbb{1} \,|\, \psi \rangle = \langle t_k \,|\, P_{\{u_i\}} \,|\, \psi \rangle \\ &= \sum_i \langle t_k \,|\, u_i \rangle \langle u_i \,|\, \psi \rangle \\ &= \sum_i S^\dagger_{ki} \langle u_i \,|\, \psi \rangle\end{aligned} \tag{C-52}$$

The inverse expressions can be derived in the same way, using (C-46):

$$\begin{aligned}\langle u_i \,|\, \psi \rangle &= \langle u_i \,|\, \mathbb{1} \,|\, \psi \rangle = \langle u_i \,|\, P_{\{t_k\}} \,|\, \psi \rangle \\ &= \sum_k \langle u_i \,|\, t_k \rangle \langle t_k \,|\, \psi \rangle \\ &= \sum_k S_{ik} \langle t_k \,|\, \psi \rangle\end{aligned} \tag{C-53}$$

C-5-c. Transformation of the components of a bra

The principle of the calculation is exactly the same. For example:

$$\begin{aligned}\langle\,\psi\,|\,t_k\,\rangle &= \langle\,\psi\,|\,\mathbb{1}\,|\,t_k\,\rangle = \langle\,\psi\,|\,P_{\{u_i\}}\,|\,t_k\,\rangle \\ &= \sum_i \langle\,\psi\,|\,u_i\,\rangle\,\langle\,u_i\,|\,t_k\,\rangle \\ &= \sum_i \langle\,\psi\,|\,u_i\,\rangle\,S_{ik}\end{aligned} \tag{C-54}$$

C-5-d. Transformation of the matrix elements of an operator

If, in $\langle\,t_k\,|\,A\,|\,t_l\,\rangle$, we insert (C-45) between $\langle\,t_k\,|$ and A, and again between A and $|\,t_l\,\rangle$, we obtain:

$$\begin{aligned}\langle\,t_k\,|\,A\,|\,t_l\,\rangle &= \langle\,t_k\,|\,P_{\{u_i\}}A\,P_{\{u_j\}}\,|\,t_l\,\rangle \\ &= \sum_{i,j} \langle\,t_k\,|\,u_i\,\rangle\,\langle\,u_i\,|\,A\,|\,u_j\,\rangle\,\langle\,u_j\,|\,t_l\,\rangle\end{aligned} \tag{C-55}$$

that is:

$$A_{kl} = \sum_{i,j} S^\dagger_{ki}\,A_{ij}\,S_{jl} \tag{C-56}$$

In the same way:

$$\begin{aligned}A_{ij} &= \langle\,u_i\,|\,A\,|\,u_j\,\rangle = \langle\,u_i\,|\,P_{\{t_k\}}A\,P_{\{t_l\}}\,|\,u_j\,\rangle \\ &= \sum_{k,l} \langle\,u_i\,|\,t_k\,\rangle\,\langle\,t_k\,|\,A\,|\,t_l\,\rangle\,\langle\,t_l\,|\,u_j\,\rangle \\ &= \sum_{k,l} S_{ik}\,A_{kl}\,S^\dagger_{lj}\end{aligned} \tag{C-57}$$

D. Eigenvalue equations. Observables

D-1. Eigenvalues and eigenvectors of an operator

D-1-a. Definitions

$|\,\psi\,\rangle$ is said to be an eigenvector (or eigenket) of the linear operator A if:

$$\boxed{A\,|\,\psi\,\rangle = \lambda\,|\,\psi\,\rangle} \tag{D-1}$$

where λ is a complex number. We are going to study a certain number of properties of equation (D-1), *the eigenvalue equation of the linear operator A*. In general, this equation possesses solutions only when λ takes on certain values, called *eigenvalues* of A. The set of the eigenvalues is called the *spectrum* of A.

Note that, if $|\,\psi\,\rangle$ is an eigenvector of A with the eigenvalue λ, $\alpha\,|\,\psi\,\rangle$ (where α is an arbitrary complex number) is also an eigenvector of A with the same eigenvalue:

$$A(\alpha\,|\,\psi\,\rangle) = \alpha A\,|\,\psi\,\rangle = \alpha\lambda\,|\,\psi\,\rangle = \lambda(\alpha\,|\,\psi\,\rangle) \tag{D-2}$$

To rid ourselves of this ambiguity, we could agree to normalize the eigenvectors to 1:

$$\langle \psi \,|\, \psi \rangle = 1 \tag{D-3}$$

But this does not completely remove the ambiguity, since $e^{i\theta} \,|\, \psi \rangle$, where θ is an arbitrary real number, has the same norm as $|\, \psi \rangle$. We shall see below that, in quantum mechanics, the physical predictions obtained using $|\, \psi \rangle$ or $e^{i\theta} \,|\, \psi \rangle$ are identical.

The eigenvalue λ is called *non-degenerate* (or *simple*) when its corresponding eigenvector is unique to within a constant factor, that is, when all its associated eigenkets are collinear. On the other hand, if there exist at least two linearly independent kets which are eigenvectors of A with the same eigenvalue, this eigenvalue is said to be *degenerate*. Its *degree (or order) of degeneracy* is then the number of linearly independent eigenvectors associated with it (the degree of degeneracy of an eigenvalue can be finite or infinite). For example, if λ is g-fold degenerate, there correspond to it g independent kets $|\, \psi^i \rangle$ ($i = 1, 2, \ldots g$) such that:

$$A \,|\, \psi^i \rangle = \lambda \,|\, \psi^i \rangle \tag{D-4}$$

But then every ket $|\, \psi \rangle$ of the form:

$$|\, \psi \rangle = \sum_{i=1}^{g} c_i \,|\, \psi^i \rangle \tag{D-5}$$

is an eigenvector of A with the eigenvalue λ, whatever the coefficients c_i, since:

$$A \,|\, \psi \rangle = \sum_{i=1}^{g} c_i \, A \,|\, \psi^i \rangle = \lambda \sum_{i=1}^{g} c_i \,|\, \psi^i \rangle = \lambda \,|\, \psi \rangle \tag{D-6}$$

Consequently, the set of eigenkets of A associated with λ constitutes a *g-dimensional vector space* (which can be infinite-dimensional), called the "*eigensubspace*" of the eigenvalue λ. In particular, it is equivalent to say that λ is non-degenerate or to say that its degree of degeneracy is $g = 1$.

To illustrate these definitions, let us choose the example of a projector (§ B-3-b): $P_\psi = |\, \psi \rangle \langle \psi \,|$ (with $\langle \psi \,|\, \psi \rangle = 1$). Its eigenvalue equation is written:

$$P_\psi \,|\, \varphi \rangle = \lambda \,|\, \varphi \rangle$$

that is:

$$|\, \psi \rangle \langle \psi \,|\, \varphi \rangle = \lambda \,|\, \varphi \rangle \tag{D-7}$$

The ket on the left-hand side is always collinear with $|\, \psi \rangle$, or zero. Consequently, the eigenvectors of P_ψ are: on the one hand, $|\, \psi \rangle$ itself, with an eigenvalue of $\lambda = 1$; on the other hand, all the kets $|\, \varphi \rangle$ orthogonal to $|\, \psi \rangle$, for which the associated eigenvalue is $\lambda = 0$. The spectrum of P_ψ therefore includes only two values: 1 and 0. The first one is simple, the second, infinitely degenerate (if the state space considered is infinite-dimensional). The eigensubspace associated with $\lambda = 0$ is the supplement[7] of $|\, \psi \rangle$ (see § D-2-c).

[7] In a vector space $\mathscr{E}$, two subspaces $\mathscr{E}_1$ and $\mathscr{E}_2$ are said to be supplementary if all kets $|\, \psi \rangle$ of $\mathscr{E}$ can be written $|\, \psi \rangle = |\, \psi_1 \rangle + |\, \psi_2 \rangle$ where $|\, \psi_1 \rangle$ and $|\, \psi_2 \rangle$ belong, respectively, to $\mathscr{E}_1$ and $\mathscr{E}_2$, and if $\mathscr{E}_1$ and $\mathscr{E}_2$ are disjoint (no common non-zero ket; the expansion $|\, \psi \rangle = |\, \psi_1 \rangle + |\, \psi_2 \rangle$ is then unique). Actually, there

Comments:

(*i*) Taking the Hermitian conjugate of both sides of equation (D-1), we obtain:

$$\langle \psi | A^\dagger = \lambda^* \langle \psi | \tag{D-8}$$

Therefore, if $|\psi\rangle$ is an eigenket of A with an eigenvalue λ, it can also be said that $\langle \psi |$ is an eigenbra of $A^\dagger$ with an eigenvalue λ^*. However, let us stress the fact that, except in the case where A is Hermitian (§ D-2-a), nothing can be said *a priori* about $\langle \psi | A$.

(*ii*) To be completely rigorous, one should solve the eigenvalue equation (D-1) in the space $\mathscr{E}$. That is, one should consider only those eigenvectors $|\psi\rangle$ which have a finite norm. In fact, we shall be obliged to use operators for which the eigenkets do not satisfy this condition (§ E). Therefore, we shall grant that vectors which are solutions of (D-1) can be "generalized kets".

D-1-b. Finding the eigenvalues and eigenvectors of an operator

Given a linear operator A, how does one find all its eigenvalues and the corresponding eigenvectors? We are concerned with this question from a purely practical point of view. We shall consider the case where the state space is of finite dimension N, and we shall grant that the results can be generalized to an infinite-dimensional state space.

Let us choose a representation, for example $\{|u_i\rangle\}$, and let us project the vector equation (D-1) onto the various orthonormal basis vectors $|u_i\rangle$:

$$\langle u_i | A | \psi \rangle = \lambda \langle u_i | \psi \rangle \tag{D-9}$$

Inserting the closure relation between A and $|\psi\rangle$, we obtain:

$$\sum_j \langle u_i | A | u_j \rangle \langle u_j | \psi \rangle = \lambda \langle u_i | \psi \rangle \tag{D-10}$$

With the usual notation:

$$\begin{aligned} \langle u_i | \psi \rangle &= c_i \\ \langle u_i | A | u_j \rangle &= A_{ij} \end{aligned} \tag{D-11}$$

equations (D-10) can be written:

$$\sum_j A_{ij}\, c_j = \lambda\, c_i \tag{D-12}$$

or

$$\sum_j [A_{ij} - \lambda \delta_{ij}]\, c_j = 0 \tag{D-13}$$

(D-13) can be considered to be a system of equations where the unknowns are the c_j, the components of the eigenvector in the chosen representation. This system is linear and homogeneous.

---

exists an infinity of sub-subspaces $\mathscr{E}_2$ supplementary to a given subspace $\mathscr{E}_1$. One can fix $\mathscr{E}_2$ by forcing it to be orthogonal to $\mathscr{E}_1$. This shall be done throughout this book, even though the word "orthogonal" will not be explicitly written before supplement.

Example: In ordinary three-dimensional space, if $\mathscr{E}_1$ is a plane P, $\mathscr{E}_2$ can be any arbitrary straight line, not contained in P. The orthogonal supplement of $\mathscr{E}_1$ is the straight line passing through the origin and orthogonal to P.

α. The characteristic equation

The system (D-13) consists of N equations ($i = 1, 2, ..., N$) with N unknowns c_j ($j = 1, 2, ..., N$). Since it is linear and homogeneous, it has a non-trivial solution (the trivial solution is the one for which all the c_j are zero) if and only if the determinant of the coefficients is zero. This condition is written:

$$\boxed{\text{Det } [\mathscr{A} - \lambda I] = 0} \tag{D-14}$$

where $\mathscr{A}$ is the $N \times N$ matrix of elements A_{ij} and I is the unit matrix.

Equation (D-14), called the characteristic equation (or secular equation), enables us to determine all the eigenvalues of the operator A, that is, its spectrum. (D-14) can be written explicitly in the form:

$$\begin{vmatrix} A_{11} - \lambda & A_{12} & A_{13} & \cdots & A_{1N} \\ A_{21} & A_{22} - \lambda & A_{23} & \cdots & A_{2N} \\ \vdots & \vdots & \vdots & & \vdots \\ A_{N1} & A_{N2} & A_{N3} & \cdots & A_{NN} - \lambda \end{vmatrix} = 0 \tag{D-15}$$

This is an Nth order equation in λ; consequently, it has N roots, real or complex, distinct or identical. It is easy to show, by performing an arbitrary change of basis, that the characteristic equation is independent of the representation chosen. Therefore, *the eigenvalues of an operator are the roots of its characteristic equation.*

β. Determination of the eigenvectors

Now let us choose an eigenvalue λ_0, a solution of the characteristic equation (D-14), and let us look for the corresponding eigenvectors. We are going to distinguish between two cases:

(i) First, assume that λ_0 is a simple root of the characteristic equation. We can then show that the system (D-13), when $\lambda = \lambda_0$, is comprised of $(N - 1)$ independent equations, the Nth one following from the preceding ones and hence redundant. But we have N unknowns; there is therefore an infinite number of solutions, but all the c_j can be determined in a unique way in terms of one of them, say c_1. If we fix c_1, we obtain for the $(N-1)$ other c_j a system of $(N-1)$ linear, inhomogeneous equations (the "right-hand side" of each equation is the term in c_1) with a non-zero determinant [the $(N - 1)$ equations are independent]. The solution of this system is of the form:

$$c_j = \alpha_j^0 \, c_1 \tag{D-16}$$

since the initial system (D-13) is linear and homogeneous. α_1^0 is, of course, equal to 1 by definition, and the $(N - 1)$ coefficients α_j^0 for $j \neq 1$ are determined from the matrix elements A_{ij} and λ_0. The eigenvectors associated with λ_0 differ only by the value chosen for c_1. They are therefore all given by:

$$|\, \psi_0(c_1) \,\rangle = \sum_j \, \alpha_j^0 \, c_1 |\, u_j \,\rangle = c_1 \, |\, \psi_0 \,\rangle \tag{D-17}$$

with:

$$|\, \psi_0 \,\rangle = \sum_j \, \alpha_j^0 \, |\, u_j \,\rangle \tag{D-18}$$

Therefore, when λ_0 is a simple root of the characteristic equation, only one eigenvector corresponds to it (to within a constant factor): it is a non-degenerate eigenvalue.

(*ii*) When λ_0 is a multiple root of order $q > 1$ of the characteristic equation, there are two possibilities:

– in general, when $\lambda = \lambda_0$, the system (D-13) is still composed of $(N-1)$ independent equations. Only one eigenvector then corresponds to the eigenvalue λ_0. The operator A cannot be diagonalized in this case: the eigenvectors of A are not sufficiently numerous for one to be able to construct with them alone a basis of the state space.

– nevertheless, when $\lambda = \lambda_0$, it may happen that the system (D-13) has only $(N-p)$ independent equations (where p is a number greater than 1 but not larger than q). To the eigenvalue λ_0 there then corresponds an eigensubspace of dimension p, and λ_0 is a p-fold degenerate eigenvalue. Let us assume, for example, that, for $\lambda = \lambda_0$, (D-13) is composed of $(N-2)$ linearly independent equations. These equations enable us to calculate the coefficients c_j in terms of any two of them, for example c_1 and c_2:

$$c_j = \beta_j^0 \, c_1 + \gamma_j^0 \, c_2 \tag{D-19}$$

(obviously: $\beta_1^0 = \gamma_2^0 = 1$; $\gamma_1^0 = \beta_2^0 = 0$). All the eigenvectors associated with λ_0 are then of the form:

$$| \, \psi_0(c_1, c_2) \, \rangle = c_1 \, | \, \psi_0^1 \, \rangle + c_2 \, | \, \psi_0^2 \, \rangle \tag{D-20}$$

with:

$$\begin{aligned} | \, \psi_0^1 \, \rangle &= \sum_j \beta_j^0 \, | \, u_j \, \rangle \\ | \, \psi_0^2 \, \rangle &= \sum_j \gamma_j^0 \, | \, u_j \, \rangle \end{aligned} \tag{D-21}$$

The vectors $| \, \psi_0(c_1, c_2) \, \rangle$ do indeed constitute a two-dimensional vector space, this being characteristic of a two-fold degenerate eigenvalue.

When an operator is Hermitian, it can be shown that the degree of degeneracy p of an eigenvalue λ is always equal to the multiplicity q of the corresponding root in the characteristic equation. Since, in most cases, we shall be studying only Hermitian operators, we shall only need to know the multiplicity of each root of (D-14) to obtain immediately the dimension of the corresponding eigensubspace. Thus, in a space of finite dimension N, a Hermitian operator always has N linearly independent eigenvectors (we shall see later that they can be chosen to be orthonormal): this operator can therefore be diagonalized (§ D-2-b).

D-2. Observables

D-2-a. Properties of the eigenvalues and eigenvectors of a Hermitian operator

We shall now consider the very important case in which the operator A is Hermitian:

$$A^\dagger = A \tag{D-22}$$

(*i*) *The eigenvalues of a Hermitian operator are real.*

Taking the scalar product of the eigenvalue equation (D-1) by $|\,\psi\,\rangle$, we obtain:

$$\langle\,\psi\,|\,A\,|\,\psi\,\rangle = \lambda\,\langle\,\psi\,|\,\psi\,\rangle \tag{D-23}$$

But $\langle\,\psi\,|\,A\,|\,\psi\,\rangle$ is a real number if A is Hermitian, as we see from:

$$\langle\,\psi\,|\,A\,|\,\psi\,\rangle^* = \langle\,\psi\,|\,A^\dagger\,|\,\psi\,\rangle = \langle\,\psi\,|\,A\,|\,\psi\,\rangle \tag{D-24}$$

where the last equation follows from hypothesis (D-22). Since $\langle\,\psi\,|\,A\,|\,\psi\,\rangle$ and $\langle\,\psi\,|\,\psi\,\rangle$ are real, equation (D-23) implies that λ must also be real.

If A is Hermitian, we can, in (D-8), replace A by $A^\dagger$ and λ by λ^*, since we have just shown that λ is real. Thus we obtain:

$$\langle\,\psi\,|\,A = \lambda\,\langle\,\psi\,| \tag{D-25}$$

which shows that $\langle\,\psi\,|$ is also an eigenbra of A with the real eigenvalue λ. Therefore, whatever the ket $|\,\varphi\,\rangle$:

$$\langle\,\psi\,|\,A\,|\,\varphi\,\rangle = \lambda\,\langle\,\psi\,|\,\varphi\,\rangle \tag{D-26}$$

The Hermitian operator A is said to act on the left in (D-26).

(*ii*) *Two eigenvectors of a Hermitian operator corresponding to two different eigenvalues are orthogonal.*

Consider two eigenvectors $|\,\psi\,\rangle$ and $|\,\varphi\,\rangle$ of the Hermitian operator A:

$$A\,|\,\psi\,\rangle = \lambda\,|\,\psi\,\rangle \tag{D-27-a}$$
$$A\,|\,\varphi\,\rangle = \mu\,|\,\varphi\,\rangle \tag{D-27-b}$$

Since A is Hermitian, (D-27-b) can be written in the form:

$$\langle\,\varphi\,|\,A = \mu\,\langle\,\varphi\,| \tag{D-28}$$

Then multiply (D-27-a) by $\langle\,\varphi\,|$ on the left and (D-28) by $|\,\psi\,\rangle$ on the right:

$$\langle\,\varphi\,|\,A\,|\,\psi\,\rangle = \lambda\,\langle\,\varphi\,|\,\psi\,\rangle \tag{D-29-a}$$
$$\langle\,\varphi\,|\,A\,|\,\psi\,\rangle = \mu\,\langle\,\varphi\,|\,\psi\,\rangle \tag{D-29-b}$$

Subtracting (D-29-b) from (D-29-a), we find:

$$(\lambda - \mu)\,\langle\,\varphi\,|\,\psi\,\rangle = 0 \tag{D-30}$$

Consequently, if $(\lambda - \mu) \neq 0$, $|\,\varphi\,\rangle$ and $|\,\psi\,\rangle$ are orthogonal.

D-2-b. Definition of an observable

When $\mathscr{E}$ is finite-dimensional, we have seen (§ D-1-b) that it is always possible to form a basis with the eigenvectors of a Hermitian operator. When $\mathscr{E}$ is infinite-dimensional, this is no longer necessarily the case. This is why it is useful to introduce a new concept, that of an observable.

Consider a Hermitian operator A. For simplicity, we shall assume that the set of its eigenvalues forms a discrete spectrum $\{ a_n ; n = 1, 2, ... \}$, and we shall indicate later the modifications that must be made when all or part of this spectrum is continuous. The degree of degeneracy of the eigenvalue a_n will be denoted by g_n (if $g_n = 1$, a_n is non-degenerate). We shall denote by $| \psi_n^i \rangle$ $(i = 1, 2, ... \ g_n)$ g_n linearly independent vectors chosen in the eigensubspace $\mathscr{E}_n$ of a_n:

$$A | \psi_n^i \rangle = a_n | \psi_n^i \rangle \quad ; \quad i = 1, 2, ..., g_n \tag{D-31}$$

We have just shown that every vector belonging to $\mathscr{E}_n$ is orthogonal to every vector of another subspace $\mathscr{E}_{n'}$, associated with $a_{n'} \neq a_n$; therefore:

$$\langle \psi_n^i | \psi_{n'}^j \rangle = 0 \quad \text{for } n \neq n' \text{ and arbitrary } i \text{ and } j \tag{D-32}$$

Inside each subspace $\mathscr{E}_n$, the $| \psi_n^i \rangle$ can always be chosen orthonormal, that is, such that:

$$\langle \psi_n^i | \psi_n^j \rangle = \delta_{ij} \tag{D-33}$$

If such a choice is made, the result is an *orthonormal system of eigenvectors of A*: the $| \psi_n^i \rangle$ satisfy the relations:

$$\boxed{\langle \psi_n^i | \psi_{n'}^{i'} \rangle = \delta_{nn'} \delta_{ii'}} \tag{D-34}$$

obtained by regrouping (D-32) and (D-33).

By definition, the Hermitian operator A is an *observable* if this orthonormal system of vectors *forms a basis* in the state space. This can be expressed by the closure relation:

$$\boxed{\sum_{n=1}^{\infty} \sum_{i=1}^{g_n} | \psi_n^i \rangle \langle \psi_n^i | = \mathbb{1}} \tag{D-35}$$

Comments:

(*i*) Since the g_n vectors $| \psi_n^i \rangle$ $(i = 1, 2, ..., g_n)$ which span the eigensubspace $\mathscr{E}_n$ of a_n are orthonormal, the projector P_n onto this subspace $\mathscr{E}_n$ can be written (*cf.* § B-3-b-γ):

$$P_n = \sum_{i=1}^{g_n} | \psi_n^i \rangle \langle \psi_n^i | \tag{D-36-a}$$

The observable A is then given by:

$$A = \sum_n a_n P_n \tag{D-36-b}$$

(it is easy to verify that the action of both sides of this equation on all the kets $| \psi_n^i \rangle$ gives the same result).

(*ii*) Relation (D-35) can be generalized to include cases where the spectrum of eigenvalues is continuous by using the rules given in table (II-3). For example, consider a Hermitian operator whose spectrum is composed of a discrete part $\{a_n$ (degree of degeneracy $g_n)\}$ and a continuous part $a(\nu)$ (assumed to be non-degenerate):

$$A \,|\, \psi_n^i \,\rangle = a_n \,|\, \psi_n^i \,\rangle \quad ; \quad n = 1,\, 2,\, ... \qquad i = 1,\, 2,\, ...,\, g_n \tag{D-37-a}$$

$$A \,|\, \psi_\nu \,\rangle = a(\nu) \,|\, \psi_\nu \,\rangle \quad ; \quad \nu_1 < \nu < \nu_2 \tag{D-37-b}$$

The vectors can always be chosen in such a way that they form an "orthonormal" system:

$$\begin{aligned} &\langle\, \psi_n^i \,|\, \psi_{n'}^{i'} \,\rangle = \delta_{nn'}\delta_{ii'} \\ &\langle\, \psi_\nu \,|\, \psi_{\nu'} \,\rangle = \delta(\nu - \nu') \\ &\langle\, \psi_n^i \,|\, \psi_\nu \,\rangle = 0 \end{aligned} \tag{D-38}$$

A will be said to be an observable if this system forms a basis, that is, if:

$$\sum_n \sum_{i=1}^{g_n} |\, \psi_n^i \,\rangle\, \langle\, \psi_n^i \,| \; + \int_{\nu_1}^{\nu_2} \mathrm{d}\nu \,|\, \psi_\nu \,\rangle\, \langle\, \psi_\nu \,| = \mathbb{1} \tag{D-39}$$

D-2-c. Example: the projector P_ψ

Let us show that $P_\psi = |\, \psi \,\rangle\, \langle\, \psi \,|$ (with $\langle\, \psi \,|\, \psi \,\rangle = 1$) is an observable. We have already pointed out (§ B-4-e) that it is Hermitian, and that its eigenvalues are 1 and 0 (§ D-1-a); the first one is simple (associated eigenvector: $|\, \psi \,\rangle$), the second one is infinitely degenerate (associated eigenvectors: all kets orthogonal to $|\, \psi \,\rangle$).

Consider an arbitrary ket $|\, \varphi \,\rangle$ in the state space. It can always be written in the form:

$$|\, \varphi \,\rangle = P_\psi \,|\, \varphi \,\rangle + (\mathbb{1} - P_\psi) \,|\, \varphi \,\rangle \tag{D-40}$$

Now, $P_\psi \,|\, \varphi \,\rangle$ is an eigenket of P_ψ with the eigenvalue 1. Since $P_\psi^2 = P_\psi$, we have:

$$P_\psi(P_\psi \,|\, \varphi \,\rangle) = P_\psi^2 \,|\, \varphi \,\rangle = P_\psi \,|\, \varphi \,\rangle \tag{D-41}$$

Moreover, $(\mathbb{1} - P_\psi) \,|\, \varphi \,\rangle$ is also an eigenket of P_ψ, but with the eigenvalue 0, as we see from:

$$P_\psi(\mathbb{1} - P_\psi) \,|\, \varphi \,\rangle = (P_\psi - P_\psi^2) \,|\, \varphi \,\rangle = 0 \tag{D-42}$$

Every ket $|\, \varphi \,\rangle$ can thus be expanded on these eigenkets of P_ψ; therefore, P_ψ is an observable.

We shall see in § E-2 two other important examples of observables.

D-3. Sets of commuting observables

D-3-a. Important theorems

α. Theorem I

If two operators A and B commute, and if $|\, \psi \,\rangle$ is an eigenvector of A, $B \,|\, \psi \,\rangle$ is also an eigenvector of A, with the same eigenvalue.

We know that, if $|\psi\rangle$ is an eigenvector of A, we have:

$$A|\psi\rangle = a|\psi\rangle \tag{D-43}$$

Applying B to both sides of this equation, we obtain:

$$BA|\psi\rangle = aB|\psi\rangle \tag{D-44}$$

Since we assumed that A and B commute, we also have, replacing BA on the left-hand side by AB:

$$A(B|\psi\rangle) = a(B|\psi\rangle) \tag{D-45}$$

This equation expresses the fact that $B|\psi\rangle$ is an eigenvector of A, with the eigenvalue a; the theorem is therefore proved.

Two cases may then arise:

(*i*) If a is a nondegenerate eigenvalue, all the eigenvectors associated with it are by definition colinear, and $B|\psi\rangle$ is necessarily proportional to $|\psi\rangle$. Thus $|\psi\rangle$ is also an eigenvector of B.

(*ii*) If a is a degenerate eigenvalue, it can only be said that $B|\psi\rangle$ belongs to the eigensubspace $\mathscr{E}_a$ of A, corresponding to the eigenvalue a. Therefore, for any $|\psi\rangle \in \mathscr{E}_a$, we have:

$$B|\psi\rangle \in \mathscr{E}_a \tag{D-46}$$

$\mathscr{E}_a$ is said to be globally invariant (or stable) under the action of B. Theorem I can also be stated in another form:

Theorem I': If two operators A and B commute, every eigensubspace of A is globally invariant under the action of B.

β. Theorem II

If two observables A and B commute, and if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two eigenvectors of A with different eigenvalues, the matrix element $\langle\psi_1|B|\psi_2\rangle$ is zero.

If $|\psi_1\rangle$ and $|\psi_2\rangle$ are eigenvectors of A, we can write:

$$\begin{aligned} A|\psi_1\rangle &= a_1|\psi_1\rangle \\ A|\psi_2\rangle &= a_2|\psi_2\rangle \end{aligned} \tag{D-47}$$

According to theorem I, the fact that A and B commute means that $B|\psi_2\rangle$ is an eigenvector of A, with the eigenvalue a_2. $B|\psi_2\rangle$ is therefore (*cf.* § D-2-a) orthogonal to $|\psi_1\rangle$ (eigenvector of eigenvalue $a_1 \neq a_2$), which can be written:

$$\langle\psi_1|B|\psi_2\rangle = 0 \tag{D-48}$$

The theorem is then proved. Another proof can be given, which does not involve theorem I: since the operator $[A, B]$ is zero, we have:

$$\langle\psi_1|(AB - BA)|\psi_2\rangle = 0 \tag{D-49}$$

Using (D-47) and the Hermiticity of A [*cf.* equation (D-25)], we obtain:

$$\begin{aligned} \langle \psi_1 \,|\, AB \,|\, \psi_2 \rangle &= a_1 \, \langle \psi_1 \,|\, B \,|\, \psi_2 \rangle \\ \langle \psi_1 \,|\, BA \,|\, \psi_2 \rangle &= a_2 \, \langle \psi_1 \,|\, B \,|\, \psi_2 \rangle \end{aligned} \tag{D-50}$$

and (D-49) can be rewritten in the form:

$$(a_1 - a_2) \, \langle \psi_1 \,|\, B \,|\, \psi_2 \rangle = 0 \tag{D-51}$$

Since, by hypothesis, $(a_1 - a_2)$ is not zero, we can deduce (D-48) from this equality.

γ. Theorem III (fundamental)

If two observables A and B commute, one can construct an orthonormal basis of the state space with eigenvectors common to A and B.

Consider two commuting observables, A and B. In order to simplify the notation, we shall assume that their spectra are entirely discrete. Since A is an observable, there exists at least one orthonormal system of eigenvectors of A which forms a basis in the state space. We shall denote these vectors by $|\, u_n^i \,\rangle$:

$$\begin{aligned} A \,|\, u_n^i \,\rangle = a_n \,|\, u_n^i \,\rangle \quad ; \quad & n = 1,\, 2,\, ... \\ & i = 1,\, 2,\, ...,\, g_n \end{aligned} \tag{D-52}$$

g_n is the degree of degeneracy of the eigenvalue a_n, that is, the dimension of the corresponding eigensubspace $\mathscr{E}_n$. We have:

$$\langle \, u_n^i \,|\, u_{n'}^{i'} \,\rangle = \delta_{nn'} \, \delta_{ii'} \tag{D-53}$$

What does the matrix look like which represents B in the $\{\, |\, u_n^i \,\rangle \,\}$ basis? We know (*cf.* theorem II) that the matrix elements $\langle \, u_n^i \,|\, B \,|\, u_{n'}^{i'} \,\rangle$ are zero when $n \neq n'$ (on the other hand, we can say nothing *a priori* about what happens for $n = n'$ and $i \neq i'$). Let us arrange the basis vectors $|\, u_n^i \,\rangle$ in the order:

$|\, u_1^1 \,\rangle$, $|\, u_1^2 \,\rangle$, ..., $|\, u_1^{g_1} \,\rangle$; $|\, u_2^1 \,\rangle$, ..., $|\, u_2^{g_2} \,\rangle$; $|\, u_3^1 \,\rangle$, ...

We then obtain for B a "block-diagonal" matrix, that is, of the form:

$$\begin{array}{c} \\ \mathscr{E}_1 \\ \\ \mathscr{E}_2 \\ \\ \mathscr{E}_3 \\ .. \end{array} \begin{array}{c} \begin{array}{ccccc} \mathscr{E}_1 & & \mathscr{E}_2 & & \mathscr{E}_3 \quad .. \end{array} \\ \left(\begin{array}{c|c|c|c} \ddots \;\; \ddots & & & \\ \ddots \;\; \ddots & 0 & 0 & 0 \\ \hline & \ddots \;\; \ddots \;\; \ddots & & \\ 0 & \ddots \;\; \ddots \;\; \ddots & 0 & 0 \\ & \ddots \;\; \ddots \;\; \ddots & & \\ \hline 0 & 0 & \ddots & 0 \\ \hline 0 & 0 & 0 & \ddots \;\; \ddots \\ & & & \ddots \;\; \ddots \end{array} \right) \end{array} \tag{D-54}$$

(only the dotted square parts contain non-zero matrix elements). The fact that the eigensubspaces $\mathscr{E}_n$ are globally invariant under the action of B (*cf.* § D-3-a-α) is evident from this matrix.

Two cases can then arise:

(*i*) When a_n is a nondegenerate eigenvalue of A, there exists only one eigenvector $|\, u_n \,\rangle$ of A, of eigenvalue a_n (the index i in $|\, u_n \,\rangle$ is then unnecessary): the dimension g_n of $\mathscr{E}_n$ is then equal to 1. In the matrix (D-54), the corresponding "block", then reduces to a 1×1 matrix, that is, to a simple number. In the column associated with $|\, u_n \,\rangle$, all the other matrix elements are zero. This expresses the fact (*cf.* § D-3-a-α-*i*) that $|\, u_n \,\rangle$ is an eigenvector common to A and B.

(*ii*) When a_n is a degenerate eigenvalue of A ($g_n > 1$), the "block" which represents B in $\mathscr{E}_n$ is not, in general, diagonal: the $|\, u_n^i \,\rangle$ are not, in general, eigenvectors of B.

It can be seen, nevertheless, that, since the action of A on each of the g_n vectors $|\, u_n^i \,\rangle$ reduces to a simple multiplication by a_n, the matrix representing the restriction of A to within $\mathscr{E}_n$ is equal to $a_n I$ (where I is the $g_n \times g_n$ unit matrix). This expresses the fact that an arbitrary ket of $\mathscr{E}_n$ is an eigenvector of A with the eigenvalue a_n. The choice in $\mathscr{E}_n$ of a basis such as $\{\, |\, u_n^i \,\rangle \;;\; i = 1,\ 2,\ ...,\ g_n \,\}$ is therefore arbitrary. Whatever this basis, the matrix representing the operator A in $\mathscr{E}_n$ is always diagonal and equal to $a_n I$. We shall use this property to obtain a basis of $\mathscr{E}_n$ composed of vectors that are also eigenvectors of B.

The matrix representing B in $\mathscr{E}_n$, when the basis chosen is:

$\{\, |\, u_n^i \,\rangle \;;\; i = 1,\ 2,\ ...,\ g_n \,\}$,

has for its elements:

$$\beta_{ij}^{(n)} = \langle\, u_n^i \,|\, B \,|\, u_n^j \,\rangle \tag{D-55}$$

This matrix is Hermitian $\left(\beta_{ji}^{(n)*} = \beta_{ij}^{(n)}\right)$, since B is a Hermitian operator. It is therefore diagonalizable, that is, *one can find in $\mathscr{E}_n$ a new basis $\{\, |\, v_n^i \,\rangle \;;\; i = 1,\ 2,\ ...,\ g_n \,\}$ in which B is represented by a diagonal matrix*:

$$\langle\, v_n^i \,|\, B \,|\, v_n^j \,\rangle = \beta_i^{(n)} \delta_{ij} \tag{D-56}$$

This means that the new basis vectors in $\mathscr{E}_n$ are eigenvectors of B:

$$B \,|\, v_n^i \,\rangle = \beta_i^{(n)} \,|\, v_n^i \,\rangle \tag{D-57}$$

As we saw above, these vectors are automatically eigenvectors of A with an eigenvalue a_n since they belong to $\mathscr{E}_n$. Let us stress the fact that *the eigenvectors of A associated with degenerate eigenvalues are not necessarily eigenvectors of B*. What we have just shown is that *it is always possible to choose, in every eigensubspace of A, a basis of eigenvectors common to A and B*.

If we perform this operation in all the subspaces $\mathscr{E}_n$, we obtain a basis of $\mathscr{E}$, formed by eigenvectors common to A and B. The theorem is therefore proved.

Comments:

(*i*) From now on, we shall denote by $|\, u_{n,p}^i \,\rangle$ the eigenvectors common to A and B.

$$\begin{aligned} A \,|\, u_{n,p}^i \,\rangle &= a_n \,|\, u_{n,p}^i \,\rangle \\ B \,|\, u_{n,p}^i \,\rangle &= b_p \,|\, u_{n,p}^i \,\rangle \end{aligned} \tag{D-58}$$

The indices n and p which appear in $|\, u^i_{n,p}\,\rangle$ enable us to specify the eigenvalues a_n and b_p of A and B. The additional index i will eventually be used to distinguish between the different basis vectors which correspond to the same eigenvalues a_n and b_p (§ D-3-b below).

(*ii*) The converse of theorem III is very simple to prove: if there exists a basis of eigenvectors common to A and B, these two observables commute. From (D-58), it is easy to deduce:

$$\begin{aligned} AB\,|\,u^i_{n,p}\,\rangle &= b_p\,A\,|\,u^i_{n,p}\,\rangle = b_p\,a_n\,|\,u^i_{n,p}\,\rangle \\ BA\,|\,u^i_{n,p}\,\rangle &= a_n\,B\,|\,u^i_{n,p}\,\rangle = a_n\,b_p\,|\,u^i_{n,p}\,\rangle \end{aligned} \tag{D-59}$$

and, subtracting these equations:

$$[A,\,B]\,|\,u^i_{n,p}\,\rangle = 0 \tag{D-60}$$

This relation is valid for all i, n and p. Since, by hypothesis, the vectors $|\,u^i_{n,p}\,\rangle$ form a basis, (D-60) entails $[A,\,B] = 0$.

(*iii*) We shall occasionally solve the eigenvalue equation of an observable C such that:

$$C = A + B \quad \text{avec} \quad [A,\,B] = 0 \tag{D-61}$$

where A and B are also observables.

When one has found a basis $\{\,|\,u^i_{n,p}\,\rangle\,\}$ of eigenvectors common to A and B, the problem is solved, since we see immediately that $|\,u^i_{n,p}\,\rangle$ is also an eigenvector of C, with an eigenvalue $a_n + b_p$. The fact that $\{\,|\,u^i_{n,p}\,\rangle\,\}$ constitutes a basis is obviously essential: this allows us, for example, to show simply that all the eigenvalues of C are of the form $a_n + b_p$.

D-3-b. Complete[8] sets of commuting observables (C.S.C.O.)[9]

Consider an observable A and a basis of $\mathscr{E}$ composed of eigenvectors $|\,u^i_n\,\rangle$ of A. If none of the eigenvalues of A is degenerate, the various basis vectors of $\mathscr{E}$ can be labelled by the eigenvalue a_n (the index i in $|\,u^i_n\,\rangle$ being in this case unnecessary). All the eigensubspaces $\mathscr{E}_n$ are then one-dimensional. Therefore, specifying the eigenvalue determines in a unique way the corresponding eigenvector (to within a constant factor). In other words, there exists only one basis of $\mathscr{E}$ formed by the eigenvectors of A (we shall not consider here as distinct two bases whose vectors are proportional). It is then said that the observable A constitutes, by itself, a C.S.C.O.

If, on the other hand, one or several eigenvalues of A are degenerate, the situation is different. Specifying a_n is no longer always sufficient to characterize a basis vector, since several independent vectors correspond to any degenerate eigenvalue. In this case,

[8] The word "complete" is used here in a sense which is totally unrelated to those referred to in the note 3 of § A-2-a, p. 91. This use of the word "complete" is customary in quantum mechanics.

[9] To have a good understanding of the important concepts introduced in this section, the reader should apply them to a concrete example such as the one discussed in Complement H_{II} (solved exercises 11 and 12).

the basis of eigenvectors of A is obviously not unique. One can choose any basis inside each of the eigensubspaces $\mathscr{E}_n$ of dimension greater than 1.

Let us then choose another observable B which commutes with A, and let us construct an orthonormal basis of eigenvectors common to A and B. By definition, A and B form a C.S.C.O. if this basis is unique (to within a phase factor for each of the basis vectors), that is, if, to each of the possible pairs of eigenvalues $\{ a_n, b_p \}$, there corresponds only one basis vector.

Comment:

In § D-3-a, we constructed a basis of eigenvectors common to A and B by solving the eigenvalue equation of B inside each eigensubspace $\mathscr{E}_n$ of $\mathscr{E}$. For A and B to constitute a C.S.C.O., it is necessary and sufficient that, inside each of these subspaces, all the g_n eigenvalues of B be distinct. Since all the vectors of $\mathscr{E}_n$ correspond to the same eigenvalue a_n of A, the g_n vectors $| v_n^i \rangle$ can then be distinguished by the eigenvalue of B which is associated with them. Note that it is not necessary that all the eigenvalues of B be non-degenerate: vectors $| v_n^i \rangle$ belonging to two distinct subspaces $\mathscr{E}_n$ can have the same eigenvalue for B. Moreover, if all the eigenvalues of B were non-degenerate, B alone would constitute a C.S.C.O.

If, for at least one of the possible pairs $\{ a_n, b_p \}$, there exist several independent vectors which are eigenvectors of A and B with these eigenvalues, the set $\{ A, B \}$ is not complete. Let us add to it, then, a third observable C, which commutes with both A and B. We can then use the same argument as in § D-3-a above, generalizing it in the following way. When to a pair $\{ a_n, b_p \}$, there corresponds only one vector, this vector is necessarily an eigenvector of C. If there are several vectors, they form an eigensubspace $\mathscr{E}_{n,p}$, in which it is possible to choose a basis formed by vectors which are also eigenvectors of C. One can thus construct, in the state space, an orthonormal basis formed by eigenvectors common to A, B and C. A, B and C form a C.S.C.O. if this basis is unique (to within multiplicative factors). Specifying a possible set of eigenvalues $\{ a_n, b_p, c_r \}$ of A, B, C then characterizes only one of the vectors of this basis. If this is not the case, one adds to A, B, C an observable D which commutes with each of these three operators, and so on. In general, we are thus led to the following:
By definition, a set of observables A, B, C... is called a complete set of commuting observables if

(*i*) *all the observables A, B, C... commute by pairs,*

(*ii*) *specifying the eigenvalues of all the operators A, B, C... determines a unique* (to within a multiplicative factor) *common eigenvector.*

An equivalent way of saying this is the following:
A set of observables A, B, C... is a complete set of commuting observables if there exists a unique orthonormal basis of common eigenvectors (to within phase factors).

C.S.C.O.'s play an important role in quantum mechanics. We shall see numerous examples of them (see, in particular, § E-2-d).

Comments:

(*i*) If $\{ A, B \}$ is a C.S.C.O., another C.S.C.O. can be obtained by adding to it any observable C, on the condition, of course, that it commutes with A and B. However, it is generally understood that one is confined to "minimal" sets, that is, those which cease to be complete when any one of the observables is omitted.

(*ii*) Let $\{ A, B, C... \}$ be a C.S.C.O.. Since the specification of the eigenvalues a_n, b_n, c_r... determines a unique ket of the corresponding basis (to within a constant factor), this ket is sometimes denoted by $| a_n, b_p, c_r, ... \rangle$.

(*iii*) For a given physical system, there exist several distinct complete sets of commuting observables. We shall see a particular example of this in § E-2-d.

E. Two important examples of representations and observables

In this paragraph, we shall return to the $\mathscr{F}$-space of wave functions of a particle, or, more exactly, to the state space $\mathscr{E}_{\mathbf{r}}$ which is associated with it, and which we shall define in the following way. Let there correspond to every wave function $\psi(\mathbf{r})$ a ket $| \psi \rangle$ belonging to $\mathscr{E}_{\mathbf{r}}$; this correspondence is linear. Moreover, the scalar product of two kets coincides with that of the functions which are associated with them:

$$\langle \varphi | \psi \rangle = \int \mathrm{d}^3 r \; \varphi^*(\mathbf{r}) \; \psi(\mathbf{r}) \tag{E-1}$$

$\mathscr{E}_{\mathbf{r}}$ is thus the state space of a (spinless) particle.

We are going to define and study, in this space, two representations and two operators which are particularly important. In Chapter III we shall associate them with the position and the momentum of the particle under consideration. They will enable us, moreover, to apply and illustrate the concepts which we have introduced in the preceding sections.

E-1. The $\{ | \mathbf{r} \rangle \}$ and $\{ | \mathbf{p} \rangle \}$ representations

E-1-a. Definition

In §§ A-3-a and A-3-b, we introduced two particular "bases" of $\mathscr{F}$: $\{ \xi_{\mathbf{r}_0}(\mathbf{r}) \}$ and $\{ v_{\mathbf{p}_0}(\mathbf{r}) \}$. They are not composed of functions belonging to $\mathscr{F}$:

$$\xi_{\mathbf{r}_0}(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0) \tag{E-2-a}$$

$$v_{\mathbf{p}_0}(\mathbf{r}) = (2\pi\hbar)^{-3/2} \; \mathrm{e}^{\frac{i}{\hbar}\mathbf{p}_0 \cdot \mathbf{r}} \tag{E-2-b}$$

However, every sufficiently regular square-integrable function can be expanded in one or the other of these "bases".

This is why we shall remove the quotation marks and associate a ket with each of the functions of these bases (*cf.* § B-2-c). The ket associated with $\xi_{\mathbf{r}_0}(\mathbf{r})$ will be denoted simply by $| \mathbf{r}_0 \rangle$, and that associated with $v_{\mathbf{p}_0}(\mathbf{r})$, by $| \mathbf{p}_0 \rangle$:

$$\xi_{\mathbf{r}_0}(\mathbf{r}) \Longleftrightarrow | \mathbf{r}_0 \rangle \tag{E-3-a}$$

$$v_{\mathbf{p}_0}(\mathbf{r}) \Longleftrightarrow | \mathbf{p}_0 \rangle \tag{E-3-b}$$

Using the bases $\{\xi_{\mathbf{r}_0}(\mathbf{r})\}$ and $\{v_{\mathbf{p}_0}(\mathbf{r})\}$ of $\mathscr{F}$, we thus define in $\mathscr{E}_\mathbf{r}$ two representations: the $\{|\mathbf{r}_0\rangle\}$ representation and the $\{|\mathbf{p}_0\rangle\}$ representation. A basis vector of the first one is characterized by three "continuous indices" x_0, y_0 and z_0, which are the coordinates of a point in three-dimensional space; for the second, the three indices are also the components of an ordinary vector $\mathbf{p}_0$.

E-1-b. Orthonormalization and closure relations

Let us calculate $\langle \mathbf{r}_0 | \mathbf{r}'_0 \rangle$. Using the definition of the scalar product in $\mathscr{E}_\mathbf{r}$:

$$\langle \mathbf{r}_0 | \mathbf{r}'_0 \rangle = \int \mathrm{d}^3 r \ \xi^*_{\mathbf{r}_0}(\mathbf{r})\, \xi_{\mathbf{r}'_0}(\mathbf{r}) = \delta(\mathbf{r}_0 - \mathbf{r}'_0) \tag{E-4-a}$$

where relation (A-55) has been used. In the same way:

$$\langle \mathbf{p}_0 | \mathbf{p}'_0 \rangle = \int \mathrm{d}^3 r \ v^*_{\mathbf{p}_0}(\mathbf{r})\, v_{\mathbf{p}'_0}(\mathbf{r}) = \delta(\mathbf{p}_0 - \mathbf{p}'_0) \tag{E-4-b}$$

using (A-47). The bases which we have just defined are therefore orthonormal in the extended sense.

The fact that the set of the $|\mathbf{r}_0\rangle$ or that of the $|\mathbf{p}_0\rangle$ constitutes a basis in $\mathscr{E}_\mathbf{r}$ can be expressed by a closure relation in $\mathscr{E}_\mathbf{r}$. This is written in an analogous manner to (C-10), integrating here, however, over three indices instead of one.

We therefore have the fundamental relations:

$\langle \mathbf{r}_0 \| \mathbf{r}'_0 \rangle = \delta(\mathbf{r}_0 - \mathbf{r}'_0)$ (a)	$\langle \mathbf{p}_0 \| \mathbf{p}'_0 \rangle = \delta(\mathbf{p}_0 - \mathbf{p}'_0)$ (c)
$\int \mathrm{d}^3 r_0 \ \| \mathbf{r}_0 \rangle \langle \mathbf{r}_0 \| = \mathbb{1}$ (b)	$\int \mathrm{d}^3 p_0 \ \| \mathbf{p}_0 \rangle \langle \mathbf{p}_0 \| = \mathbb{1}$ (d)

(E-5)

E-1-c. Components of a ket

Consider an arbitrary ket $|\psi\rangle$, corresponding to the wave function $\psi(\mathbf{r})$. The preceding closure relations enable us to write it in either of these two forms:

$$|\psi\rangle = \int \mathrm{d}^3 r_0 \ |\mathbf{r}_0\rangle \langle \mathbf{r}_0 | \psi \rangle \tag{E-6-a}$$

$$|\psi\rangle = \int \mathrm{d}^3 p_0 \ |\mathbf{p}_0\rangle \langle \mathbf{p}_0 | \psi \rangle \tag{E-6-b}$$

The coefficients $\langle \mathbf{r}_0 | \psi \rangle$ and $\langle \mathbf{p}_0 | \psi \rangle$ can be calculated using the formulas:

$$\langle \mathbf{r}_0 | \psi \rangle = \int \mathrm{d}^3 r \ \xi^*_{\mathbf{r}_0}(\mathbf{r})\, \psi(\mathbf{r}) \tag{E-7-a}$$

$$\langle \mathbf{p}_0 | \psi \rangle = \int \mathrm{d}^3 r \ v^*_{\mathbf{p}_0}(\mathbf{r})\, \psi(\mathbf{r}) \tag{E-7-b}$$

We then find:

$$\langle \mathbf{r}_0 | \psi \rangle = \psi(\mathbf{r}_0) \tag{E-8-a}$$

$$\langle \mathbf{p}_0 | \psi \rangle = \overline{\psi}(\mathbf{p}_0) \tag{E-8-b}$$

where $\overline{\psi}(\mathbf{p}$ is the Fourier transform of $\psi(\mathbf{r})$.

The value $\psi(\mathbf{r}_0)$ of the wave function at the point $\mathbf{r}_0$ is thus shown to be the component of the ket $|\,\psi\,\rangle$ on the basis vector $|\,\mathbf{r}_0\,\rangle$ of the $\{\,|\,\mathbf{r}_0\,\rangle\,\}$ representation. The "wave function in momentum space" $\overline{\psi}(\mathbf{p})$ can be interpreted analogously. The possibility of characterizing $|\,\psi\,\rangle$ by $\psi(\mathbf{r})$ is thus simply a special case of the results of § C-3-a.

For example, for $|\,\psi\,\rangle = |\,\mathbf{p}_0\,\rangle$, formula (E-8-a) gives:

$$\langle\,\mathbf{r}_0\,|\,\mathbf{p}_0\,\rangle = v_{\mathbf{p}_0}(\mathbf{r}_0) = (2\pi\hbar)^{-3/2}\,\mathrm{e}^{\frac{i}{\hbar}\mathbf{p}_0\cdot\mathbf{r}_0} \tag{E-9}$$

For $|\,\psi\,\rangle = |\,\mathbf{r'}_0\,\rangle$, the result is indeed in agreement with the orthonormalization relation (E-5-a):

$$\langle\,\mathbf{r}_0\,|\,\mathbf{r'}_0\,\rangle = \xi_{\mathbf{r'}_0}(\mathbf{r}_0) = \delta(\mathbf{r}_0 - \mathbf{r'}_0) \tag{E-10}$$

Now that we have reinterpreted the wave function $\psi(\mathbf{r})$ and its Fourier transform $\overline{\psi}(\mathbf{p})$, we shall denote the basis vectors of the two representations we are studying here by $|\,\mathbf{r}\,\rangle$ and $|\,\mathbf{p}\,\rangle$, instead of $|\,\mathbf{r}_0\,\rangle$ and $|\,\mathbf{p}_0\,\rangle$. Formulas (E-8) can then be written:

$$\langle\,\mathbf{r}\,|\,\psi\,\rangle = \psi(\mathbf{r}) \tag{E-8-a}$$
$$\langle\,\mathbf{p}\,|\,\psi\,\rangle = \overline{\psi}(\mathbf{p}) \tag{E-8-b}$$

and the orthonormalization and closure relations (E-5) become:

$$\langle\,\mathbf{r}\,|\,\mathbf{r'}\,\rangle = \delta(\mathbf{r}-\mathbf{r'}) \quad \text{(a)} \qquad \langle\,\mathbf{p}\,|\,\mathbf{p'}\,\rangle = \delta(\mathbf{p}-\mathbf{p'}) \quad \text{(c)}$$
$$\int \mathrm{d}^3r\;|\,\mathbf{r}\,\rangle\,\langle\,\mathbf{r}\,| = \mathbb{1} \quad \text{(b)} \qquad \int \mathrm{d}^3p\;|\,\mathbf{p}\,\rangle\,\langle\,\mathbf{p}\,| = \mathbb{1} \quad \text{(d)} \tag{E-5}$$

Of course, $\mathbf{r}$ and $\mathbf{p}$ are still considered to represent two sets of *continuous indices*, $\{\,x,\,y,\,z\,\}$ and $\{\,p_x,\,p_y,\,p_z\,\}$, which fix the basis kets of the $\{\,|\,\mathbf{r}\,\rangle\,\}$ and $\{\,|\,\mathbf{p}\,\rangle\,\}$ representations respectively.

Now let $\{\,u_i(\mathbf{r})\,\}$ be an orthonormal basis of $\mathscr{F}$. With each $u_i(\mathbf{r})$ is associated a ket $|\,u_i\,\rangle$ of $\mathscr{E}_\mathbf{r}$. The set $\{\,|\,u_i\,\rangle\,\}$ forms an orthonormal basis in $\mathscr{E}_\mathbf{r}$; it therefore satisfies the closure relation:

$$\sum_i |\,u_i\,\rangle\,\langle\,u_i\,| = \mathbb{1} \tag{E-11}$$

Evaluate the matrix element of both sides of (E-11) between $|\,\mathbf{r}\,\rangle$ and $|\,\mathbf{r'}\,\rangle$:

$$\sum_i \langle\,\mathbf{r}\,|\,u_i\,\rangle\,\langle\,u_i\,|\,\mathbf{r'}\,\rangle = \langle\,\mathbf{r}\,|\,\mathbb{1}\,|\,\mathbf{r'}\,\rangle = \langle\,\mathbf{r}\,|\,\mathbf{r'}\,\rangle \tag{E-12}$$

According to (E-8-a) and (E-5-a), this relation can be written:

$$\sum_i u_i(\mathbf{r})\,u_i^*(\mathbf{r'}) = \delta(\mathbf{r}-\mathbf{r'}) \tag{E-13}$$

The closure relation for the $\{\,u_i(\mathbf{r})\,\}$ [formula (A-32)] is therefore simply the expression in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation of the vectorial closure relation (E-11).

E-1-d. The scalar product of two vectors

We have defined the scalar product of two kets of $\mathscr{E}_\mathbf{r}$ as being equal to that of the associated wave functions in $\mathscr{F}$ [equation (E-1)]. In light of the discussion in § E-1-c,

this definition appears simply as a special case of formula (C-21). (E-1) can, in fact, be derived by inserting the closure relation (E-5-b) between $\langle \varphi |$ and $| \psi \rangle$:

$$\langle \varphi | \psi \rangle = \int \mathrm{d}^3 r \, \langle \varphi | \mathbf{r} \rangle \, \langle \mathbf{r} | \psi \rangle \tag{E-14}$$

and by interpreting the components $\langle \mathbf{r} | \psi \rangle$ and $\langle \mathbf{r} | \varphi \rangle$) as in (E-8-a).

If we place ourselves in the $\{ | \mathbf{p} \rangle \}$ representation, a well-known property of the Fourier transform is demonstrated (Appendix I, § 2-c):

$$\begin{aligned} \langle \varphi | \psi \rangle &= \int \mathrm{d}^3 p \, \langle \varphi | \mathbf{p} \rangle \, \langle \mathbf{p} | \psi \rangle \\ &= \int \mathrm{d}^3 p \, \overline{\varphi}^*(\mathbf{p}) \, \overline{\psi}(\mathbf{p}) \end{aligned} \tag{E-15}$$

E-1-e. Changing from the $\{ | \mathbf{r} \rangle \}$ representation to the $\{ | \mathbf{p} \rangle \}$ representation

This is accomplished using the method indicated in § C-5, the only difference arising from the fact that we are dealing here with two continuous bases. Changing from one basis to the other brings in the numbers:

$$\langle \mathbf{r} | \mathbf{p} \rangle = \langle \mathbf{p} | \mathbf{r} \rangle^* = (2\pi\hbar)^{-3/2} \, \mathrm{e}^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}} \tag{E-16}$$

A given ket $| \psi \rangle$ is represented by $\langle \mathbf{r} | \psi \rangle = \psi(\mathbf{r})$ in the $\{ | \mathbf{r} \rangle \}$ representation and by $\langle \mathbf{p} | \psi \rangle = \overline{\psi}(\mathbf{p})$ in the $\{ | \mathbf{p} \rangle \}$ representation. We already know [relations (E-2-b) and (E-7-b)] that $\psi(\mathbf{r})$ and $\overline{\psi}(\mathbf{p})$ are related by a Fourier transform. This is indeed what the formulas for the representation change yield:

$$\langle \mathbf{r} | \psi \rangle = \int \mathrm{d}^3 p \, \langle \mathbf{r} | \mathbf{p} \rangle \, \langle \mathbf{p} | \psi \rangle$$

that is:

$$\psi(\mathbf{r}) = (2\pi\hbar)^{-3/2} \int \mathrm{d}^3 p \;\; \mathrm{e}^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}} \;\; \overline{\psi}(\mathbf{p}) \tag{E-17}$$

Inversely:

$$\langle \mathbf{p} | \psi \rangle = \int \mathrm{d}^3 r \, \langle \mathbf{p} | \mathbf{r} \rangle \, \langle \mathbf{r} | \psi \rangle$$

that is:

$$\overline{\psi}(\mathbf{p}) = (2\pi\hbar)^{-3/2} \int \mathrm{d}^3 r \;\; \mathrm{e}^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}} \;\; \psi(\mathbf{r}) \tag{E-18}$$

By applying the general formula (C-56), one can easily pass from the matrix elements $\langle \mathbf{r}' | A | \mathbf{r} \rangle = A(\mathbf{r}', \mathbf{r})$ of an operator A in the $\{ | \mathbf{r} \rangle \}$ representation to the matrix elements $\langle \mathbf{p}' | A | \mathbf{p} \rangle = A(\mathbf{p}', \mathbf{p})$ of the same operator in the $\{ | \mathbf{p} \rangle \}$ representation:

$$A(\mathbf{p}', \mathbf{p}) = (2\pi\hbar)^{-3} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \;\; \mathrm{e}^{\frac{i}{\hbar} (\mathbf{p} \cdot \mathbf{r} - \mathbf{p}' \cdot \mathbf{r}')} \;\; A(\mathbf{r}', \mathbf{r}) \tag{E-19}$$

An analogous formula enables one to calculate $A(\mathbf{r}', \mathbf{r})$ from $A(\mathbf{p}', \mathbf{p})$.

E-2. The R and P operators

E-2-a. Definition

Let $|\,\psi\,\rangle$ be an arbitrary ket of $\mathscr{E}_{\mathbf{r}}$ and let $\langle\,\mathbf{r}\,|\,\psi\,\rangle = \psi(\mathbf{r}) \equiv \psi(x,y,z)$ be the corresponding wave function. Using the definition of the operator X, the ket:

$$|\,\psi'\,\rangle = X\,|\,\psi\,\rangle \tag{E-20}$$

is represented, in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ basis, by the function $\langle\,\mathbf{r}\,|\,\psi'\,\rangle = \psi'(\mathbf{r}) \equiv \psi'(x,y,z)$ such that:

$$\psi'(x,y,z) = x\ \psi(x,y,z) \tag{E-21}$$

In the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation, the X operator therefore coincides with the operator which multiplies by x. Although we characterize X by the way in which it transforms the wave functions, it is an operator which acts in the state space $\mathscr{E}_{\mathbf{r}}$. We can introduce two other operators, Y and Z, in an analogous manner. Thus we define X, Y and Z by the formulas:

$$\langle\,\mathbf{r}\,|\,X\,|\,\psi\,\rangle = x\,\langle\,\mathbf{r}\,|\,\psi\,\rangle \tag{E-22-a}$$
$$\langle\,\mathbf{r}\,|\,Y\,|\,\psi\,\rangle = y\,\langle\,\mathbf{r}\,|\,\psi\,\rangle \tag{E-22-b}$$
$$\langle\,\mathbf{r}\,|\,Z\,|\,\psi\,\rangle = z\,\langle\,\mathbf{r}\,|\,\psi\,\rangle \tag{E-22-c}$$

where the numbers x, y, z are precisely the three indices which label the ket $|\,\mathbf{r}\,\rangle$. X, Y and Z will be considered to be the "components" of a "vector operator" $\mathbf{R}$: for the moment, we shall treat this simply as a condensed notation, suggested by the fact that x, y, z are the components of the ordinary vector $\mathbf{r}$.

Manipulation of the X, Y, Z operators is particularly simple in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation. For example, in order to calculate the matrix element $\langle\,\varphi\,|\,X\,|\,\psi\,\rangle$, all we need to do is insert the closure relation (E-5-b) between $\langle\,\varphi\,|$ and X and use definition (E-22-a):

$$\begin{aligned}\langle\,\varphi\,|\,X\,|\,\psi\,\rangle &= \int \mathrm{d}^3r\,\langle\,\varphi\,|\,\mathbf{r}\,\rangle\,\langle\,\mathbf{r}\,|\,X\,|\,\psi\,\rangle \\ &= \int \mathrm{d}^3r\,\varphi^*(\mathbf{r})\,x\,\psi(\mathbf{r})\end{aligned} \tag{E-23}$$

Similarly, we define the vector operator $\mathbf{P}$ by its components P_x, P_y, P_z, whose action, in the $\{\,|\,\mathbf{p}\,\rangle\,\}$ representation, is given by:

$$\langle\,\mathbf{p}\,|\,P_x\,|\,\psi\,\rangle = p_x\,\langle\,\mathbf{p}\,|\,\psi\,\rangle \tag{E-24-a}$$
$$\langle\,\mathbf{p}\,|\,P_y\,|\,\psi\,\rangle = p_y\,\langle\,\mathbf{p}\,|\,\psi\,\rangle \tag{E-24-b}$$
$$\langle\,\mathbf{p}\,|\,P_z\,|\,\psi\,\rangle = p_z\,\langle\,\mathbf{p}\,|\,\psi\,\rangle \tag{E-24-c}$$

where p_x, p_y, p_z are the three indices which appear in the ket $|\,\mathbf{p}\,\rangle$.

Let us ascertain how the $\mathbf{P}$ operator acts in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation. To do so (*cf.* § C-5-d), we use the closure relation (E-5-d) and the transformation matrix (E-16)

to obtain:

$$\begin{aligned}\langle \mathbf{r} \,|\, P_x \,|\, \psi \rangle &= \int \mathrm{d}^3p \, \langle \mathbf{r} \,|\, \mathbf{p} \rangle \, \langle \mathbf{p} \,|\, P_x \,|\, \psi \rangle \\ &= (2\pi\hbar)^{-3/2} \int \mathrm{d}^3p \;\; \mathrm{e}^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} \; p_x \; \bar{\psi}(\mathbf{p}) \end{aligned} \tag{E-25}$$

We recognize in (E-25) the Fourier transform of $p_x \, \overline{\psi}(\mathbf{p})$, that is $\dfrac{\hbar}{i}\dfrac{\partial}{\partial x}\psi(\mathbf{r})$ [Appendix I, relation (38a)]. Therefore:

$$\langle \mathbf{r} \,|\, \mathbf{P} \,|\, \psi \rangle = \frac{\hbar}{i} \boldsymbol{\nabla} \langle \mathbf{r} \,|\, \psi \rangle \tag{E-26}$$

In the $\{\, |\, \mathbf{r} \rangle \,\}$ representation, the $\mathbf{P}$ operator coincides with the differential operator $(\hbar/i)\boldsymbol{\nabla}$ applied to the wave functions. The calculation of a matrix element such as $\langle \varphi \,|\, P_x \,|\, \psi \rangle$ in the $\{\, |\, \mathbf{r} \rangle \,\}$ representation is therefore performed in the following manner:

$$\begin{aligned}\langle \varphi \,|\, P_x \,|\, \psi \rangle &= \int \mathrm{d}^3r \, \langle \varphi \,|\, \mathbf{r} \rangle \, \langle \mathbf{r} \,|\, P_x \,|\, \psi \rangle \\ &= \int \mathrm{d}^3r \, \varphi^*(\mathbf{r}) \left[\frac{\hbar}{i} \frac{\partial}{\partial x} \right] \psi(\mathbf{r}) \end{aligned} \tag{E-27}$$

Placing ourselves in the $\{\, |\, \mathbf{r} \rangle \,\}$ representation, we can also calculate the commutators between the X, Y, Z, P_x, P_y, P_z operators. For example:

$$\begin{aligned}\langle \mathbf{r} \,|\, [X, P_x] \,|\, \psi \rangle &= \langle \mathbf{r} \,|\, (XP_x - P_x X) \,|\, \psi \rangle \\ &= x \, \langle \mathbf{r} \,|\, P_x \,|\, \psi \rangle - \frac{\hbar}{i} \frac{\partial}{\partial x} \, \langle \mathbf{r} \,|\, X \,|\, \psi \rangle \\ &= \frac{\hbar}{i} \, x \, \frac{\partial}{\partial x} \langle \mathbf{r} \,|\, \psi \rangle - \frac{\hbar}{i} \frac{\partial}{\partial x} \, x \, \langle \mathbf{r} \,|\, \psi \rangle \\ &= i\hbar \, \langle \mathbf{r} \,|\, \psi \rangle \end{aligned} \tag{E-28}$$

This calculation is valid for all $|\, \psi \rangle$ and for any ket of the $|\, \mathbf{r} \rangle$ basis. Thus one finds[10]:

$$[X, \, P_x] = i\hbar \tag{E-29}$$

In the same way, we find all the other commutators between the components of $\mathbf{R}$ and those of $\mathbf{P}$. The result can be written in the form:

$$\boxed{\left.\begin{array}{l} [R_i, \, R_j] \, = \, 0 \\ [P_i, \, P_j] \, = \, 0 \\ [R_i, \, P_j] \, = \, i\hbar\delta_{ij} \end{array}\right\} \; i, j = 1, \, 2, \, 3} \tag{E-30}$$

where R_1, R_2, R_3, and P_1, P_2, P_3 designate respectively X, Y, Z and P_x, P_y, P_z. Formulas (E-30) are called *canonical commutation relations.*

[10] The commutator $[X, \, P_x]$ is an operator, and it should, actually, be written $[X, \, P_x] = i\hbar\mathbb{1}$. However, we shall often replace the identity operator $\mathbb{1}$ by the number 1, except when it is important to make the distinction.

E-2-b. $\boldsymbol{R}$ and $\boldsymbol{P}$ are Hermitian

In order to show that X, for example, is a Hermitian operator, we can use formula (E-23):

$$\begin{aligned}\langle\,\varphi\,|\,X\,|\,\psi\,\rangle &= \int \mathrm{d}^3r\ \varphi^*(\mathbf{r})\ x\ \psi(\mathbf{r}) = \left[\int \mathrm{d}^3r\ \varphi(\mathbf{r})\ x\ \psi^*(\mathbf{r})\right]^* \\ &= \langle\,\psi\,|\,X\,|\,\varphi\,\rangle^* \end{aligned} \tag{E-31}$$

From § B-4-e, we know that equation (E-31) is characteristic of a Hermitian operator.

Similar proofs show that Y and Z are also Hermitian. For P_x, P_y and P_z, the $\{\,|\,\mathbf{p}\,\rangle\,\}$ representation can be used, and the calculations are then analogous to the preceding ones.

It is interesting to show that $\mathbf{P}$ is Hermitian by using equation (E-26), which gives its action in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation. Consider, for example, formula (E-27) and integrate it by parts:

$$\begin{aligned}\langle\,\varphi\,|\,P_x\,|\,\psi\,\rangle &= \frac{\hbar}{i}\int \mathrm{d}y\ \mathrm{d}z\ \int_{-\infty}^{+\infty}\mathrm{d}x\ \varphi^*(\mathbf{r})\ \frac{\partial}{\partial x}\ \psi(\mathbf{r}) \\ &= \frac{\hbar}{i}\int \mathrm{d}y\ \mathrm{d}z\ \left\{\Big[\varphi^*(\mathbf{r})\ \psi(\mathbf{r})\Big]_{x=-\infty}^{x=+\infty} - \int_{-\infty}^{+\infty}\mathrm{d}x\ \psi(\mathbf{r})\ \frac{\partial}{\partial x}\ \varphi^*(\mathbf{r})\right\}\end{aligned} \tag{E-32}$$

Since the integral which yields the scalar product $\langle\,\varphi\,|\,\psi\,\rangle$ is convergent, $\varphi^*(\mathbf{r})\,\psi(\mathbf{r})$ approaches zero when $x \longrightarrow \pm\infty$. The first term on the right hand side of (E-32) is therefore equal to zero, and:

$$\begin{aligned}\langle\,\varphi\,|\,P_x\,|\,\psi\,\rangle &= -\frac{\hbar}{i}\int \mathrm{d}^3r\ \ \psi(\mathbf{r})\ \frac{\partial}{\partial x}\,\varphi^*(\mathbf{r}) \\ &= \left[\frac{\hbar}{i}\int \mathrm{d}^3r\ \ \psi^*(\mathbf{r})\ \frac{\partial}{\partial x}\,\varphi(\mathbf{r})\right]^* \\ &= \langle\,\psi\,|\,P_x\,|\,\varphi\,\rangle^*\end{aligned} \tag{E-33}$$

It can be seen that the presence of the imaginary number i is essential. The differential operator $\dfrac{\partial}{\partial x}$, acting on the functions of $\mathscr{F}$, is *not* Hermitian, because of the sign change which is introduced by the integration by parts. However, $i\dfrac{\partial}{\partial x}$ is Hermitian, as is $\dfrac{\hbar}{i}\dfrac{\partial}{\partial x}$.

E-2-c. Eigenvectors of $\boldsymbol{R}$ and $\boldsymbol{P}$

Consider the action of the X operator on the ket $|\,\mathbf{r}_0\,\rangle$; according to (E-22-a), we have:

$$\langle\,\mathbf{r}\,|\,X\,|\,\mathbf{r}_0\,\rangle = x\,\langle\,\mathbf{r}\,|\,\mathbf{r}_0\,\rangle = x\,\delta(\mathbf{r}-\mathbf{r}_0) = x_0\,\delta(\mathbf{r}-\mathbf{r}_0) = x_0\,\langle\,\mathbf{r}\,|\,\mathbf{r}_0\,\rangle \tag{E-34}$$

This equation expresses the fact that the components, in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation, of the ket $X\,|\,\mathbf{r}_0\,\rangle$ are equal to those of the ket $|\,\mathbf{r}_0\,\rangle$ multiplied by x_0. We therefore have:

$$X\,|\,\mathbf{r}_0\,\rangle = x_0\,|\,\mathbf{r}_0\,\rangle \tag{E-35}$$

An analogous argument shows that the kets $|\,\mathbf{r}_0\,\rangle$ are also eigenvectors of the Y and Z operators. Omitting the index zero which then becomes unnecessary, we can write:

$$\boxed{\begin{aligned} X\,|\,\mathbf{r}\,\rangle &= x\,|\,\mathbf{r}\,\rangle \\ Y\,|\,\mathbf{r}\,\rangle &= y\,|\,\mathbf{r}\,\rangle \\ Z\,|\,\mathbf{r}\,\rangle &= z\,|\,\mathbf{r}\,\rangle \end{aligned}} \tag{E-36}$$

The kets $|\,\mathbf{r}\,\rangle$ are therefore the eigenkets common to X, Y and Z. Thus the notation $|\,\mathbf{r}\,\rangle$ which we chose above is justified: each eigenvector is labelled by a vector $\mathbf{r}$, whose components x, y, z represent three continuous indices which correspond to the eigenvalues of X, Y, Z.

Similar arguments can be elaborated for the $\mathbf{P}$ operator, placing ourselves, this time, in the representation $\{\,|\,\mathbf{p}\,\rangle\,\}$. We then obtain:

$$\boxed{\begin{aligned} P_x\,|\,\mathbf{p}\,\rangle &= p_x\,|\,\mathbf{p}\,\rangle \\ P_y\,|\,\mathbf{p}\,\rangle &= p_y\,|\,\mathbf{p}\,\rangle \\ P_z\,|\,\mathbf{p}\,\rangle &= p_z\,|\,\mathbf{p}\,\rangle \end{aligned}} \tag{E-37}$$

Comment:

This result can also be derived from equation (E-26), which gives the action of $\mathbf{P}$ in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation. Using (E-9), we find:

$$\begin{aligned} \langle\,\mathbf{r}\,|\,P_x\,|\,\mathbf{p}\,\rangle &= \frac{\hbar}{i}\,\frac{\partial}{\partial x}\,\langle\,\mathbf{r}\,|\,\mathbf{p}\,\rangle = \frac{\hbar}{i}\,\frac{\partial}{\partial x}\,(2\pi\hbar)^{-3/2}\,\mathrm{e}^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} \\ &= p_x\,(2\pi\hbar)^{-3/2}\,\mathrm{e}^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} = p_x\,\langle\,\mathbf{r}\,|\,\mathbf{p}\,\rangle \end{aligned} \tag{E-38}$$

All the components of the ket $P_x\,|\,\mathbf{p}\,\rangle$ in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation can be obtained by multiplying those of $|\,\mathbf{p}\,\rangle$ by the constant p_x: $|\,\mathbf{p}\,\rangle$ is an eigenket of P_x with the eigenvalue p_x.

E-2-d. *R* and *P* are observables

Relations (E-5-b) and (E-5-d) express the fact that the $\{\,|\,\mathbf{r}\,\rangle\,\}$ vectors and the $\{\,|\,\mathbf{p}\,\rangle\,\}$ vectors constitute bases in $\mathscr{E}_{\mathbf{r}}$. Therefore, $\mathbf{R}$ and $\mathbf{P}$ are observables.

Moreover, the specification of the three eigenvalues x_0, y_0, z_0 of X, Y, Z uniquely determines the corresponding eigenvector $|\,\mathbf{r}_0\,\rangle$: in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation, its coordinates are $\delta(x-x_0)\delta(y-y_0)\delta(z-z_0)$. The set of the three operators X, Y, Z therefore constitutes a C.S.C.O. in $\mathscr{E}_{\mathbf{r}}$.

It can be shown in the same way that the three components P_x, P_y, P_z of $\mathbf{P}$ also constitute a C.S.C.O. in $\mathscr{E}_{\mathbf{r}}$.

Note that, in $\mathscr{E}_{\mathbf{r}}$, X does not constitute a C.S.C.O. by itself. When the x_0 index is fixed, y_0 and z_0 can take on any real values. Thus, each eigenvalue x_0 is infinitely degenerate. On the other hand, in the state space $\mathscr{E}_x$ of a one-dimensional problem, X constitutes a C.S.C.O.: the eigenvalue x_0 uniquely determines the corresponding eigenket $|\,x_0\,\rangle$, its coordinates being $\delta(x-x_0)$ in the $\{\,|\,x\,\rangle\,\}$ representation.

Comment:

We have found two C.S.C.O.'s in $\mathscr{E}_{\mathbf{r}}$, $\{ X, Y, Z \}$ and $\{ P_x, P_y, P_z \}$. We shall encounter others later. Consider, for example, the set $\{ X, P_y, P_z \}$: these three observables commute (equations (E-30)); moreover, if the three eigenvalues x_0, p_{0_y} and p_{0_z} are fixed, there corresponds to them only one ket, whose associated wave function is written:

$$\psi_{x_0, p_{0_y}, p_{0_z}}(x, y, z) = \delta(x - x_0) \, \frac{1}{2\pi\hbar} \, \mathrm{e}^{\frac{i}{\hbar}(p_{0_y} y + p_{0_z} z)} \tag{E-39}$$

F. Tensor product of state spaces[11]

F-1. Introduction

We introduced the state space of a physical system using the concept of a one-particle wave function. However, our reasoning has involved sometimes one- and sometimes three-dimensional wave functions. Now it is clear that the space of square-integrable functions is not the same for functions of one variable $\psi(x)$ as for functions of three variables $\psi(\mathbf{r})$: $\mathscr{E}_{\mathbf{r}}$ and $\mathscr{E}_x$ are therefore different spaces. Nevertheless, $\mathscr{E}_{\mathbf{r}}$ appears to be essentially a generalization of $\mathscr{E}_x$. Does there exist a more precise relation between these two spaces?

In this section, we are going to define and study the operation of taking the tensor product of vector spaces[12], and apply it to state spaces. This will answer, in particular, the question we have just asked: $\mathscr{E}_{\mathbf{r}}$ can be constructed from $\mathscr{E}_x$ and two other spaces, $\mathscr{E}_y$ and $\mathscr{E}_z$, which are isomorphic to it (§ F-4-a below).

In the same way, we shall be concerned later (Chapters IV and IX) with the existence, for certain particles, of an intrinsic angular momentum or spin. In addition to the external degrees of freedom (position, momentum), which are treated using the observables $\mathbf{R}$ and $\mathbf{P}$ defined in $\mathscr{E}_{\mathbf{r}}$, it will be necessary to take into account the internal degrees of freedom and to introduce spin observables which act in a spin state space $\mathscr{E}_s$. The state space $\mathscr{E}$ of a particle with spin will then be seen to be the tensor product of $\mathscr{E}_{\mathbf{r}}$ and $\mathscr{E}_s$.

Finally, the concept of a tensor product of state spaces allows us to solve the following problem. Let (S_1) and (S_2) be two isolated physical systems (they are, for example, sufficiently far apart that their interactions are perfectly negligible). The state spaces which correspond to (S_1) and (S_2) are, respectively, $\mathscr{E}_1$ and $\mathscr{E}_2$. Now let us assume that we consider the set of these two systems to form one physical system (S) (this becomes indispensable when they are close enough to interact). What is then the state space $\mathscr{E}$ of the global system (S)?

It can be seen from these examples how useful the definitions and results of this section are in quantum mechanics.

F-2. Definition and properties of the tensor product

Let $\mathscr{E}_1$ and $\mathscr{E}_2$ be two[13] spaces, of dimension N_1 and N_2 respectively (N_1 and N_2 can be finite or infinite). Vectors and operators of these spaces will be assigned an index,

[11] This section is not necessary for the understanding of Chapter III. One can study it later when it becomes necessary to use tensor products (Complement $\mathrm{D_{IV}}$, or Chapter IX).

[12] This operation is sometimes called the "Kronecker product"

[13] The following definitions can casily be extended to the tensor product of a finite number of spaces.

(1) or (2), depending on whether they belong to $\mathscr{E}_1$ or $\mathscr{E}_2$.

F-2-a. Tensor product space $\mathscr{E}$

α. *Definition*

By definition, the vector space $\mathscr{E}$ is called the *tensor product of* $\mathscr{E}_1$ *and* $\mathscr{E}_2$:

$$\mathscr{E} = \mathscr{E}_1 \otimes \mathscr{E}_2 \tag{F-1}$$

if there is associated with each pair of vectors, $|\,\varphi(1)\,\rangle$ belonging to $\mathscr{E}_1$ and $|\,\chi(2)\,\rangle$ belonging to $\mathscr{E}_2$, a vector of $\mathscr{E}$, denoted by[14]:

$$|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle \tag{F-2}$$

which is called the tensor product of $|\,\varphi(1)\,\rangle$ and $|\,\chi(2)\,\rangle$, this correspondence satisfying the following conditions:

(*i*) It is *linear* with respect to multiplication by complex numbers:

$$\begin{aligned} \big[\lambda\,|\,\varphi(1)\,\rangle\big] \otimes |\,\chi(2)\,\rangle &= \lambda\big[\,|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle\,\big] \\ |\,\varphi(1)\,\rangle \otimes \big[\mu\,|\,\chi(2)\,\rangle\big] &= \mu\big[\,|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle\,\big] \end{aligned} \tag{F-3}$$

(*ii*) It is *distributive* with respect to vector addition:

$$\begin{aligned} |\,\varphi(1)\,\rangle \otimes \big[|\,\chi_1(2)\,\rangle + |\,\chi_2(2)\,\rangle\big] &= |\,\varphi(1)\,\rangle \otimes |\,\chi_1(2)\,\rangle + |\,\varphi(1)\,\rangle \otimes |\,\chi_2(2)\,\rangle \\ \big[|\,\varphi_1(1)\,\rangle + |\,\varphi_2(1)\,\rangle\big] \otimes |\,\chi(2)\,\rangle &= |\,\varphi_1(1)\,\rangle \otimes |\,\chi(2)\,\rangle + |\,\varphi_2(1)\,\rangle \otimes |\,\chi(2)\,\rangle \end{aligned} \tag{F-4}$$

(*iii*) When a basis has been chosen in each of the spaces $\mathscr{E}_1$ and $\mathscr{E}_2$, $\{\,|\,u_i(1)\,\rangle\,\}$ for $\mathscr{E}_1$ and $\{\,|\,v_l(2)\,\rangle\,\}$ for $\mathscr{E}_2$, the set of vectors $|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle$ constitutes a basis in $\mathscr{E}$. If N_1 and N_2 are finite, the dimension of $\mathscr{E}$ is consequently $N_1 N_2$.

β. *Vectors of* $\mathscr{E}$

(*i*) Let us first consider a *tensor product vector*, $|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle$. Whatever $|\,\varphi(1)\,\rangle$ and $|\,\chi(2)\,\rangle$ may be, they can be expressed in the $\{\,|\,u_i(1)\,\rangle\,\}$ and $\{\,|\,v_l(2)\,\rangle\,\}$ bases respectively:

$$\begin{aligned} |\,\varphi(1)\,\rangle &= \sum_i a_i\,|\,u_i(1)\,\rangle \\ |\,\chi(2)\,\rangle &= \sum_l b_l\,|\,v_l(2)\,\rangle \end{aligned} \tag{F-5}$$

Using the properties described in § F-2-a-α, the expansion of the vector $|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle$ in the $\{\,|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle\,\}$ basis can be written:

$$|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle = \sum_{i,l} a_i b_l\,|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle \tag{F-6}$$

[14] This vector can be written either $|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle$ or $|\,\chi(2)\,\rangle \otimes |\,\varphi(1)\,\rangle$; the order of the two vectors is of no importance.

Therefore, *the components of a tensor product vector are the products of the components of the two vectors of the product.*

(*ii*) *There exist in* $\mathscr{E}$ *vectors which are not tensor products* of a vector of $\mathscr{E}_1$ by a vector of $\mathscr{E}_2$. Since $\{\,|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle\,\}$ constitutes by hypothesis a basis in $\mathscr{E}$, the most general vector of $\mathscr{E}$ is expressed by:

$$|\,\psi\,\rangle = \sum_{i,l} c_{i,l}\,|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle \tag{F-7}$$

Given $N_1 N_2$ arbitrary complex numbers $c_{i,l}$, it is not always possible to put them in the form of products, $a_i b_l$, of N_1 numbers a_i and N_2 numbers b_l. Therefore, in general, vectors $|\,\varphi(1)\,\rangle$ and $|\,\chi(2)\,\rangle$ of which $|\,\psi\,\rangle$ is the tensor product do not exist. However, *an arbitrary vector of* $\mathscr{E}$ *can always be decomposed into a linear combination of tensor product vectors*, as is shown by formula (F-7).

γ. The scalar product in $\mathscr{E}$

The existence of scalar products in $\mathscr{E}_1$ and $\mathscr{E}_2$ permits us to define one in $\mathscr{E}$ as well. We first define the scalar product of $|\,\varphi(1)\,\chi(2)\,\rangle = |\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle$ by $|\,\varphi'(1)\,\chi'(2)\,\rangle = |\,\varphi'(1)\,\rangle \otimes |\,\chi'(2)\,\rangle$ by setting:

$$\langle\,\varphi'(1)\,\chi'(2)\,|\,\varphi(1)\,\chi(2)\,\rangle = \langle\,\varphi'(1)\,|\,\varphi(1)\,\rangle\,\langle\,\chi'(2)\,|\,\chi(2)\,\rangle \tag{F-8}$$

For two arbitrary vectors of $\mathscr{E}$, we simply use the fundamental properties of the scalar product [equations (B-9), (B-10) and (B-11)], since each of these vectors is a linear combination of tensor product vectors.

Notice, in particular, that the basis $\{\,|\,u_i(1)\,v_l(2)\,\rangle = |\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle\,\}$ is orthonormal if each of the bases $\{\,|\,u_i(1)\,\rangle\,\}$ and $\{\,|\,v_l(2)\,\rangle\,\}$ is:

$$\begin{aligned}\langle\,u_{i'}(1)\,v_{l'}(2)\,|\,u_i(1)\,v_l(2)\,\rangle &= \langle\,u_{i'}(1)\,|\,u_i(1)\,\rangle\,\langle\,v_{l'}(2)\,|\,v_l(2)\,\rangle \\ &= \delta_{ii'}\,\delta_{ll'}\end{aligned} \tag{F-9}$$

F-2-b. Tensor product of operators

(*i*) First, consider a linear operator $A(1)$ defined in $\mathscr{E}_1$. We associate with it a linear operator $\widetilde{A}(1)$ acting in $\mathscr{E}$, which we call the *extension of* $A(1)$ *in* $\mathscr{E}$, and which is characterized in the following way: when $\widetilde{A}(1)$ is applied to a tensor product vector $|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle$, one obtains, by definition:

$$\widetilde{A}(1)\big[|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle\big] = \big[A(1)|\,\varphi(1)\,\rangle\big] \otimes |\,\chi(2)\,\rangle \tag{F-10}$$

The hypothesis that $\widetilde{A}(1)$ is linear is then sufficient for determining it completely. An arbitrary vector $|\,\psi\,\rangle$ of $\mathscr{E}$ can be written in the form (F-7). Definition (F-10) then gives the action of $\widetilde{A}(1)$ on $|\,\psi\,\rangle$:

$$\widetilde{A}(1)|\,\psi\,\rangle = \sum_{i,l} c_{i,l}\big[A(1)\,|\,u_i(1)\,\rangle\big] \otimes |\,v_l(2)\,\rangle \tag{F-11}$$

We obtain in an analogous manner the extension $\widetilde{B}(2)$ of an operator $B(2)$ initially defined in $\mathscr{E}_2$.

(*ii*) Now let $A(1)$ and $B(2)$ be two linear operators acting respectively in $\mathscr{E}_1$ and $\mathscr{E}_2$. Their *tensor product* $A(1) \otimes B(2)$ is the linear operator in $\mathscr{E}$, defined by the following relation which describes its action on the tensor product vectors:

$$\big[A(1) \otimes B(2)\big]\,\big[\,|\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle\,\big] = \big[A(1)\,|\,\varphi(1)\,\rangle\,\big] \otimes \big[B(2)|\,\chi(2)\,\rangle\,\big] \tag{F-12}$$

Here also, this definition is sufficient for characterizing $A(1) \otimes B(2)$.

Comments:

(*i*) The extensions of operators are special cases of tensor products: if $\mathbb{1}(1)$ and $\mathbb{1}(2)$ are the identity operators in $\mathscr{E}_1$ and $\mathscr{E}_2$ respectively, $\widetilde{A}(1)$ and $\widetilde{B}(2)$ can be written:

$$\begin{aligned}\widetilde{A}(1) &= A(1) \otimes \mathbb{1}(2)\\ \widetilde{B}(2) &= \mathbb{1}(1) \otimes B(2)\end{aligned} \tag{F-13}$$

Inversely, the tensor product $A(1) \otimes B(2)$ coincides with the ordinary product of two operators $\widetilde{A}(1)$ and $\widetilde{B}(2)$ of $\mathscr{E}$:

$$A(1) \otimes B(2) = \widetilde{A}(1)\widetilde{B}(2) \tag{F-14}$$

(*ii*) It is easy to show that two operators such as $\widetilde{A}(1)$ and $\widetilde{B}(2)$ commute in $\mathscr{E}$:

$$\big[\widetilde{A}(1), \widetilde{B}(2)\big] = 0 \tag{F-15}$$

We must verify that $\widetilde{A}(1)\widetilde{B}(2)$ and $\widetilde{B}(2)\widetilde{A}(1)$ yield the same result when they act on an arbitrary vector of the $\{\,|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle\,\}$ basis:

$$\begin{aligned}\widetilde{A}(1)\widetilde{B}(2)\,|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle &= \widetilde{A}(1)\Big[|\,u_i(1)\,\rangle \otimes \big[B(2)\,|\,v_l(2)\,\rangle\big]\Big]\\ &= \big[A(1)\,|\,u_i(1)\,\rangle\big] \otimes \big[B(2)\,|\,v_l(2)\,\rangle\big] \qquad \text{(F-16)}\\ \widetilde{B}(2)\widetilde{A}(1)\,|\,u_i(1)\,\rangle \otimes |\,v_l(2)\,\rangle &= \widetilde{B}(2)\Big[\big[A(1)\,|\,u_i(1)\,\rangle\big] \otimes |\,v_l(2)\,\rangle\Big]\\ &= \big[A(1)\,|\,u_i(1)\,\rangle\big] \otimes \big[B(2)\,|\,v_l(2)\,\rangle\big] \qquad \text{(F-17)}\end{aligned}$$

(*iii*) The projector onto the tensor product vector $|\,\varphi(1)\,\chi(2)\,\rangle = |\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle$, which is an operator acting in $\mathscr{E}$, is obtained by taking the tensor product of the projectors onto $|\,\varphi(1)\,\rangle$ and $|\,\chi(2)\,\rangle$:

$$|\,\varphi(1)\,\chi(2)\,\rangle\,\langle\,\varphi(1)\,\chi(2)\,| = |\,\varphi(1)\,\rangle\,\langle\,\varphi(1)\,| \otimes |\,\chi(2)\,\rangle\,\langle\,\chi(2)\,| \tag{F-18}$$

This relation follows immediately from the definition of the scalar product in $\mathscr{E}$.

(*iv*) Just as with vectors, there exist operators in $\mathscr{E}$ which are not tensor products of an operator of $\mathscr{E}_1$ and an operator of $\mathscr{E}_2$.

F-2-c. Notations

In quantum mechanics, the notation generally used is a simplified version of the one which we have defined here. This is the one we shall adopt, but it is important to interpret it correctly in the light of the preceding discussion.

First of all, the symbol $\otimes$ which indicates the tensor product is omitted, and the vectors or operators which are to be multiplied tensorially are simply juxtaposed:

$$|\,\varphi(1)\,\rangle|\,\chi(2)\,\rangle \quad \text{means} \quad |\,\varphi(1)\,\rangle \otimes |\,\chi(2)\,\rangle \tag{F-19}$$

$$A(1)B(2) \quad \text{means} \quad A(1) \otimes B(2) \tag{F-20}$$

Moreover, the extension in $\mathscr{E}$ of an operator of $\mathscr{E}_1$ or $\mathscr{E}_2$ is written in the same way as this operator itself:

$$A(1) \quad \text{means} \quad \widetilde{A}(1) \quad \text{or} \quad A(1) \tag{F-21}$$

No confusion is possible in (F-19): until now we have never written two kets one after the other as we do here. Notice in particular that the expression $|\,\psi\,\rangle\,|\,\varphi\,\rangle$, where $|\,\psi\,\rangle$ and $|\,\varphi\,\rangle$ belong to the same space $\mathscr{E}$, is not defined in this space: it represents a vector of the space which is the tensor product of $\mathscr{E}$ by itself.

On the other hand, the notation in (F-20) and (F-21) is slightly ambiguous, especially in the latter, where two different operators are represented by the same symbol. However, it will be possible in practice to distinguish between them by the vector to which this symbol is applied: depending on whether it is a vector of $\mathscr{E}$ or of $\mathscr{E}_1$, we shall be dealing with $\widetilde{A}(1)$, or with $A(1)$ in a strict sense. As for formula (F-20), it poses no problem when $\mathscr{E}_1$ and $\mathscr{E}_2$ are different, since we have, until now, defined only products of operators which act in the same space. Moreover, $A(1)B(2)$ can be considered to be an ordinary product of operators of $\mathscr{E}$, if $A(1)$ and $B(2)$ are interpreted as designating, in fact, $\widetilde{A}(1)$ and $\widetilde{B}(2)$ [equation (F-14)].

F-3. Eigenvalue equations in the product space

The vectors of $\mathscr{E}$ which are tensor products of a vector of $\mathscr{E}_1$ and a vector of $\mathscr{E}_2$ play an important role in the discussion above. We shall see that this is also the case for the extensions to $\mathscr{E}$ of operators acting in $\mathscr{E}_1$ and $\mathscr{E}_2$.

F-3-a. Eigenvalues and eigenvectors of extended operators

α. *Eigenvalue equation of* $A(1)$

Consider an operator $A(1)$, for which we know, in $\mathscr{E}_1$, all the eigenstates and eigenvalues. We shall assume, for example, that the whole spectrum of $A(1)$ is discrete:

$$A(1)\,|\,\varphi_n^i(1)\,\rangle = a_n\,|\,\varphi_n^i(1)\,\rangle \quad ; \quad i = 1,\, 2,\, ...,\, g_n \tag{F-22}$$

We want to solve the eigenvalue equation of the extension of $A(1)$ in $\mathscr{E}$:

$$A(1)\,|\,\psi\,\rangle = \lambda\,|\,\psi\,\rangle \quad ; \quad |\,\psi\,\rangle \in \mathscr{E} \tag{F-23}$$

It can immediately be seen, from (F-10), that every vector of the form $|\,\varphi_n^i(1)\,\rangle\,|\,\chi(2)\,\rangle$ is an eigenvector of $A(1)$ with the eigenvalue a_n, whatever $|\,\chi(2)\,\rangle$ may be, since:

$$\begin{aligned} A(1)\,|\,\varphi_n^i(1)\,\rangle\,|\,\chi(2)\,\rangle &= \left[A(1)\,|\,\varphi_n^i(1)\,\rangle\right]|\,\chi(2)\,\rangle \\ &= a_n\,|\,\varphi_n^i(1)\,\rangle\,|\,\chi(2)\,\rangle \end{aligned} \tag{F-24}$$

Let us show that, when $A(1)$ is an observable in $\mathscr{E}_1$, all the solutions of (F-23) can be obtained in this way. The set of the $|\,\varphi_n^i(1)\,\rangle$ then forms a basis in $\mathscr{E}_1$. Consequently, the orthonormal system of vectors $|\,\psi_n^{i,l}\,\rangle$ such that:

$$|\,\psi_n^{i,l}\,\rangle = |\,\varphi_n^i(1)\,\rangle\,|\,v_l(2)\,\rangle \tag{F-25}$$

(where $\{\,|\,v_l(2)\,\rangle\,\}$ is a basis of $\mathscr{E}_2$) forms a basis in $\mathscr{E}$. We therefore have an orthonormal basis constituted by the eigenvectors of $A(1)$ in $\mathscr{E}$, $\{\,|\,\psi_n^{i,l}\,\rangle\,\}$, which means that equation (F-23) is solved.

The following conclusions can be drawn:

– *If $A(1)$ is an observable in $\mathscr{E}_1$, it is also an observable in $\mathscr{E}$.* This results from the fact that the extension of $A(1)$ is Hermitian and from the fact that $\{\,|\,\psi_n^{i,l}\,\rangle\,\}$ constitutes a basis in $\mathscr{E}$.

– *The spectrum of $A(1)$ is the same in $\mathscr{E}$ as in $\mathscr{E}_1$*: the same eigenvalues a_n appear in (F-22) and in (F-24).

– Nevertheless, an eigenvalue a_n which is g_n-fold degenerate in $\mathscr{E}_1$ has, in $\mathscr{E}$, a degree of degeneracy $N_2 \times g_n$. We known that the eigensubspace associated with a_n is spanned in $\mathscr{E}$ by the kets $|\,\psi_n^{i,l}\,\rangle = |\,\varphi_n^i(1)\,\rangle\,|\,v_l(2)\,\rangle$ with n fixed and $i = 1, 2, ..., g_n$; $l = 1, 2, ..., N_2$. Therefore, even if a_n is simple in $\mathscr{E}_1$, it is (N_2-fold) degenerate in $\mathscr{E}$.

The projector onto the eigensubspace corresponding to an eigenvalue a_n is written, in $\mathscr{E}$ [*cf.* (F-18)]:

$$\begin{aligned}\sum_{i,l} |\,\psi_n^{i,l}\,\rangle\,\langle\,\psi_n^{i,l}\,| &= \sum_{i,l} |\,\varphi_n^i(1)\,\rangle\,\langle\,\varphi_n^i(1)\,| \otimes |\,v_l(2)\,\rangle\,\langle\,v_l(2)\,| \\ &= \sum_i |\,\varphi_n^i(1)\,\rangle\,\langle\,\varphi_n^i(1)\,| \otimes \mathbb{1}(2)\end{aligned} \tag{F-26}$$

using in $\mathscr{E}_2$ the closure relation relative to the $\{\,|\,v_l(2)\,\rangle\,\}$ basis. It is therefore the extension of the projector $P_n(1) = \sum_i |\,\varphi_n^i(1)\,\rangle\,\langle\,\varphi_n^i(1)\,|$ which is associated with a_n in $\mathscr{E}_1$.

β. *Eigenvalue equation of $A(1) + B(2)$*

We shall often need to solve, in a tensor product space such as $\mathscr{E}$, eigenvalue equations for operators of the form:

$$C = A(1) + B(2) \tag{F-27}$$

where $A(1)$ and $B(2)$ are observables whose eigenvalues and eigenvectors are known in $\mathscr{E}_1$ and $\mathscr{E}_2$ respectively:

$$\begin{aligned}A(1)\,|\,\varphi_n(1)\,\rangle &= a_n\,|\,\varphi_n(1)\,\rangle \\ B(2)\,|\,\chi_p(2)\,\rangle &= b_p\,|\,\chi_p(2)\,\rangle\end{aligned} \tag{F-28}$$

[to simplify the notation, we assume the spectra of $A(1)$ and $B(2)$ to be discrete and non-degenerate in $\mathscr{E}_1$ and $\mathscr{E}_2$].

$A(1)$ and $B(2)$ commute [formulas (F-16) and (F-17)], and the $|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle$, which form a basis in $\mathscr{E}$, are eigenvectors common to $A(1)$ and $B(2)$:

$$\begin{aligned} A(1)\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle &= a_n\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle \\ B(2)\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle &= b_p\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle \end{aligned} \tag{F-29}$$

They are also eigenvectors of C:

$$C\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle = (a_n + b_p)\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle \tag{F-30}$$

This gives us directly the solution of the eigenvalue equation of C.

Therefore: *the eigenvalues of $C = A(1) + B(2)$ are the sums of an eigenvalue of $A(1)$ and an eigenvalue of $B(2)$. One can find a basis of eigenvectors of C which are tensor products of an eigenvector of $A(1)$ and an eigenvector of $B(2)$.*

Comment:

Equation (F-30) shows that the eigenvalues of C are all of the form $c_{np} = a_n + b_p$. If two different pairs of values of n and p which give the same value for c_{np} do not exist, c_{np} is non-degenerate (recall that we have assumed a_n and b_p to be non-degenerate in $\mathscr{E}_1$ and $\mathscr{E}_2$ respectively). The corresponding eigenvector of C is necessarily the tensor product $|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle$. If, on the other hand, the eigenvalue c_{np} is, for example, two-fold degenerate (there exist m and q such that $c_{mq} = c_{np}$), all that can be asserted is that every eigenvector of C corresponding to this eigenvalue is written:

$$\lambda\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle + \mu\,|\,\varphi_n(1)\,\rangle\,|\,\chi_p(2)\,\rangle \tag{F-31}$$

where λ and μ are arbitrary complex numbers. In this case, therefore, there exist eigenvectors of C which are not tensor products.

F-3-b. Complete sets of commuting observables in $\mathscr{E}$

We are finally going to show that if a C.S.C.O. has been chosen in both spaces $\mathscr{E}_1$ and $\mathscr{E}_2$, obtaining one in $\mathscr{E}$ is straightforward.

As an example, let us consider the case where $A(1)$ constitutes a C.S.C.O. by itself in $\mathscr{E}_1$, and the C.S.C.O. in $\mathscr{E}_2$ is composed of two observables, $B(2)$ and $C(2)$. This means (*cf.* § D-3-b) that all the eigenvalues a_n of $A(1)$ are nondegenerate in $\mathscr{E}_1$:

$$A(1)\,|\,\varphi_n(1)\,\rangle = a_n\,|\,\varphi_n(1)\,\rangle \tag{F-32}$$

the ket $|\,\varphi_n(1)\,\rangle$ being unique to within a constant factor. On the other hand, in $\mathscr{E}_2$, some of the eigenvalues b_p of $B(2)$ are degenerate, as are some of the eigenvalues (c_r) of $C(2)$. Nevertheless, the basis of eigenvectors common to $B(2)$ and $C(2)$ is unique in $\mathscr{E}_2$, since there exists only one ket (to within a constant factor) which is an eigenvector of $B(2)$ and of $C(2)$ with the eigenvalues b_p and c_r, fixed:

$$\begin{cases} B(2)\,|\,\chi_{pr}(2)\,\rangle = b_p\,|\,\chi_{pr}(2)\,\rangle \\ C(2)\,|\,\chi_{pr}(2)\,\rangle = c_r\,|\,\chi_{pr}(2)\,\rangle \\ |\,\chi_{pr}(2)\,\rangle \text{ unique to within a constant factor} \end{cases} \tag{F-33}$$

In $\mathscr{E}$, each of the eigenvalues a_n is N_2-fold degenerate (*cf.* § F-3-a). Therefore, $A(1)$ no longer forms a C.S.C.O. by itself. Similarly, there exist N_1 linearly independent kets which are eigenvectors of $B(2)$ and $C(2)$ with the eigenvalues b_p and c_r respectively, and the set $\{ B(2), C(2) \}$ is not complete either. However, we saw in § F-3-a that the eigenvectors which are common to the three commuting observables $A(1)$, $B(2)$ and $C(2)$ are the $|\varphi_n(1)\,\chi_{pr}(2)\rangle = |\varphi_n(1)\rangle\,|\chi_{pr}(2)\rangle$:

$$\begin{aligned} A(1)\,|\varphi_n(1)\,\chi_{pr}(2)\rangle &= a_n\,|\varphi_n(1)\,\chi_{pr}(2)\rangle \\ B(2)\,|\varphi_n(1)\,\chi_{pr}(2)\rangle &= b_p\,|\varphi_n(1)\,\chi_{pr}(2)\rangle \\ C(2)\,|\varphi_n(1)\,\chi_{pr}(2)\rangle &= c_r\,|\varphi_n(1)\,\chi_{pr}(2)\rangle \end{aligned} \tag{F-34}$$

The system $\{ |\varphi_n(1)\,\chi_{pr}(2)\rangle \}$ constitutes a basis in $\mathscr{E}$, since this is the case for $\{ |\varphi_n(1)\rangle \}$ and $\{ |\chi_{pr}(2)\rangle \}$ in $\mathscr{E}_1$ and $\mathscr{E}_2$ respectively. Moreover, if a set of three eigenvalues $\{ a_n,\ b_p,\ c_r \}$ is chosen, only one vector $|\varphi_n(1)\,\chi_{pr}(2)\rangle$ corresponds to it. $A(1)$, $B(2)$ and $C(2)$ therefore constitute a C.S.C.O. in $\mathscr{E}$.

The preceding argument can be generalized without difficulty: *by joining two sets of commuting observables which are complete in $\mathscr{E}_1$ and $\mathscr{E}_2$ respectively, one obtains a complete set of commuting observables in $\mathscr{E}$.*

F-4. Applications

F-4-a. One- and three-dimensional particle states

α. State spaces

Consider again, in the light of the preceding discussion, the problem posed in the introduction (§ F-1): how are $\mathscr{E}_x$ and $\mathscr{E}_{\mathbf{r}}$ related?

$\mathscr{E}_x$ is the state space of a particle moving in one dimension, that is, the state space associated with the wave functions $\varphi(x)$. In $\mathscr{E}_x$, the observable X which was studied in § E-2 constitutes a C.S.C.O. by itself (§ E-2-d); its eigenvectors are the basis kets of the $\{ |x\rangle \}$ representation. A vector $|\varphi\rangle$ of $\mathscr{E}_x$ is characterized, in this representation, by a wave function $\varphi(x) = \langle x|\varphi\rangle$; in particular, the basis ket $|x_0\rangle$ corresponds to the wave function $\xi_{x_0}(x) = \delta(x - x_0)$.

In the same way, it is possible to introduce the spaces $\mathscr{E}_y$ and $\mathscr{E}_z$ associated with the wave functions $\chi(y)$ and $\omega(z)$. The observable Y forms a C.S.C.O. in $\mathscr{E}_y$, as does Z in $\mathscr{E}_z$. The corresponding eigenvectors are the basis kets of the $\{ |y\rangle \}$ and $\{ |z\rangle \}$ representations of $\mathscr{E}_y$ and $\mathscr{E}_z$ respectively. A vector $|\chi\rangle$ of $\mathscr{E}_y$ (or $|\omega\rangle$ of $\mathscr{E}_z$) is characterized in the $\{ |y\rangle \}$ (or $\{ |z\rangle \}$ representation by a function $\chi(y) = \langle y|\chi\rangle$ (or $\omega(z) = \langle z|\omega\rangle$). The function which corresponds to the basis ket $|y_0\rangle$ (or $|z_0\rangle$) is $\delta(y - y_0)$ (or $\delta(z - z_0)$).

Let us then form the tensor product:

$$\mathscr{E}_{xyz} = \mathscr{E}_x \otimes \mathscr{E}_y \otimes \mathscr{E}_z \tag{F-35}$$

We obtain a basis in $\mathscr{E}_{xyz}$ from the tensor product of the $\{ |x\rangle \}$, $\{ |y\rangle \}$ and $\{ |z\rangle \}$ bases. We shall denote it by $\{ |x, y, z\rangle \}$, with:

$$|x, y, z\rangle = |x\rangle\,|y\rangle\,|z\rangle \tag{F-36}$$

The basis kets are simultaneous eigenvectors of the X, Y and Z operators extended into $\mathscr{E}_{xyz}$:

$$\begin{aligned} X\,|\,x,y,z\,\rangle &= x\,|\,x,y,z\,\rangle \\ Y\,|\,x,y,z\,\rangle &= y\,|\,x,y,z\,\rangle \\ Z\,|\,x,y,z\,\rangle &= z\,|\,x,y,z\,\rangle \end{aligned} \tag{F-37}$$

Therefore, $\mathscr{E}_{xyz}$ coincides with $\mathscr{E}_{\mathbf{r}}$, the state space of a three-dimensional particle, and $|\,x,y,z\,\rangle$ with $|\,\mathbf{r}\,\rangle$:

$$|\,x,y,z\,\rangle \equiv |\,\mathbf{r}\,\rangle = |\,x\,\rangle\,|\,y\,\rangle\,|\,z\,\rangle \tag{F-38}$$

where x, y, z are precisely the cartesian coordinates of $\mathbf{r}$.

There exist in $\mathscr{E}_{\mathbf{r}}$ kets $|\,\varphi\;\chi\;\omega\,\rangle = |\,\varphi\,\rangle\,|\,\chi\,\rangle\,|\,\omega\,\rangle$ that are the tensor products of three kets, one of $\mathscr{E}_x$, one of $\mathscr{E}_y$ and one of $\mathscr{E}_z$. The components in the $\{\,|\,\mathbf{r}\,\rangle\,\}$ representation are then [*cf.* formula (F-8)]:

$$\langle\,\mathbf{r}\,|\,\varphi\;\chi\;\omega\,\rangle = \langle\,x\,|\,\varphi\,\rangle\,\langle\,y\,|\,\chi\,\rangle\,\langle\,z\,|\,\omega\,\rangle \tag{F-39}$$

The associated wave functions are thus factorized: $\varphi(x)\;\chi(y)\;\omega(z)$. This is the case for the basis vectors themselves:

$$\langle\,\mathbf{r}\,|\,\mathbf{r}_0\,\rangle = \delta(\mathbf{r}-\mathbf{r}_0) = \delta(x-x_0)\;\delta(y-y_0)\;\delta(z-z_0) \tag{F-40}$$

Note that the most general state of $\mathscr{E}_{\mathbf{r}}$ is not such a product. It is written:

$$|\,\psi\,\rangle = \int \mathrm{d}x\;\mathrm{d}y\;\mathrm{d}z\;\;\psi(x,y,z)\;|\,x,y,z\,\rangle \tag{F-41}$$

In $\psi(x,y,z) = \langle\,x,y,z\,|\,\psi\,\rangle$, the x-, y- and z-dependences cannot, in general, be factorized: each of the wave functions associated with the kets of $\mathscr{E}_{\mathbf{r}}$ is a wave function with three variables.

The results of § F-3 thus enable us to understand why X, which constitutes a C.S.C.O. by itself in $\mathscr{E}_x$, no longer has this property in $\mathscr{E}_{\mathbf{r}}$ (*cf.* § E-2-d): the eigenvalues of its extension in $\mathscr{E}_{\mathbf{r}}$ are the same as in $\mathscr{E}_x$, but they become infinitely degenerate because $\mathscr{E}_y$ and $\mathscr{E}_z$ are infinite-dimensional. Starting with a C.S.C.O. in $\mathscr{E}_x$, $\mathscr{E}_y$ and $\mathscr{E}_z$, we construct one for $\mathscr{E}_{\mathbf{r}}$: $\{\,X,Y,Z\,\}$, for example, but also $\{\,P_x,Y,Z\,\}$ since P_x forms a C.S.C.O. in $\mathscr{E}_x$, or $\{\,P_x,P_y,Z\,\}$, etc...

β. *An important application*

Let us try to solve in $\mathscr{E}_{\mathbf{r}}$ the eigenvalue equation of an operator H such that:

$$H = H_x + H_y + H_z \tag{F-42}$$

where H_x, H_y and H_z are the extensions of observables acting respectively in $\mathscr{E}_x$, $\mathscr{E}_y$ and $\mathscr{E}_z$. In practice, one recognizes that H_x, for example, is the extension of an observable of $\mathscr{E}_x$ because it is constructed using only the operators X and P_x. Using the reasoning of

§ F-3-a-β, one first looks for the eigenvalues and eigenvectors of H_x in $\mathscr{E}_x$, H_y in $\mathscr{E}_y$ and H_z in $\mathscr{E}_z$:

$$\begin{aligned} H_x \,|\, \varphi_n \,\rangle &= E_x^n \,|\, \varphi_n \,\rangle \\ H_y \,|\, \chi_p \,\rangle &= E_y^p \,|\, \chi_p \,\rangle \\ H_z \,|\, \omega_r \,\rangle &= E_z^r \,|\, \omega_r \,\rangle \end{aligned} \tag{F-43}$$

The eigenvalues of H are then all of the form:

$$E^{n,p,r} = E_x^n + E_y^p + E_z^r \tag{F-44}$$

with an eigenvector that is the tensor product $|\, \varphi_n \,\rangle \,|\, \chi_p \,\rangle \,|\, \omega_r \,\rangle$; the wave function associated with this vector is the product:

$$\varphi_n(x)\; \chi_p(y)\; \omega_r(z) = \langle\, x \,|\, \varphi_n \,\rangle\; \langle\, y \,|\, \chi_p \,\rangle\; \langle\, z \,|\, \omega_r \,\rangle$$

This is the type of situation that was considered in Complement $\mathrm{F_I}$ (§ 2) for the justification of the study of one-dimensional models. There, we were dealing with differential operators acting on wave functions:

$$H = -\frac{\hbar^2}{2m}\,\Delta + V(\mathbf{r}) \tag{F-45}$$

This equation can be decomposed as in (F-42) in the particular case where the potential can be written:

$$V(\mathbf{r}) = V_1(x) + V_2(y) + V_3(z) \tag{F-46}$$

F-4-b. States of a two-particle system

Consider a physical system which is made up of two (spinless) particles. We shall distinguish between them by numbering them (1) and (2). To describe the system quantum mechanically, we can generalize the concept of a wave function, introduced for the case of one particle. A state of the system can be characterized, at a given time, by a function of six spatial variables $\psi(\mathbf{r}_1, \mathbf{r}_2) \equiv \psi(x_1, y_1, z_1; x_2, y_2, z_2)$. The probabilistic interpretation of such a two-particle wave function is the following: the probability $\mathrm{d}\mathscr{P}(\mathbf{r}_1, \mathbf{r}_2)$, at the given time, of finding particle (1) in the volume $\mathrm{d}^3r_1 = \mathrm{d}x_1\; \mathrm{d}y_1\; \mathrm{d}z_1$ situated at the point $\mathbf{r}_1$, *and* particle (2) in the volume $\mathrm{d}^3r_2 = \mathrm{d}x_2\; \mathrm{d}y_2\; \mathrm{d}z_2$ about $\mathbf{r}_2$, is:

$$\mathrm{d}\mathscr{P}(\mathbf{r}_1, \mathbf{r}_2) = C\; |\psi(\mathbf{r}_1, \mathbf{r}_2)|^2\; \mathrm{d}^3r_1\; \mathrm{d}^3r_2 \tag{F-47}$$

The normalization constant C is obtained by imposing the condition that the total probability must be equal to 1 (conservation of the number of particles; *cf.* § B-2 of Chapter I):

$$\frac{1}{C} = \int \mathrm{d}^3r_1\; \mathrm{d}^3r_2\; |\psi(\mathbf{r}_1, \mathbf{r}_2)|^2 \tag{F-48}$$

and the observables X_1, Y_1, Z_1 can be defined in $\mathscr{E}_{\mathbf{r}_1}$. Similarly, in the state space $\mathscr{E}_{\mathbf{r}_2}$ of particle (2), we introduce the $\{\, |\, \mathbf{r}_2 \,\rangle \,\}$ representation and the observables X_2, Y_2, Z_2. Take the tensor product:

$$\mathscr{E}_{\mathbf{r}_1\mathbf{r}_2} = \mathscr{E}_{\mathbf{r}_1} \otimes \mathscr{E}_{\mathbf{r}_2} \tag{F-49}$$

The set of vectors:

$$|\,\mathbf{r}_1,\mathbf{r}_2\,\rangle = |\,\mathbf{r}_1\,\rangle\,|\,\mathbf{r}_2\,\rangle \tag{F-50}$$

forms a basis in $\mathscr{E}_{\mathbf{r}_1\mathbf{r}_2}$. Consequently, every ket $|\,\psi\,\rangle$ of this space can be written:

$$|\,\psi\,\rangle = \int \mathrm{d}^3r_1\ \mathrm{d}^3r_2\ \ \psi(\mathbf{r}_1,\mathbf{r}_2)\,|\,\mathbf{r}_1,\mathbf{r}_2\,\rangle \tag{F-51}$$

with

$$\psi(\mathbf{r}_1,\mathbf{r}_2) = \langle\,\mathbf{r}_1,\mathbf{r}_2\,|\,\psi\,\rangle \tag{F-52}$$

Moreover, the square of the norm of $|\,\psi\,\rangle$ is equal to:

$$\langle\,\psi\,|\,\psi\,\rangle = \int \mathrm{d}^3r_1\ \mathrm{d}^3r_2\ \ |\psi(\mathbf{r}_1,\mathbf{r}_2)|^2 \tag{F-53}$$

For it to be finite, $\psi(\mathbf{r}_1,\mathbf{r}_2)$ must be square-integrable. Therefore, a wave function $\psi(\mathbf{r}_1,\mathbf{r}_2)$ is associated with each ket of $\mathscr{E}_{\mathbf{r}_1\mathbf{r}_2}$: *the state space of a two-particle system is the tensor product of the spaces which correspond to each of the particles.* A C.S.C.O. is obtained in $\mathscr{E}_{\mathbf{r}_1\mathbf{r}_2}$ by joining, for example, X_1, Y_1, Z_1 and X_2, Y_2, Z_2.

Assume that the state of the system is described by a tensor product ket:

$$|\,\psi\,\rangle = |\,\psi_1\,\rangle\,|\,\psi_2\,\rangle \tag{F-54}$$

The corresponding wave function can then be factorized:

$$\psi(\mathbf{r}_1,\mathbf{r}_2) = \langle\,\mathbf{r}_1,\mathbf{r}_2\,|\,\psi\,\rangle = \langle\,\mathbf{r}_1\,|\,\psi_1\,\rangle\,\langle\,\mathbf{r}_2\,|\,\psi_2\,\rangle = \psi_1(\mathbf{r}_1)\ \psi_2(\mathbf{r}_2) \tag{F-55}$$

In this case, one says that *there is no correlation* between the two particles. We shall analyze later (Complement $\mathrm{D_{III}}$) the physical consequences of such a situation.

The preceding can be generalized: when a physical system is composed of the union of two or several simpler systems, its state space is the tensor product of the spaces which correspond to each of the component systems.

References and suggestions for further reading:

Section 10 of the bibliography contains references to a certain number of mathematical texts, listed by subject. Under each heading, they are ranked, as much as possible, in order of increasing difficulty. See also the quantum mechanics texts (sections 1 and 2 of the bibliography), which treat the mathematical problems at many different levels. They also contain other references.

For a very simple approach to the fundamental mathematical concepts needed to understand Chapter II (vector spaces, operators, diagonalization of matrices, etc.), the reader can consult, for example: Arfken (10.4), Chap. 4; Bak and Lichtenberg (10.3), Chap. I; Bass (10.1), vol. I, Chap. II to V. A more explicit application to quantum mechanics can be found in Jackson (10.5) (see, in particular, Chap. 5), Butkov (10.8), Chap. 10 (finite-dimensional linear spaces) and Chap. 11 (infinite-dimensional vector spaces, spaces of functions). See also Meijer and Bauer (2.18), Chap. 1, particularly the table at the end of this chapter.

A_{II} : THE SCHWARZ INEQUALITY **B_{II} : REVIEW OF SOME USEFUL PROPERTIES OF LINEAR OPERATORS** **C_{II} : UNITARY OPERATORS**	Review of some definitions and useful mathematical results (elementary level) intended for readers unfamiliar with these concepts; will serve as a reference later (especially B_{II}).
D_{II} : A MORE DETAILED STUDY OF THE $\{\|\mathbf{r}\rangle\}$ AND $\{\|\mathbf{p}\rangle\}$ REPRESENTATIONS **E_{II} : SOME GENERAL PROPERTIES OF TWO OBSERVABLES, Q AND P, WHOSE COMMUTATOR IS EQUAL TO $i\hbar$**	Complete § E of Chapter II. D_{II}: remains at the level of Chapter II and can be read immediately after it. E_{II}: adopts a more general and a slightly more formal point of view. Introduces, in particular, the translation operator. May be reserved for later study.
F_{II} : THE PARITY OPERATOR	Discussion of the parity operator, particularly important in quantum mechanics; at the same time, a simple illustration of the concepts of Chapter II ; recommended for these two reasons.
G_{II} : AN APPLICATION OF THE PROPERTIES OF THE TENSOR PRODUCT: THE TWO-DIMENSIONAL INFINITE WELL	A simple application of the tensor product (§ F of Chapter II); can be considered as a worked exercise.
H_{II} : EXERCISES	Solutions are given for exercises 11 and 12; their aim is to familiarize the reader with the properties of commuting observables and the concept of a C.S.C.O. in a very simple special case. It is recommended that these exercises be done during the reading of § D-3 of Chapter II.

Complement A_{II}

The Schwarz inequality

For any ket $|\psi\rangle$ belonging to the state space $\mathcal{E}$, we have:

$$\langle\psi|\psi\rangle \quad \text{real} \geq 0 \tag{1}$$

$\langle\psi|\psi\rangle$ being equal to zero only when $|\psi\rangle$ is the null vector [*cf.* equation (B-12) of Chapter II]. Using inequality (1), we shall derive the Schwarz inequality. This inequality states that, if $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are any arbitrary vectors of $\mathcal{E}$, then:

$$\boxed{|\langle\varphi_1|\varphi_2\rangle|^2 \leq \langle\varphi_1|\varphi_1\rangle\langle\varphi_2|\varphi_2\rangle} \tag{2}$$

the equality being realized if and only if $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are proportional.

Given $|\varphi_1\rangle$ and $|\varphi_2\rangle$, consider the ket $|\psi\rangle$ defined by:

$$|\psi\rangle = |\varphi_1\rangle + \lambda|\varphi_2\rangle \tag{3}$$

where λ is an arbitrary parameter. Whatever λ may be:

$$\langle\psi|\psi\rangle = \langle\varphi_1|\varphi_1\rangle + \lambda\langle\varphi_1|\varphi_2\rangle + \lambda^*\langle\varphi_2|\varphi_1\rangle + \lambda\lambda^*\langle\varphi_2|\varphi_2\rangle \geq 0 \tag{4}$$

Let us chose for λ the value:

$$\lambda = -\frac{\langle\varphi_2|\varphi_1\rangle}{\langle\varphi_2|\varphi_2\rangle} \tag{5}$$

In (4), the second and third terms of the right-hand side are then equal, and opposite in value to the fourth term, so that (4) reduces to:

$$\langle\varphi_1|\varphi_1\rangle - \frac{\langle\varphi_1|\varphi_2\rangle\langle\varphi_2|\varphi_1\rangle}{\langle\varphi_2|\varphi_2\rangle} \geq 0 \tag{6}$$

Since $\langle\varphi_2|\varphi_2\rangle$ is positive, we can multiply this inequality by $\langle\varphi_2|\varphi_2\rangle$, to obtain:

$$\langle\varphi_1|\varphi_1\rangle\ \langle\varphi_2|\varphi_2\rangle \geq \langle\varphi_1|\varphi_2\rangle\ \langle\varphi_2|\varphi_1\rangle \tag{7}$$

which is precisely (2). In (7), the equality can only be realized if $\langle\psi|\psi\rangle = 0$, that is, according to (3), if $|\varphi_1\rangle = -\lambda|\varphi_2\rangle$. The kets $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are then proportional.

References:

Bass I (10.1), § 5-3; Arfken (10.4), § 9-4.

Complement B_{II}

Review of some useful properties of linear operators

1 Trace of an operator 163
1-a Definition 163
1-b The trace is invariant 164
1-c Important properties 164
2 Commutator algebra 165
2-a Definition 165
2-b Properties 165
3 Restriction of an operator to a subspace 165
4 Functions of operators 166
4-a Definition and simple properties 166
4-b An important example: the potential operator 168
4-c Commutators involving functions of operators 168
5 Derivative of an operator 169
5-a Definition 169
5-b Differentiation rules 170
5-c Examples 170
5-d An application: a useful formula 171

The aim of this complement is to review a certain number of definitions and useful properties of linear operators.

1. Trace of an operator

1-a. Definition

The trace of an operator A, written Tr A, is the sum of its diagonal matrix elements.

When a discrete orthonormal basis, $\{|u_i\rangle\}$, is chosen for the space $\mathcal{E}$, one has, by definition:

$$\text{Tr } A = \sum_i \langle u_i|A|u_i\rangle \tag{1}$$

For the case of a continuous orthonormal basis $\{|w_\alpha\rangle\}$, one has:

$$\text{Tr } A = \int \mathrm{d}\alpha \ \langle w_\alpha|A|w_\alpha\rangle \tag{2}$$

When $\mathcal{E}$ is an infinite-dimensional space, the trace of the operator A is defined only if expressions (1) and (2) converge.

1-b. The trace is invariant

The sum of the diagonal elements of the matrix which represents an operator A in an arbitrary basis does not depend on this basis.

Let us derive this property for the case of a change from one discrete orthonormal basis $\{|u_i\rangle\}$ to another discrete orthonormal basis $\{|t_k\rangle\}$. We have:

$$\sum_i \langle u_i|A|u_i\rangle = \sum_i \langle u_i| \left[\sum_k |t_k\rangle\langle t_k|\right] A|u_i\rangle \tag{3}$$

(where we have used the closure relation for the $|t_k\rangle$ states). The right-hand side of (3) is equal to:

$$\sum_{i,k} \langle u_i|t_k\rangle\langle t_k|A|u_i\rangle = \sum_{i,k} \langle t_k|A|u_i\rangle\langle u_i|t_k\rangle \tag{4}$$

(since it is possible to change the order of two numbers in a product). We can then replace $\sum_i |u_i\rangle\langle u_i|$ in (4) by $\mathbb{1}$ (closure relation for the $|u_i\rangle$ states), and we obtain, finally:

$$\sum_i \langle u_i|A|u_i\rangle = \sum_k \langle t_k|A|t_k\rangle \tag{5}$$

We have therefore demonstrated the property of invariance for this case.

Comment:

If the operator A is an observable, Tr A can therefore be calculated in a basis of eigenvectors of A. The diagonal matrix elements are then the eigenvalues a_n of A (degree of degeneracy g_n) and the trace can be written:

$$\mathrm{Tr}\, A = \sum_n g_n\, a_n \tag{6}$$

1-c. Important properties

$$\mathrm{Tr}\, AB = \mathrm{Tr}\, BA \tag{7a}$$

$$\mathrm{Tr}\, ABC = \mathrm{Tr}\, BCA = \mathrm{Tr}\, CAB \tag{7b}$$

In general, the trace of the product of any number of operators is invariant when a cyclic permutation is performed on these operators.

Let us prove, for example, relation (7a):

$$\begin{aligned}\mathrm{Tr}\, AB &= \sum_i \langle u_i|AB|u_i\rangle = \sum_{i,j} \langle u_i|A|u_j\rangle\, \langle u_j|B|u_i\rangle \\ &= \sum_{i,j} \langle u_j|B|u_i\rangle\, \langle u_i|A|u_j\rangle = \sum_j \langle u_j|BA|u_j\rangle = \mathrm{Tr}\, BA\end{aligned} \tag{8}$$

(twice using the closure relation on the $\{|u_i\rangle\}$ basis). Relation (7a) is thus proved; its generalization (7b) presents no difficulty.

2. Commutator algebra

2-a. Definition

The commutator $[A, B]$ of two operators is, by definition:

$$[A, B] = AB - BA \tag{9}$$

2-b. Properties

$$[A, B] = -[B, A] \tag{10}$$
$$[A, (B + C)] = [A, B] + [A, C] \tag{11}$$
$$[A, BC] = [A, B]C + B[A, C] \tag{12}$$
$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{13}$$
$$[A, B]^\dagger = [B^\dagger, A^\dagger] \tag{14}$$

The derivation of these properties is straightforward: it suffices to compare both sides of each equation after having written them out explicitly.

3. Restriction of an operator to a subspace

Let P_q be the projector onto the q-dimensional subspace $\mathcal{E}_q$ spanned by the q orthonormal vectors $|\varphi_i\rangle$:

$$P_q = \sum_{i=1}^{q} |\varphi_i\rangle\langle\varphi_i| \tag{15}$$

By definition, the restriction $\hat{A}_q$ of the operator A to the subspace $\mathcal{E}_q$ is:

$$\hat{A}_q = P_q A P_q \tag{16}$$

If $|\psi\rangle$ is an arbitrary ket, it follows from this definition that:

$$\hat{A}_q|\psi\rangle = P_q A|\hat{\psi}_q\rangle \tag{17}$$

where:

$$|\hat{\psi}_q\rangle = P_q|\psi\rangle \tag{18}$$

is the orthogonal projection of $|\psi\rangle$ onto $\mathcal{E}_q$. Consequently, to make $\hat{A}_q$ act on an arbitrary ket $|\psi\rangle$, one begins by projecting this ket onto $\mathcal{E}_q$; then one lets the operator A act on this projection, retaining only the projection in $\mathcal{E}_q$ of the resulting ket. The operator $\hat{A}_q$, which transforms any ket of $\mathcal{E}_q$ into a ket belonging to this same subspace, is therefore an operator whose action has been restricted to $\mathcal{E}_q$.

What can be said about the matrix which represents $\hat{A}_q$? Let us choose a basis $\{|u_k\rangle\}$ whose first q vectors belong to $\mathcal{E}_q$ (they are, for example, the $|\varphi_i\rangle$), the others belonging to the supplementary subspace. We have:

$$\langle u_i|\hat{A}_q|u_j\rangle = \langle u_i|P_q A P_q|u_j\rangle \tag{19}$$

that is:

$$\langle u_i|\hat{A}_q|u_j\rangle = \begin{cases} \langle u_i|A|u_j\rangle & \text{if } i,\, j \leq q \\ 0 & \text{if one of the two indices } i \text{ or } j \text{ is greater than } q \end{cases} \tag{20}$$

Therefore, the matrix which represents $\hat{A}_q$ is, as it were, "cut out" of the one which represents A. One retains only the matrix elements of A associated with basis vectors $|u_i\rangle$ and $|u_j\rangle$, both belonging to $\mathcal{E}_q$, the other matrix elements being replaced by zeros.

4. Functions of operators

4-a. Definition and simple properties

Consider an arbitrary linear operator A. It is not difficult to define the operator A^n: it is the operator which corresponds to n successive applications of the operator A. The definition of the operator A^{-1}, the inverse of A, is also well known: A^{-1} is the operator (if it exists) which satisfies the relations:

$$A^{-1}A = AA^{-1} = \mathbb{1} \tag{21}$$

How can we define, in a more general way, an arbitrary function of an operator? To do this, let us consider a function F of a variable z. Assume that, in a certain domain, F can be expanded in a power series in z:

$$F(z) = \sum_{n=0}^{\infty} f_n z^n \tag{22}$$

By definition, the corresponding function of the operator A is the operator $F(A)$ defined by a series which has the same coefficients f_n:

$$F(A) = \sum_{n=0}^{\infty} f_n A^n \tag{23}$$

For example, the operator e^A is defined by:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = \mathbb{1} + A + A^2/2! + ... + A^n/n! + ... \tag{24}$$

We shall not consider the problems concerning the convergence of the series (23), which depends on the eigenvalues of A and on the radius of convergence of the series (22).

Note that if $F(z)$ is a real function, the coefficients f_n are real. If, moreover, A is Hermitian, we see from (23) that $F(A)$ is Hermitian.

Let $|\varphi_a\rangle$ be an eigenvector of A with eigenvalue a:

$$A|\varphi_a\rangle = a|\varphi_a\rangle \tag{25}$$

Applying the operator n times in succession, we obtain:

$$A^n|\varphi_a\rangle = a^n|\varphi_a\rangle \tag{26}$$

Now let us apply series (23) to $|\varphi_a\rangle$; we obtain:

$$F(A)|\varphi_a\rangle = \sum_{n=0}^{\infty} f_n a^n |\varphi_a\rangle = F(a)|\varphi_a\rangle \tag{27}$$

This leads to the following rule: *when $|\varphi_a\rangle$ is an eigenvector of A with the eigenvalue a, $|\varphi_a\rangle$ is also an eigenvector of $F(A)$, with the eigenvalue $F(a)$.*

This property leads to a second definition of a function of an operator. Let us consider a diagonalizable operator A (this is always the case if A is an observable), and let us choose a basis where the matrix associated with A is actually diagonal (its elements are then the eigenvalues a_i of A). $F(A)$ is, by definition, the operator which is represented, in this same basis, by the diagonal matrix whose elements are $F(a_i)$.

For example, if σ_z, is the matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{28}$$

it follows directly that:

$$\mathrm{e}^{\sigma_z} = \begin{pmatrix} \mathrm{e} & 0 \\ 0 & 1/\mathrm{e} \end{pmatrix} \tag{29}$$

Comment:

Attention should be given, when functions of operators are used, with respect to the order of the operators. For example, the operators $\mathrm{e}^A\mathrm{e}^B$, $\mathrm{e}^B\mathrm{e}^A$, and e^{A+B} are not, in general, equal when A and B are operators and not numbers. Consider:

$$\mathrm{e}^A\mathrm{e}^B = \sum_p \frac{A^p}{p!} \sum_q \frac{B^q}{q!} = \sum_{p,q} \frac{A^p\, B^q}{p!\, q!} \tag{30}$$

$$\mathrm{e}^B\mathrm{e}^A = \sum_q \frac{B^q}{q!} \sum_p \frac{A^p}{p!} = \sum_{p,q} \frac{B^q\, A^p}{p!\, q!} \tag{31}$$

$$\mathrm{e}^{A+B} = \sum_p \frac{(A+B)^p}{p!} \tag{32}$$

When A and B are arbitrary, the right-hand sides of (30), (31) and (32) have no reason to be equal (see exercise 7 of Complement $\mathrm{H_{II}}$). However, *when A and B commute*, we have:

$$[A, B] = 0 \Longrightarrow \mathrm{e}^A\mathrm{e}^B = \mathrm{e}^B\mathrm{e}^A = \mathrm{e}^{A+B} \tag{33}$$

(an obvious relation if the diagonal matrices that represent e^A and e^B are considered in a basis of eigenvectors common to A and B).

4-b. An important example: the potential operator

In one-dimensional problems, we shall often have to consider "potential" operators $V(X)$ (so called because they correspond to the classical potential energy $V(x)$ of a particle placed in a force field), where $V(X)$ is a function of the position operator X.

It follows from the preceding section that $V(X)$ has as eigenvectors the eigenvectors $|x\rangle$ of X, and we have simply:

$$V(X)|x\rangle = V(x)|x\rangle \tag{34}$$

The matrix elements of $V(X)$ in the $\{|x\rangle\}$ representation are therefore:

$$\langle x|V(X)|x'\rangle = V(x)\delta(x-x') \tag{35}$$

Applying (34) and using the fact that $V(X)$ is Hermitian (the function $V(x)$ is real), we obtain:

$$\langle x|V(X)|\psi\rangle = V(x)\langle x|\psi\rangle = V(x)\psi(x) \tag{36}$$

This equation shows that in the $\{|x\rangle\}$ representation, the action of the operator $V(X)$ is simply multiplication by $V(x)$.

The generalization of (34), (35) and (36) to three-dimensional problems can be performed without difficulty; in this case, we obtain:

$$V(\mathbf{R})|\mathbf{r}\rangle = V(\mathbf{r})|\mathbf{r}\rangle \tag{37}$$

$$\langle \mathbf{r}|V(\mathbf{R})|\mathbf{r}'\rangle = V(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') \tag{38}$$

$$\langle \mathbf{r}|V(\mathbf{R})|\psi\rangle = V(\mathbf{r})\psi(\mathbf{r}) \tag{39}$$

4-c. Commutators involving functions of operators

Definition (23) shows that A commutes with every function of A:

$$[A, F(A)] = 0 \tag{40}$$

Similarly, if A and B commute, so do $F(A)$ and B:

$$[B, A] = 0 \Longrightarrow [B, F(A)] = 0 \tag{41}$$

What will be the commutator of an operator with a function of another operator that does not commute with it? We shall restrict ourselves here to the case of the X and P operators, whose commutator is equal to:

$$[X, P] = i\hbar \tag{42}$$

Using relation (12), we can calculate:

$$[X, P^2] = [X, PP] = [X, P]P + P[X, P] = 2i\hbar\, P \tag{43}$$

More generally, let us show that:

$$[X, P^n] = i\hbar\ n\ P^{n-1} \tag{44}$$

If we assume that this equation is verified, we obtain:

$$[X, P^{n+1}] = [X, PP^n] = [X, P]P^n + P[X, P^n]$$
$$= i\hbar\, P^n + i\hbar n\, PP^{n-1} = i\hbar(n+1)P^n \tag{45}$$

Relation (44) is therefore established by recurrence.

Now let us calculate the commutator $[X, F(P)]$:

$$[X, F(P)] = \sum_n [X, f_n P^n] = \sum_n i\hbar n f_n P^{n-1} \tag{46}$$

If $F'(z)$ denotes the derivative of the function $F(z)$, we recognize in (46) the definition of the operator $F'(P)$. Therefore:

$$\boxed{[X, F(P)] = i\hbar\, F'(P)} \tag{47}$$

An analogous argument would have enabled us to obtain the symmetric relation:

$$\boxed{[P, G(X)] = -i\hbar\, G'(X)} \tag{48}$$

Comments:

(*i*) The preceding argument is based on the fact that $F(P)$ (or $G(X)$) depends only on P (or on X). It is more difficult to calculate a commutator such as $[X, \Phi(X, P)]$, where $\Phi(X, P)$ is an operator which depends on both X and P: the difficulties arise from the fact that X and P do not commute.

(*ii*) Equations (47) and (48) can be generalized to the case of two operators A and B which both commute with their commutator. An argument modeled on the preceding one shows that, if we have:

$$[A, C] = [B, C] = 0 \tag{49}$$

with:

$$C = [A, B] \tag{50}$$

then:

$$[A, F(B)] = [A, B]F'(B) \tag{51}$$

5. Derivative of an operator

5-a. Definition

Let $A(t)$ be an operator which depends on an arbitrary variable t. By definition, the derivative $\dfrac{\mathrm{d}A}{\mathrm{d}t}$ of $A(t)$ with respect to t is given by the limit (if it exists):

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{A(t + \Delta t) - A(t)}{\Delta t} \tag{52}$$

The matrix elements of $A(t)$ in an arbitrary basis of t-independent vectors $|u_i\rangle$ are functions of t:

$$\langle u_i|A|u_j\rangle = A_{ij}(t) \tag{53}$$

Let us call $\left(\frac{\mathrm{d}A}{\mathrm{d}t}\right)_{ij} = \left\langle u_i \left| \frac{\mathrm{d}A}{\mathrm{d}t} \right| u_j \right\rangle$ the matrix elements of $\frac{\mathrm{d}A}{\mathrm{d}t}$. It is easy to verify the relation:

$$\left(\frac{\mathrm{d}A}{\mathrm{d}t}\right)_{ij} = \frac{\mathrm{d}}{\mathrm{d}t}A_{ij} \tag{54}$$

Thus we obtain a very simple rule: to obtain the matrix elements representing $\frac{\mathrm{d}A}{\mathrm{d}t}$, all we must do is take the matrix representing A and differentiate each of its elements (without changing their places).

5-b. Differentiation rules

They are analogous to the ones for ordinary functions:

$$\frac{\mathrm{d}}{\mathrm{d}t}(F+G) = \frac{\mathrm{d}F}{\mathrm{d}t} + \frac{\mathrm{d}G}{\mathrm{d}t} \tag{55}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(FG) = \frac{\mathrm{d}F}{\mathrm{d}t}G + F\frac{\mathrm{d}G}{\mathrm{d}t} \tag{56}$$

Nevertheless, care must be taken not to modify the order of the operators in formula (56).

Let us prove, for example, the second of these equations. The matrix elements of FG are:

$$\langle u_i|FG|u_j\rangle = \sum_k \langle u_i|F|u_k\rangle \langle u_k|G|u_j\rangle \tag{57}$$

We have seen that the matrix elements of $\mathrm{d}(FG)/\mathrm{d}t$ are the derivatives with respect to t of those of (FG). Thus we have, taking the derivative of the right-hand side of (57):

$$\begin{aligned}\left\langle u_i \left| \frac{\mathrm{d}}{\mathrm{d}t}(FG) \right| u_j \right\rangle &= \sum_k \left[\langle u_i \left| \frac{\mathrm{d}F}{\mathrm{d}t} \right| u_k\rangle \langle u_k|G|u_j\rangle + \langle u_i|F|u_k\rangle \left\langle u_k \left| \frac{\mathrm{d}G}{\mathrm{d}t} \right| u_j \right\rangle \right] \\ &= \left\langle u_i \left| \frac{\mathrm{d}F}{\mathrm{d}t}G + F\frac{\mathrm{d}G}{\mathrm{d}t} \right| u_j \right\rangle \end{aligned} \tag{58}$$

This equation is valid for any i and j. Formula (56) is thus established.

5-c. Examples

Let us calculate the derivative of the operator e^{At}. By definition, we have:

$$\mathrm{e}^{At} = \sum_{n=0}^{\infty} \frac{(At)^n}{n!} \tag{59}$$

Taking the derivative of the series term by term, we obtain:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At} &= \sum_{n=0}^{\infty} n\frac{t^{n-1}A^n}{n!} \\ &= A\sum_{n=1}^{\infty}\frac{(At)^{n-1}}{(n-1)!} \\ &= \left[\sum_{n=1}^{\infty}\frac{(At)^{n-1}}{(n-1)!}\right]A \end{aligned} \tag{60}$$

We recognize inside the brackets the series that defines e^{At} (taking as the summation index $p = n - 1$). The result is therefore:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At} = A\,\mathrm{e}^{At} = \mathrm{e}^{At}A \tag{61}$$

In this simple case involving only one operator, it is unnecessary to pay attention to the order of the factors: e^{At} and A commute.

This is not the case if one is interested in taking the derivative of an operator such as $\mathrm{e}^{At}\mathrm{e}^{Bt}$. Applying (56) and (61), we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{At}\mathrm{e}^{Bt}) = A\mathrm{e}^{At}\mathrm{e}^{Bt} + \mathrm{e}^{At}\,B\mathrm{e}^{Bt} \tag{62}$$

The right-hand side of this equation can be transformed into $\mathrm{e}^{At}\,A\mathrm{e}^{Bt} + \mathrm{e}^{At}\,B\mathrm{e}^{Bt}$ or $\mathrm{e}^{At}\,A\mathrm{e}^{Bt} + \mathrm{e}^{At}\,\mathrm{e}^{Bt}B$, for example. However, we can never obtain (unless, of course, A and B commute) an expression such as $(A + B)\mathrm{e}^{At}\mathrm{e}^{Bt}$. In this case, the order of the operators is therefore important.

Comment:

Even when the function involves only one operator, taking the derivative cannot always be performed according to the rules valid for ordinary functions. For example, when $A(t)$ has an arbitrary time-dependence, the derivative $\dfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{A(t)}$ *is generally not equal to* $\dfrac{\mathrm{d}A}{\mathrm{d}t}\mathrm{e}^{A(t)}$. It can be seen by expanding $\mathrm{e}^{A(t)}$ in a power series in $A(t)$ that $A(t)$ and $\dfrac{\mathrm{d}A}{\mathrm{d}t}$ must commute for the equality to hold.

5-d. An application: a useful formula

Consider two operators A and B which, by hypothesis, both commute with their commutator. In this case, we shall derive the relation:

$$\mathrm{e}^{A}\mathrm{e}^{B} = \mathrm{e}^{A+B}\,\mathrm{e}^{\frac{1}{2}[A,B]} \tag{63}$$

(sometimes called Glauber's formula).

Let us define the operator $F(t)$, a function of the real variable t, by:

$$F(t) = \mathrm{e}^{At}\,\mathrm{e}^{Bt} \tag{64}$$

We have:

$$\frac{\mathrm{d}F}{\mathrm{d}t} = A\mathrm{e}^{At}\,\mathrm{e}^{Bt} + \mathrm{e}^{At}\,B\mathrm{e}^{Bt} = (A + \mathrm{e}^{At}\,B\mathrm{e}^{-At})F(t) \tag{65}$$

Since A and B commute with their commutator, formula (51) can be applied in order to calculate:

$$[\mathrm{e}^{At}, B] = t[A, B]\mathrm{e}^{At} \tag{66}$$

Therefore:

$$\mathrm{e}^{At}\,B = B\,\mathrm{e}^{At} + t[A, B]\mathrm{e}^{At} \tag{67}$$

Multiply both sides of this equation on the right by e^{-At}. Substituting the relation so obtained into (65), we obtain:

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \Big(A + B + t[A, B]\Big)F(t) \tag{68}$$

The operators $A + B$ and $[A, B]$ commute by hypothesis. We can therefore integrate the differential equation (68) as if $A + B$ and $[A, B]$ were numbers. This yields:

$$F(t) = F(0)\mathrm{e}^{(A+B)t+\frac{1}{2}[A,B]t^2} \tag{69}$$

Setting $t = 0$ in (64), we see that $F(0) = \mathbb{1}$, and we obtain for any time t:

$$F(t) = \mathrm{e}^{(A+B)t+\frac{1}{2}[A,B]t^2} \tag{70}$$

Let us then set $t = 1$; we obtain equation (63), which is thus proven.

Comment:

When the operators A and B are arbitrary, equation (63) is not in general valid: it is necessary that both A and B commute with $[A, B]$. This condition may seem very restrictive. Actually, in quantum mechanics, one often encounters operators whose commutator is a number: for example, X and P, or the operators a and $a^\dagger$ of the harmonic oscillator (*cf.* Chap. V).

References:

See the subsections "General texts" and "Linear algebra – Hilbert spaces" of section 10 of the bibliography.

Complement C_{II}

Unitary operators

1 General properties of unitary operators 173
1-a Definition and simple properties 173
1-b Unitary operators and change of bases 174
1-c Unitary matrices 176
1-d Eigenvalues and eigenvectors of a unitary operator 176
2 Unitary transformations of operators 177
3 The infinitesimal unitary operator 178

1. General properties of unitary operators

1-a. Definition and simple properties

By definition, an operator U is unitary if its inverse U^{-1} is equal to its adjoint $U^\dagger$:

$$U^\dagger U = U U^\dagger = \mathbb{1} \tag{1}$$

Consider two arbitrary vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ of $\mathcal{E}$, and their transforms $|\tilde{\psi}_1\rangle$ and $|\tilde{\psi}_2\rangle$ under the action of U:

$$\begin{aligned} |\tilde{\psi}_1\rangle &= U|\psi_1\rangle \\ |\tilde{\psi}_2\rangle &= U|\psi_2\rangle \end{aligned} \tag{2}$$

Let us calculate the scalar product $\langle\tilde{\psi}_1|\tilde{\psi}_2\rangle$; we obtain:

$$\langle\tilde{\psi}_1|\tilde{\psi}_2\rangle = \langle\psi_1|U^\dagger U|\psi_2\rangle = \langle\psi_1|\psi_2\rangle \tag{3}$$

The unitary transformation associated with the operator U therefore conserves the scalar product (and, consequently, the norm) in $\mathcal{E}$. When $\mathcal{E}$ is finite-dimensional, moreover, this property is characteristic of a unitary operator.

Comments:

(*i*) If A is a Hermitian operator, the operator $T = \mathrm{e}^{iA}$ is unitary, since:

$$T^\dagger = \mathrm{e}^{-iA^\dagger} = \mathrm{e}^{-iA} \tag{4}$$

and therefore:

$$\begin{aligned} T^\dagger T &= \mathrm{e}^{-iA}\mathrm{e}^{iA} = \mathbb{1} \\ T T^\dagger &= \mathrm{e}^{iA}\mathrm{e}^{-iA} = \mathbb{1} \end{aligned} \tag{5}$$

(obviously, $-iA$ commutes with iA).

(*ii*) The product of two unitary operators is also unitary. If U and V are unitary, we have:

$$U^\dagger U = UU^\dagger = \mathbb{1}$$
$$V^\dagger V = VV^\dagger = \mathbb{1} \tag{6}$$

Let us now calculate:

$$(UV)^\dagger(UV) = V^\dagger U^\dagger UV = V^\dagger V = \mathbb{1}$$
$$(UV)(UV)^\dagger = UVV^\dagger U^\dagger = UU^\dagger = \mathbb{1} \tag{7}$$

These equations indeed show that the product operator UV is unitary. This property, moreover, was foreseeable: when two transformations conserve the scalar product, so does the successive application of these two transformations.

(*iii*) In the ordinary three-dimensional space of real vectors, we are familiar with operators which conserve the norm and the scalar product: rotations, symmetry operations with respect to a point, to a plane, etc. In this case, where the space is real, these operators are said to be orthogonal. Unitary operators constitute the generalization of orthogonal operators to complex spaces (with an arbitrary number of dimensions).

1-b. Unitary operators and change of bases

α. Let $\{|v_i\rangle\}$ be an orthonormal basis of the state space $\mathcal{E}$, assumed to be discrete. Call $|\tilde{v}_i\rangle$ the transform of the vector $|v_i\rangle$ under the action of a unitary operator U:

$$|\tilde{v}_i\rangle = U|v_i\rangle \tag{8}$$

Since the operator U is unitary, we have:

$$\langle\tilde{v}_i|\tilde{v}_j\rangle = \langle v_i|v_j\rangle = \delta_{ij} \tag{9}$$

The $|\tilde{v}_i\rangle$ vectors are therefore orthonormal. Let us show that they constitute a basis of $\mathcal{E}$. To do so, consider an arbitrary vector $|\psi\rangle$ of $\mathcal{E}$. Since the set $\{|v_i\rangle\}$ constitutes a basis, the vector $U^\dagger|\psi\rangle$ can be expanded on the $|v_i\rangle$:

$$U^\dagger|\psi\rangle = \sum_i c_i|v_i\rangle \tag{10}$$

Applying the operator U to this equation, we obtain:

$$UU^\dagger|\psi\rangle = \sum_i c_i U|v_i\rangle \tag{11}$$

and, therefore:

$$|\psi\rangle = \sum_i c_i|\tilde{v}_i\rangle \tag{12}$$

This equation expresses the fact that any vector $|\psi\rangle$ can be expanded on the vectors $|\tilde{v}_i\rangle$, which therefore constitute a basis. Thus we can state the following result: a necessary condition for an operator U to be unitary is that the vectors of an orthonormal basis of $\mathcal{E}$, transformed by U, constitute another orthonormal basis.

β. Conversely, let us show that this condition is sufficient. By hypothesis, we then have:

$$\begin{aligned} &|\tilde{v}_i\rangle = U|v_i\rangle \\ &\langle\tilde{v}_i|\tilde{v}_j\rangle = \delta_{ij} \\ &\sum_i |\tilde{v}_i\rangle \langle\tilde{v}_i| = \mathbb{1} \end{aligned} \tag{13}$$

and therefore:

$$\langle\tilde{v}_j|U^\dagger = \langle\tilde{v}_j| \tag{14}$$

Let us calculate:

$$\begin{aligned} U^\dagger U|v_i\rangle = U^\dagger|\tilde{v}_i\rangle &= \sum_j |v_j\rangle \langle v_j|U^\dagger|\tilde{v}_i\rangle \\ &= \sum_j |v_j\rangle \langle\tilde{v}_j|\tilde{v}_i\rangle = \sum_j |v_j\rangle\delta_{ij} \\ &= |v_i\rangle \end{aligned} \tag{15}$$

Relation (15), which is valid for all i, expresses the fact that the operator $U^\dagger U$ is the identity operator. Let us show, in the same way, that $UU^\dagger = \mathbb{1}$. To do this, consider the action of $U^\dagger$ on a vector $|v_i\rangle$:

$$\begin{aligned} U^\dagger|v_i\rangle &= \sum_j |v_j\rangle \langle v_j|U^\dagger|v_i\rangle \\ &= \sum_j |v_j\rangle \langle\tilde{v}_j|v_i\rangle \end{aligned} \tag{16}$$

We then have:

$$\begin{aligned} UU^\dagger|v_i\rangle &= \sum_j U|v_j\rangle \langle\tilde{v}_j|v_i\rangle \\ &= \sum_j |\tilde{v}_j\rangle \langle\tilde{v}_j|v_i\rangle \\ &= |v_i\rangle \end{aligned} \tag{17}$$

We deduce from this that $UU^\dagger = \mathbb{1}$: the operator U is therefore unitary.

1-c. Unitary matrices

Let:

$$U_{ij} = \langle v_i|U|v_j\rangle \tag{18}$$

be the matrix elements of U. How can one see from the matrix representing U if this operator is unitary?

Relation (1) gives us:

$$\langle v_i|U^\dagger U|v_j\rangle = \sum_k \langle v_i|U^\dagger|v_k\rangle \langle v_k|U|v_j\rangle \tag{19}$$

that is:

$$\sum_k U^*_{ki}U_{kj} = \delta_{ij} \tag{20}$$

When a matrix is unitary, the sum of the products of the elements of one column and the complex conjugates of the elements of another column is

– zero if the two columns are different,

– equal to 1 if they are not.

Let us cite some examples in which this rule can be easily verified.

Examples:

(*i*) The matrix which represents a rotation through an angle θ about Oz, in ordinary three-dimensional space:

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{21}$$

(*ii*) The rotation matrix in the state space of a spin $\frac{1}{2}$ particle (*cf.* Chap. IX):

$$R^{(1/2)}(\alpha,\beta,\gamma) = \begin{pmatrix} \mathrm{e}^{-\frac{i}{2}(\alpha+\gamma)}\cos\dfrac{\beta}{2} & -\mathrm{e}^{\frac{i}{2}(\gamma-\alpha)}\sin\dfrac{\beta}{2} \\ \mathrm{e}^{\frac{i}{2}(\gamma-\alpha)}\sin\dfrac{\beta}{2} & \mathrm{e}^{\frac{i}{2}(\alpha+\gamma)}\cos\dfrac{\beta}{2} \end{pmatrix} \tag{22}$$

1-d. Eigenvalues and eigenvectors of a unitary operator

Let $|\psi_u\rangle$ be a normalized eigenvector of the unitary operator U with eigenvalue u:

$$U|\psi_u\rangle = u|\psi_u\rangle \tag{23}$$

The square of the norm of the vector $U|\psi_u\rangle$ is:

$$\langle\psi_u|U^\dagger U|\psi_u\rangle = u^*u\langle\psi_u|\psi_u\rangle = u^*u \tag{24}$$

Since the unitary operator conserves the norm, we have, necessarily, $u^*u = 1$. The eigenvalues of a unitary operator must therefore be complex numbers of modulus 1:

$$u = \mathrm{e}^{i\varphi_u} \quad \text{where} \quad \varphi_u \quad \text{is real} \tag{25}$$

Consider two eigenvectors $|\psi_u\rangle$ and $|\psi_{u'}\rangle$ of U; we then have:

$$\begin{aligned}\langle\psi_u|\psi_{u'}\rangle &= \langle\psi_u|U^\dagger U|\psi_{u'}\rangle = u^*u'\langle\psi_u|\psi_{u'}\rangle \\ &= \mathrm{e}^{i(\varphi_{u'}-\varphi_u)}\langle\psi_u|\psi_{u'}\rangle\end{aligned} \tag{26}$$

When the eigenvalues u and u' are different, we see from (26) that the scalar product $\langle\psi_u|\psi_{u'}\rangle$ must be zero: two eigenvectors of a unitary operator corresponding to different eigenvalues are orthogonal.

2. Unitary transformations of operators

We saw in § 1-b that a unitary operator U permits the construction, starting with one orthonormal basis $\{|v_i\rangle\}$ of $\mathcal{E}$, of another one, $\{|\tilde{v}_i\rangle\}$. In this section, we are going to define a transformation that acts, not on the vectors, but on the operators.

By definition, the transform $\tilde{A}$ of the operator A will be the operator which, in the $\{|\tilde{v}_i\rangle\}$ basis, has the same matrix elements as the operator A in the $\{|v_i\rangle\}$ basis:

$$\langle\tilde{v}_i|\tilde{A}|\tilde{v}_j\rangle = \langle v_i|A|v_j\rangle \tag{27}$$

Substituting (8) into this equation, we obtain:

$$\langle v_i|U^\dagger\tilde{A}U|v_j\rangle = \langle v_i|A|v_j\rangle \tag{28}$$

Since i and j are arbitrary, we have:

$$U^\dagger\tilde{A}U = A \tag{29}$$

or, multiplying this equation on the left by U and on the right by $U^\dagger$:

$$\tilde{A} = UAU^\dagger \tag{30}$$

Equation (30) can be taken to be the definition of the transform $\tilde{A}$ of the operator A by the unitary transformation U. In quantum mechanics, such transformations are often used: a first example is given in Complement $\mathrm{F_{II}}$ of this chapter (§ 2-a).

How can the eigenvectors of $\tilde{A}$ be obtained from those of A? Let us consider an eigenvector $|\varphi_a\rangle$ of A, with an eigenvalue a:

$$A|\varphi_a\rangle = a|\varphi_a\rangle \tag{31}$$

Let $|\tilde{\varphi}_a\rangle$ be the transform of $|\varphi_a\rangle$ by the operator U: $|\tilde{\varphi}_a\rangle = U|\varphi_a\rangle$. We then have:

$$\begin{aligned}\tilde{A}|\tilde{\varphi}_a\rangle &= (UAU^\dagger)U|\varphi_a\rangle = UA(U^\dagger U)|\varphi_a\rangle \\ &= UA|\varphi_a\rangle = aU|\varphi_a\rangle \\ &= a|\tilde{\varphi}_a\rangle\end{aligned} \tag{32}$$

$|\tilde{\varphi}_a\rangle$ is therefore an eigenvector of $\tilde{A}$, with eigenvalue a. This can be generalized to the following rule: the eigenvectors of the transform $\tilde{A}$ of A are the transforms $|\tilde{\varphi}_a\rangle$ of the eigenvectors $|\varphi_a\rangle$ of A; the eigenvalues are unchanged.

Comments:

(*i*) The adjoint of the transform $\tilde{A}$ of A by U is the transform of $A^\dagger$ by U:

$$(\tilde{A})^\dagger = (UAU^\dagger)^\dagger = UA^\dagger U^\dagger = \tilde{A}^\dagger \tag{33}$$

In particular, it follows from this relation that, if A is Hermitian, $\tilde{A}$ is also.

(*ii*) Similarly, we have:

$$(\tilde{A})^2 = UAU^\dagger UAU^\dagger = UAAU^\dagger = \tilde{A}^2$$

and, in general:

$$(\tilde{A})^n = \tilde{A}^n \tag{34}$$

Using definition (23) of Complement B_{II}, we can easily show that:

$$\tilde{F}(A) = F(\tilde{A}) \tag{35}$$

where $F(A)$ is a function of the operator A.

3. The infinitesimal unitary operator

Let $U(\varepsilon)$ be a unitary operator which depends on an infinitely small real quantity ε; by hypothesis, $U(\varepsilon) \longrightarrow \mathbb{1}$ when $\varepsilon \longrightarrow 0$. Expand $U(\varepsilon)$ in a power series in ε:

$$U(\varepsilon) = \mathbb{1} + \varepsilon\, G + ... \tag{36}$$

We then have:

$$U^\dagger(\varepsilon) = \mathbb{1} + \varepsilon\, G^\dagger + ... \tag{37}$$

and:

$$U(\varepsilon)\, U^\dagger(\varepsilon) = U^\dagger(\varepsilon)U(\varepsilon) = \mathbb{1} + \varepsilon(G + G^\dagger) + ... \tag{38}$$

Since $U(\varepsilon)$ is unitary, the first-order terms in ε on the right-hand side of (38) are zero; we therefore have:

$$G + G^\dagger = 0 \tag{39}$$

This relation expresses the fact that the operator G is anti-Hermitian. It is convenient to set:

$$F = iG \tag{40}$$

so as to obtain the equation:

$$F - F^\dagger = 0 \tag{41}$$

which states that F is Hermitian. An infinitesimal unitary operator can therefore be written in the form:

$$U(\varepsilon) = \mathbb{1} - i\varepsilon F \tag{42}$$

where F is a Hermitian operator.

Substituting (42) into (30), we obtain:

$$\tilde{A} = (\mathbb{1} - i\varepsilon F)A(\mathbb{1} + i\varepsilon F^\dagger) = (\mathbb{1} - i\varepsilon F)A(\mathbb{1} + i\varepsilon F) \tag{43}$$

and, therefore:

$$\tilde{A} - A = -i\varepsilon[F, A] \tag{44}$$

The variation of the operator A under the transformation U is, to first order in ε, proportional to the commutator $[F, A]$.

Complement D_{II}

A more detailed study of the $\{|\mathrm{r}\rangle\}$ and $\{|\mathrm{p}\rangle\}$ representations

1 The $\{|\mathbf{r}\rangle\}$ representation 181

1-a The $\mathbf{R}$ operator and functions of $\mathbf{R}$ 181

1-b The $\mathbf{P}$ operator and functions of $\mathbf{P}$ 182

1-c The Schrödinger equation in the $\{|\mathbf{r}\rangle\}$ representation 183

2 The $\{|\mathbf{p}\rangle\}$ representation 184

2-a The $\mathbf{P}$ operator and functions of $\mathbf{P}$ 184

2-b The $\mathbf{R}$ operator and functions of $\mathbf{R}$ 184

2-c The Schrödinger equation in the $\{|\mathbf{p}\rangle\}$ representation 184

1. The $\{|\mathrm{r}\rangle\}$ representation

1-a. The $\mathbf{R}$ operator and functions of $\mathbf{R}$

Let us calculate the matrix elements, in the $\{|\mathbf{r}\rangle\}$ representation, of the X, Y, Z operators. Using formula (E-2-c) of Chapter II and the orthogonality relations of the kets $|\mathbf{r}\rangle$, we immediately obtain:

$$\begin{aligned}
&\langle \mathbf{r}|X|\mathbf{r}'\rangle = x\ \delta(\mathbf{r}-\mathbf{r}')\\
&\langle \mathbf{r}|Y|\mathbf{r}'\rangle = y\ \delta(\mathbf{r}-\mathbf{r}')\\
&\langle \mathbf{r}|Z|\mathbf{r}'\rangle = z\ \delta(\mathbf{r}-\mathbf{r}')
\end{aligned} \tag{1}$$

These three equations can be condensed into one:

$$\langle \mathbf{r}|\mathbf{R}|\mathbf{r}'\rangle = \mathbf{r}\ \delta(\mathbf{r}-\mathbf{r}') \tag{2}$$

The matrix elements, in the $\{|\mathbf{r}\rangle\}$ representation, of a function $F(\mathbf{R})$ are also very simple [*cf.* equation (27) of Complement B_{II}]:

$$\langle \mathbf{r}|F(\mathbf{R})|\mathbf{r}'\rangle = F(\mathbf{r})\ \delta(\mathbf{r}-\mathbf{r}') \tag{3}$$

1-b. The P operator and functions of P

Let us calculate the matrix element $\langle \mathbf{r}|P_x|\mathbf{r}'\rangle$:

$$\begin{aligned}\langle \mathbf{r}|P_x|\mathbf{r}'\rangle &= \int \mathrm{d}^3p\ \langle \mathbf{r}|P_x|\mathbf{p}\rangle\ \langle \mathbf{p}|\mathbf{r}'\rangle \\ &= \int \mathrm{d}^3p\ p_x\langle \mathbf{r}|\mathbf{p}\rangle\ \langle \mathbf{p}|\mathbf{r}'\rangle \\ &= (2\pi\hbar)^{-3}\int \mathrm{d}^3p\ p_x\ \mathrm{e}^{\frac{i}{\hbar}\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')} \\ &= \left[\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty}\mathrm{d}p_x\ p_x\ \mathrm{e}^{\frac{i}{\hbar}p_x(x-x')}\right] \\ &\quad\times\left[\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty}\mathrm{d}p_y\ \mathrm{e}^{\frac{i}{\hbar}p_y(y-y')}\right] \\ &\quad\times\left[\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty}\mathrm{d}p_z\ \mathrm{e}^{\frac{i}{\hbar}p_z(z-z')}\right]\end{aligned} \tag{4}$$

From this, it follows that, using the integral form of the "delta function" and its derivative [*cf.* Appendix II, equations (34) and (53)]:

$$\langle \mathbf{r}|P_x|\mathbf{r}'\rangle = \frac{\hbar}{i}\ \delta'(x-x')\ \delta(y-y')\ \delta(z-z') \tag{5}$$

The matrix elements of the other components of **P** could be obtained in an analogous fashion.

Let us verify that the action of P_x in the $\{|\mathbf{r}\rangle\}$ representation can indeed be derived from formula (5). To do so, let us calculate:

$$\langle \mathbf{r}|P_x|\psi\rangle = \int \mathrm{d}^3r'\ \langle \mathbf{r}|P_x|\mathbf{r}'\rangle\ \langle \mathbf{r}'|\psi\rangle \tag{6}$$

From (5):

$$\langle \mathbf{r}|P_x|\psi\rangle = \frac{\hbar}{i}\int \delta'(x-x')\mathrm{d}x'\int \delta(y-y')\ \mathrm{d}y'\int \delta(z-z')\ \psi(x',y',z')\ \mathrm{d}z' \tag{7}$$

Using the relation (48) of Appendix II:

$$\int \delta'(-u)\ f(u)\ \mathrm{d}u = -\int \delta'(u)\ f(u)\ \mathrm{d}u = f'(0) \tag{8}$$

and taking $u = x' - x$, we obtain:

$$\langle \mathbf{r}|P_x|\psi\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x,y,z) \tag{9}$$

which is indeed equation (E-26) of Chapter II.

What is the value of the matrix element $\langle \mathbf{r}|G(\mathbf{P})|\mathbf{r}'\rangle$ of a function $G(\mathbf{P})$ of the $\mathbf{P}$ operator? An analogous calculation gives us:

$$\begin{aligned}\langle \mathbf{r}|G(\mathbf{P})|\mathbf{r}'\rangle &= \int \mathrm{d}^3p\ \langle \mathbf{r}|G(\mathbf{P})|\mathbf{p}\rangle\,\langle \mathbf{p}|\mathbf{r}'\rangle \\ &= (2\pi\hbar)^{-3}\int \mathrm{d}^3p\ G(\mathbf{p})\ \mathrm{e}^{i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} \\ &= (2\pi\hbar)^{-3/2}\ \tilde{G}(\mathbf{r}-\mathbf{r}') \end{aligned} \tag{10}$$

where $\tilde{G}(\mathbf{r})$ is the inverse Fourier transform of the function $G(\mathbf{p})$:

$$\tilde{G}(\mathbf{r}) = (2\pi\hbar)^{-3/2}\int \mathrm{d}^3p\ \mathrm{e}^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}\ G(\mathbf{p}) \tag{11}$$

1-c. The Schrödinger equation in the $\{|\mathbf{r}\rangle\}$ representation

In Chapter III, we shall introduce the Schrödinger equation, which is of fundamental importance in quantum mechanics:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{12}$$

where H is the Hamiltonian operator, which we shall define in that chapter. For a (spinless) particle in a scalar potential $V(\mathbf{r})$ [*cf.* equation (B-42) of Chapter III]:

$$H = \frac{1}{2m}\mathbf{P}^2 + V(\mathbf{R}) \tag{13}$$

Let us write this equation in the $\{|\mathbf{r}\rangle\}$ representation, that is, using the wave function $\psi(\mathbf{r},t)$, defined by:

$$\psi(\mathbf{r},t) = \langle \mathbf{r}|\psi(t)\rangle \tag{14}$$

Projecting (12) onto $|\mathbf{r}\rangle$, in the case where H is given by formula (13), we obtain:

$$i\hbar\frac{\partial}{\partial t}\langle \mathbf{r}|\psi(t)\rangle = \frac{1}{2m}\langle \mathbf{r}|\mathbf{P}^2|\psi(t)\rangle + \langle \mathbf{r}|V(\mathbf{R})|\psi(t)\rangle \tag{15}$$

The quantities involved in this equation can be expressed in terms of $\psi(\mathbf{r},t)$, since:

$$\frac{\partial}{\partial t}\langle \mathbf{r}|\psi(t)\rangle = \frac{\partial}{\partial t}\psi(\mathbf{r},t) \tag{16}$$

$$\langle \mathbf{r}|V(\mathbf{R})|\psi(t)\rangle = V(\mathbf{r})\,\psi(\mathbf{r},t) \tag{17}$$

The matrix element $\langle \mathbf{r}|\mathbf{P}^2|\psi\rangle$ can be calculated by using the fact that $\mathbf{P}$ acts like $\dfrac{\hbar}{i}\boldsymbol{\nabla}$ in the $\{|\mathbf{r}\rangle\}$ representation:

$$\begin{aligned}\langle \mathbf{r}|\mathbf{P}^2|\psi(t)\rangle &= \langle \mathbf{r}|(P_x^2+P_y^2+P_z^2)|\psi(t)\rangle \\ &= -\hbar^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi(x,y,z,t) \\ &= -\hbar^2\Delta\psi(\mathbf{r},t)\end{aligned} \tag{18}$$

The Schrödinger equation then becomes:

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V(\mathbf{r})\right]\psi(\mathbf{r},t)} \tag{19}$$

This is indeed the wave equation introduced in Chapter I (§ B-2).

2. The $\{|\mathbf{p}\rangle\}$ representation

2-a. The P operator and functions of P

We obtain without difficulty formulas analogous to (2) and (3):

$$\langle\mathbf{p}|\mathbf{P}|\mathbf{p}'\rangle=\mathbf{p}\;\delta(\mathbf{p}-\mathbf{p}') \tag{20}$$

$$\langle\mathbf{p}|G(\mathbf{P})|\mathbf{p}'\rangle=G(\mathbf{p})\;\delta(\mathbf{p}-\mathbf{p}') \tag{21}$$

2-b. The R operator and functions of R

Arguments analogous to those of § 1 yield the formulas corresponding to (5) and (10):

$$\langle\mathbf{p}|X|\mathbf{p}'\rangle=i\hbar\,\delta'(p_x-p'_x)\,\delta(p_y-p'_y)\,\delta(p_z-p'_z) \tag{22}$$

and:

$$\langle\mathbf{p}|F(\mathbf{R})|\mathbf{p}'\rangle=(2\pi\hbar)^{-3/2}\;\bar{F}(\mathbf{p}-\mathbf{p}') \tag{23}$$

with:

$$\bar{F}(\mathbf{p})=(2\pi\hbar)^{-3/2}\int \mathrm{d}^3r\,\mathrm{e}^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}F(\mathbf{r}) \tag{24}$$

2-c. The Schrödinger equation in the $\{|\mathbf{p}\rangle\}$ representation

Let us introduce the "wave function in the $\{|\mathbf{p}\rangle\}$ representation" by:

$$\overline{\psi}(\mathbf{p},t)=\langle\mathbf{p}|\psi(t)\rangle \tag{25}$$

Using (12), we shall look for the equation giving the time evolution of $\overline{\psi}(\mathbf{p},t)$. Projecting (12) onto the ket $\{|\mathbf{p}\rangle\}$, we obtain:

$$i\hbar\frac{\partial}{\partial t}\langle\mathbf{p}|\psi(t)\rangle=\frac{1}{2m}\langle\mathbf{p}|\mathbf{P}^2|\psi(t)\rangle+\langle\mathbf{p}|V(\mathbf{R})|\psi(t)\rangle \tag{26}$$

Now we have:

$$\frac{\partial}{\partial t}\langle\mathbf{p}|\psi(t)\rangle=\frac{\partial}{\partial t}\overline{\psi}(\mathbf{p},t) \tag{27}$$

$$\langle\mathbf{p}|\mathbf{P}^2|\psi(t)\rangle=\mathbf{p}^2\,\overline{\psi}(\mathbf{p},t) \tag{28}$$

The quantity which remains to be calculated is:

$$\langle \mathbf{p}|V(\mathbf{R})|\psi(t)\rangle = \int \mathrm{d}^3p' \langle \mathbf{p}|V(\mathbf{R})|\mathbf{p}'\rangle \langle \mathbf{p}'|\psi(t)\rangle \tag{29}$$

Using (23), we find:

$$\langle \mathbf{p}|V(\mathbf{R})|\psi(t)\rangle = (2\pi\hbar)^{-3/2} \int \mathrm{d}^3p' \, \overline{V}(\mathbf{p}-\mathbf{p}') \, \overline{\psi}(\mathbf{p}',t) \tag{30}$$

where $\overline{V}(\mathbf{p})$ is the Fourier transform of $V(\mathbf{r})$:

$$\overline{V}(\mathbf{p}) = (2\pi\hbar)^{-3/2} \int \mathrm{d}^3r \, \mathrm{e}^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} \, V(\mathbf{r}) \tag{31}$$

The Schrödinger equation in the $\{|\mathbf{p}\rangle\}$ representation is therefore written:

$$\boxed{i\hbar\frac{\partial}{\partial t}\,\overline{\psi}(\mathbf{p},t) = \frac{\mathbf{p}^2}{2m}\overline{\psi}(\mathbf{p},t) + (2\pi\hbar)^{-3/2} \int \mathrm{d}^3p' \, \overline{V}(\mathbf{p}-\mathbf{p}') \, \overline{\psi}(\mathbf{p}',t)} \tag{32}$$

Comment:

Since $\overline{\psi}(\mathbf{p},t)$ is the Fourier transform of $\psi(\mathbf{r},t)$ [*cf.* formula (E-18) of Chapter II], it would have been possible to find equation (32) by taking the Fourier transforms of both sides of equation (19).

Complement E_{II}

Some general properties of two observables, Q and P, whose commutator is equal to $i\hbar$

1 The operator $S(\lambda)$: definition, properties 187

2 Eigenvalues and eigenvectors of Q 188

2-a Spectrum of Q 188

2-b Degree of degeneracy 188

2-c Eigenvectors 189

3 The $\{|\mathbf{q}\rangle\}$ representation 189

3-a The action of $\boldsymbol{Q}$ in the $\{|\boldsymbol{q}\rangle\}$ representation 189

3-b The action of $S(\lambda)$ in the $\{|\boldsymbol{q}\rangle\}$ representation; the translation operator 190

3-c The action of $\boldsymbol{P}$ in the $\{|\boldsymbol{q}\rangle\}$ representation 190

4 The $\{|\boldsymbol{p}\rangle\}$ representation. The symmetric nature of the $\boldsymbol{P}$ and $\boldsymbol{Q}$ observables 190

In quantum mechanics, one often encounters two operators whose commutator is equal to $i\hbar$. This is the case, for example, when these two operators correspond to the two classical conjugate quantities q_i and p_i (q_i, the coordinate in a system of orthonormal axes, and the conjugate momentum $p_i = \dfrac{\partial \mathcal{L}}{\partial q_i}$ *cf.* § 3-a of Appendix III). In quantum mechanics, one associates with q_i and p_i operators Q_i and P_i that satisfy the relation:

$$[Q_i, P_i] = i\hbar \tag{1}$$

In § E of Chapter II, we encountered such operators: X and P_x. In this complement, we shall take a more general point of view and show that it is possible to establish a whole series of important properties relative to two observables P and Q whose commutator is equal to $i\hbar$. All these properties are just consequences of the commutation relation (1).

1. The operator $S(\lambda)$: definition, properties

We shall consider two observables P and Q, satisfying the relation:

$$[Q, P] = i\hbar \tag{2}$$

and we shall define the operator $S(\lambda)$, which depends on the real parameter λ, by:

$$S(\lambda) = \mathrm{e}^{-i\lambda P/\hbar} \tag{3}$$

This operator is unitary; it is easy to verify the relations:

$$S^{\dagger}(\lambda) = S^{-1}(\lambda) = S(-\lambda) \tag{4}$$

Let us calculate the commutator $[Q, S(\lambda)]$. We can apply formula (51) of Complement B_{II}, since $[Q, P] = i\hbar$ commutes with Q and P:

$$[Q, S(\lambda)] = i\hbar \left(-\frac{i\lambda}{\hbar} \right) e^{-i\lambda P/\hbar} = \lambda \, S(\lambda) \tag{5}$$

This relation can also be written:

$$Q \, S(\lambda) = S(\lambda)[Q + \lambda] \tag{6}$$

Finally, note that:

$$S(\lambda) \, S(\mu) = S(\lambda + \mu) \tag{7}$$

2. Eigenvalues and eigenvectors of Q

2-a. Spectrum of Q

Assume that Q has a non-zero eigenvector $|q\rangle$, with eigenvalue q:

$$Q|q\rangle = q|q\rangle \tag{8}$$

Apply equation (6) to the vector $|q\rangle$. This yields:

$$\begin{aligned} Q \, S(\lambda)|q\rangle &= S(\lambda)(Q + \lambda)|q\rangle \\ &= S(\lambda)(q + \lambda)|q\rangle = (q + \lambda)S(\lambda)|q\rangle \end{aligned} \tag{9}$$

This equation expresses the fact that $S(\lambda)|q\rangle$ is another non-zero eigenvector of Q, with an eigenvalue of $(q+\lambda)$ ($S(\lambda)|q\rangle$ is non-zero because $S(\lambda)$ is unitary). Thus, starting with an eigenvector of Q, one can, by applying $S(\lambda)$, construct another eigenvector of λ, with any real eigenvalue (λ can indeed take on any real value). The spectrum of Q is therefore a continuous spectrum, composed of all possible values on the real axis[1].

2-b. Degree of degeneracy

From now on, we shall assume, for simplicity, that the eigenvalue q of Q is non-degenerate (the results which we shall derive can be generalized to the case where q is degenerate). Let us show that if q is non-degenerate, all the other eigenvalues of Q are also non-degenerate. Let us assume, for example, that the eigenvalue $q + \lambda$ is two-fold degenerate, and we shall show that we arrive at a contradiction. There would then exist two orthogonal eigenvectors, $|q + \lambda, \alpha\rangle$ and $|q + \lambda, \beta\rangle$, corresponding to the eigenvalue $q + \lambda$:

$$\langle q + \lambda, \beta | q + \lambda, \alpha \rangle = 0 \tag{10}$$

[1] This shows that in a space $\mathcal{E}$ of finite dimension N, there are no observables Q and P, whose commutator is equal to $i\hbar$. The number of eigenvalues of Q could not be simultaneously less than or equal to N and infinite.

This result can be derived directly, moreover, by taking the trace of relation (2): $\mathrm{Tr}\,\{QP\} - \mathrm{Tr}\,\{PQ\} = \mathrm{Tr}\, i\hbar$. When N is finite, the traces on the left-hand side of this equation exist: they are finite and equal numbers [*cf.* Complement B_{II}, formula (7a)]. The equation becomes $0 = \mathrm{Tr}\,\{i\hbar\} = Ni\hbar$, which is impossible.

Consider the two vectors $S(-\lambda)|q+\lambda,\alpha\rangle$ and $S(-\lambda)|q+\lambda,\beta\rangle$. They are, according to (9), two eigenvectors of Q, with an eigenvalue of $q+\lambda-\lambda=q$. They are not collinear, since they are orthogonal; their scalar product can be written, using the fact that $S(\lambda)$ is unitary:

$$\langle q+\lambda,\beta|S^{\dagger}(-\lambda)S(-\lambda)|q+\lambda,\alpha\rangle=\langle q+\lambda,\beta|q+\lambda,\alpha\rangle=0 \tag{11}$$

We reach the conclusion that q is at least two-fold degenerate, which is contrary to the initial hypothesis. Consequently, all the eigenvalues of Q must have the same degree of degeneracy.

2-c. Eigenvectors

We shall fix the relative phases of the different eigenvectors of Q with respect to the eigenvector $|0\rangle$, of eigenvalue 0, by setting:

$$|q\rangle=S(q)|0\rangle \tag{12}$$

Applying $S(\lambda)$ to both sides of (12) and using (7), we obtain:

$$S(\lambda)|q\rangle=S(\lambda)S(q)|0\rangle=S(\lambda+q)|0\rangle=|q+\lambda\rangle \tag{13}$$

The adjoint expression of (13) is written:

$$\langle q|S^{\dagger}(\lambda)=\langle q+\lambda| \tag{14}$$

or, using (4) and replacing λ by $-\lambda$:

$$\langle q|S(\lambda)=\langle q-\lambda| \tag{15}$$

3. The $\{|q\rangle\}$ representation

Since Q is an observable, the set of its eigenvectors $\{|q\rangle\}$ constitutes a basis of $\mathcal{E}$. It is possible to characterize each ket by its "wave function in the $\{|q\rangle\}$ representation":

$$\psi(q)=\langle q|\psi\rangle \tag{16}$$

3-a. The action of Q in the $\{|q\rangle\}$ representation

Let us calculate, in the $\{|q\rangle\}$ representation, the wave function associated with the ket $Q|\psi\rangle$. It is written:

$$\langle q|Q|\psi\rangle=q\langle q|\psi\rangle=q\psi(q) \tag{17}$$

[using (8) and the fact that Q is Hermitian]. The action of Q in the $\{|q\rangle\}$ representation is therefore simply a multiplication by q.

3-b. The action of $S(\lambda)$ in the $\{|q\rangle\}$ representation; the translation operator

The wave function in the $\{|q\rangle\}$ representation associated with the ket $S(\lambda)|\psi\rangle$ is written [formula (15)]:

$$\langle q|S(\lambda)|\psi\rangle = \langle q-\lambda|\psi\rangle = \psi(q-\lambda) \tag{18}$$

The action of the operator $S(\lambda)$ in the $\{|q\rangle\}$ representation is therefore a translation of the wave function over a distance λ parallel to the q-axis[2]. For this reason, $S(\lambda)$ is called the *translation operator*.

3-c. The action of P in the $\{|q\rangle\}$ representation

When ε is an infinitely small quantity, we have:

$$S(-\varepsilon) = \mathrm{e}^{i\varepsilon P/\hbar} = \mathbb{1} + i\frac{\varepsilon}{\hbar}P + O(\varepsilon^2) \tag{19}$$

Consequently:

$$\langle q|S(-\varepsilon)|\psi\rangle = \psi(q) + i\frac{\varepsilon}{\hbar}\langle q|P|\psi\rangle + O(\varepsilon^2) \tag{20}$$

On the other hand, equation (18) yields:

$$\langle q|S(-\varepsilon)|\psi\rangle = \psi(q+\varepsilon) \tag{21}$$

Comparison of (20) and (21) shows that:

$$\psi(q+\varepsilon) = \psi(q) + i\frac{\varepsilon}{\hbar}\langle q|P|\psi\rangle + O(\varepsilon^2) \tag{22}$$

It follows that:

$$\begin{aligned}\langle q|P|\psi\rangle &= \frac{\hbar}{i}\,\mathrm{Lim}_{\varepsilon\to 0}\frac{\psi(q+\varepsilon)-\psi(q)}{\varepsilon} \\ &= \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}q}\psi(q)\end{aligned} \tag{23}$$

The action of P in the $\{|q\rangle\}$ representation is therefore that of $\frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}q}$. Equation (E-26) of Chapter II is thus generalized.

4. The $\{|p\rangle\}$ representation. The symmetric nature of the P and Q observables

Relation (23) enables us to obtain easily the wave function $v_p(q)$ associated, in the $\{|q\rangle\}$ representation, with the eigenvector $|p\rangle$ of P with an eigenvalue of p:

$$v_p(q) = \langle q|p\rangle = (2\pi\hbar)^{-1/2}\,\mathrm{e}^{\frac{i}{\hbar}pq} \tag{24}$$

[2]The function $f(x-a)$ is the function which, at the point $x = x_0 + a$, takes on the value $f(x_0)$. It is therefore the function obtained from $f(x)$ by a translation of $+a$.

We can therefore write:

$$|p\rangle = (2\pi\hbar)^{-1/2} \int_{-\infty}^{+\infty} \mathrm{d}q \ \mathrm{e}^{\frac{i}{\hbar}pq} |q\rangle \tag{25}$$

A ket $|\psi\rangle$ can be defined by its "wave function in the $\{|p\rangle\}$ representation":

$$\overline{\psi}(p) = \langle p|\psi\rangle \tag{26}$$

Using the adjoint relation of (25), we obtain:

$$\overline{\psi}(p) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{+\infty} \mathrm{d}q \ \mathrm{e}^{-\frac{i}{\hbar}pq} \psi(q) \tag{27}$$

$\overline{\psi}(p)$ is therefore the Fourier transform of $\psi(q)$.

The action of the P operator in the $\{|p\rangle\}$ representation corresponds to a multiplication by p; that of the Q operator corresponds, as can easily be shown using (27), to the operation $i\hbar \dfrac{\mathrm{d}}{\mathrm{d}p}$.

Thus we obtain symmetrical results in the $\{|q\rangle\}$ and $\{|p\rangle\}$ representations. This is not surprising: in our hypotheses, it is possible to exchange the P and Q operators, simply changing the sign of the commutator in relation (2). Instead of introducing the operator $S(\lambda)$, we could therefore have considered $T(\lambda')$ defined by:

$$T(\lambda') = \mathrm{e}^{i\lambda' Q/\hbar} \tag{28}$$

and we could have developed the same arguments, replacing P by Q and i by $-i$ everywhere.

References:

Messiah (1.17), Vol. I, § VIII-6; Dirac (1.13), § 25; Merzbacher (1.16), Chap. 14, § 7.

Complement F_{II}

The parity operator

1 The parity operator 193
1-a Definition 193
1-b Simple properties of Π 194
1-c Eigensubspaces of Π 195
2 Even and odd operators 196
2-a Definitions 196
2-b Selection rules 197
2-c Examples 197
2-d Functions of operators 198
3 Eigenstates of an even observable B_+ 199
4 Application to an important special case 199

1. The parity operator

1-a. Definition

Consider a physical system whose state space is $\mathcal{E}_{\mathbf{r}}$. The parity operator Π is defined by its action on the basis vectors[1] $|\mathbf{r}\rangle$ of $\mathcal{E}_{\mathbf{r}}$:

$$\Pi|\mathbf{r}\rangle = |-\mathbf{r}\rangle \tag{1}$$

The matrix elements of Π are therefore, in the $\{|\mathbf{r}\rangle\}$ representation:

$$\langle\mathbf{r}|\Pi|\mathbf{r}'\rangle = \langle\mathbf{r}|-\mathbf{r}'\rangle = \delta(\mathbf{r}+\mathbf{r}') \tag{2}$$

Consider an arbitrary vector $|\psi\rangle$ of $\mathcal{E}_{\mathbf{r}}$:

$$|\psi\rangle = \int \mathrm{d}^3r\, \psi(\mathbf{r})|\mathbf{r}\rangle \tag{3}$$

If the variable change $\mathbf{r}' = -\mathbf{r}$ is performed, $|\psi\rangle$ can be written:

$$|\psi\rangle = \int \mathrm{d}^3r'\, \psi(-\mathbf{r}')|-\mathbf{r}'\rangle \tag{4}$$

Now calculate $\Pi|\psi\rangle$; we obtain:

$$\Pi|\psi\rangle = \int \mathrm{d}^3r'\, \psi(-\mathbf{r}')|\mathbf{r}'\rangle \tag{5}$$

[1] Care must be taken not to confuse $|-\mathbf{r}_0\rangle$ and $-|\mathbf{r}_0\rangle$. The former is an eigenvector of $\mathbf{R}$, with eigenvalue $-\mathbf{r}_0$ and wavefunction $\xi_{-\mathbf{r}_0}(\mathbf{r}) = \delta(\mathbf{r}+\mathbf{r}_0)$. The latter is an eigenvector of $\mathbf{R}$ with eigenvalue $\mathbf{r}_0$ and wavefunction $-\xi_{\mathbf{r}_0}(\mathbf{r}) = -\delta(\mathbf{r}-\mathbf{r}_0)$.

Comparison of (3) and (5) shows that the action of Π in the $\{|\mathbf{r}\rangle\}$ representation is to change $\mathbf{r}$ to $-\mathbf{r}$:

$$\langle \mathbf{r}|\Pi|\psi\rangle = \psi(-\mathbf{r}) \tag{6}$$

Now let us consider a physical system $\mathscr{S}$ whose state vector is $|\psi\rangle$; $\Pi|\psi\rangle$ describes the physical system obtained from $\mathscr{S}$ by reflection through the origin of the axes.

1-b. Simple properties of Π

The operator Π^2 is the identity operator. From (1) we have:

$$\Pi^2|\mathbf{r}\rangle = \Pi(\Pi|\mathbf{r}\rangle) = \Pi|-\mathbf{r}\rangle = |\mathbf{r}\rangle \tag{7}$$

that is, since the kets $|\mathbf{r}\rangle$ form a basis of $\mathcal{E}_{\mathbf{r}}$:

$$\Pi^2 = \mathbb{1} \tag{8a}$$

or:

$$\Pi = \Pi^{-1} \tag{8b}$$

It is easy to show by recurrence that the operator Π^n is:

$$\begin{cases} - \text{ equal to } \mathbb{1} \text{ when } n \text{ is even} \\ - \text{ equal to } \Pi \text{ when } n \text{ is odd} \end{cases}$$

We can rewrite (6) in the form:

$$\langle \mathbf{r}|\Pi|\psi\rangle = \langle -\mathbf{r}|\psi\rangle \tag{9}$$

Since this equation is valid for all $|\psi\rangle$, it can be deduced that:

$$\langle \mathbf{r}|\Pi = \langle -\mathbf{r}| \tag{10}$$

Moreover, the Hermitian conjugate expression of (1) is written:

$$\langle \mathbf{r}|\Pi^\dagger = \langle -\mathbf{r}| \tag{11}$$

Since the kets $|\mathbf{r}\rangle$ form a basis, it follows from (10) and (11) that Π is Hermitian:

$$\Pi^\dagger = \Pi \tag{12}$$

Combining this equation with (8b), we obtain:

$$\Pi^{-1} = \Pi^\dagger \tag{13}$$

Π is therefore unitary as well.

1-c. Eigensubspaces of Π

Let $|\varphi_\pi\rangle$ be an eigenvector of Π, with an eigenvalue of p_π. Applying (8a), we obtain:

$$|\varphi_\pi\rangle = \Pi^2|\varphi_\pi\rangle = p_\pi^2|\varphi_\pi\rangle \tag{14}$$

We therefore have $p_\pi^2 = 1$: the eigenvalues of Π are limited to 1 and -1. Since the space $\mathcal{E}_\mathbf{r}$ is infinite-dimensional, we immediately see that these eigenvalues are degenerate. An eigenvector of Π with the eigenvalue $+1$ will be said to be even; an eigenvector with the eigenvalue -1, odd.

Consider the two operators P_+ and P_- defined by:

$$P_+ = \frac{1}{2}(\mathbb{1} + \Pi)$$

$$P_- = \frac{1}{2}(\mathbb{1} - \Pi) \tag{15}$$

These operators are Hermitian; using (8a), it is easy to show that:

$$P_+^2 = P_+$$

$$P_-^2 = P_- \tag{16}$$

P_+ and P_- are thus the projectors onto two subspaces of $\mathcal{E}_\mathbf{r}$, which we shall call $\mathcal{E}_+$ and $\mathcal{E}_-$. Let us calculate the products P_+P_- and P_-P_+; we obtain:

$$P_+P_- = \frac{1}{4}(\mathbb{1} + \Pi - \Pi - \Pi^2) = 0$$

$$P_-P_+ = \frac{1}{4}(\mathbb{1} - \Pi + \Pi - \Pi^2) = 0 \tag{17}$$

The two subspaces $\mathcal{E}_+$ and $\mathcal{E}_-$ are therefore orthogonal. Let us show that they are also supplementary. We see immediately from definition (15) that:

$$P_+ + P_- = \mathbb{1} \tag{18}$$

For all $|\psi\rangle \in \mathcal{E}$, we have, therefore:

$$|\psi\rangle = (P_+ + P_-)|\psi\rangle = |\psi_+\rangle + |\psi_-\rangle \tag{19}$$

with:

$$|\psi_+\rangle = P_+|\psi\rangle$$

$$|\psi_-\rangle = P_-|\psi\rangle \tag{20}$$

Let us calculate the products ΠP_+ and ΠP_-; we obtain:

$$\Pi P_+ = \frac{1}{2}\Pi(\mathbb{1} + \Pi) = \frac{1}{2}(\Pi + \mathbb{1}) = P_+$$

$$\Pi P_- = \frac{1}{2}\Pi(\mathbb{1} - \Pi) = \frac{1}{2}(\Pi - \mathbb{1}) = -P_- \tag{21}$$

These equations enable us to show that the vectors $|\psi_+\rangle$ and $|\psi_-\rangle$ introduced in (20) are even and odd, respectively:

$$\Pi|\psi_+\rangle = \Pi P_+|\psi\rangle = P_+|\psi\rangle = |\psi_+\rangle$$

$$\Pi|\psi_-\rangle = \Pi P_-|\psi\rangle = -P_-|\psi\rangle = -|\psi_-\rangle \tag{22}$$

The spaces $\mathcal{E}_+$ and $\mathcal{E}_-$ are therefore the eigensubspaces of Π, with the eigenvalues $+1$ and -1. In the $\{|\mathbf{r}\rangle\}$ representation, equations (22) can be written:

$$\langle\mathbf{r}|\psi_+\rangle = \psi_+(\mathbf{r}) = \langle\mathbf{r}|\Pi|\psi_+\rangle = \psi_+(-\mathbf{r})$$

$$\langle\mathbf{r}|\psi_-\rangle = \psi_-(\mathbf{r}) = -\langle\mathbf{r}|\Pi|\psi_-\rangle = -\psi_-(-\mathbf{r}) \tag{23}$$

The wave functions $\psi_+(\mathbf{r})$ and $\psi_-(\mathbf{r})$ are even and odd, respectively.

Relation (19) expresses the fact that any ket $|\psi\rangle$ of $\mathcal{E}_\mathbf{r}$ can be decomposed into a sum of two eigenvectors of Π, $|\psi_+\rangle$ and $|\psi_-\rangle$, belonging respectively to the even subspace $\mathcal{E}_+$ and the odd subspace $\mathcal{E}_-$. Therefore, Π is an observable.

2. Even and odd operators

2-a. Definitions

In § 2 of Complement C_{II}, we defined the concept of a unitary transformation of operators. In the case of Π [which is indeed unitary; see (13)], the transformed operator of an arbitrary operator B is written:

$$\tilde{B} = \Pi B \Pi \tag{24}$$

and satisfies the relation [*cf.* equation (27) of Complement C_{II}]:

$$\langle\mathbf{r}|\tilde{B}|\mathbf{r}'\rangle = \langle -\mathbf{r}|B|-\mathbf{r}'\rangle \tag{25}$$

The operator $\tilde{B}$ is said to be the parity transform of B.

In particular, if: $\tilde{B} = +B$ the operator B is said to be even

if: $\tilde{B} = -B$ the operator B is said to be odd.

An even operator B_+ therefore satisfies:

$$B_+ = \Pi B_+ \Pi \tag{26}$$

or, multiplying this equation on the left by Π and using (8a):

$$\Pi B_+ = B_+ \Pi \tag{27}$$

$$[\Pi, B_+] = 0 \tag{28}$$

An even operator is therefore an operator that commutes with Π. It can be seen, similarly, that an odd operator B_- is an operator that anticommutes with Π:

$$\Pi B_- + B_- \Pi = 0 \tag{29}$$

2-b. Selection rules

Let B_+ be an even operator. Let us calculate the matrix element $\langle\varphi|B_+|\psi\rangle$; by hypothesis, we have:

$$\langle\varphi|B_+|\psi\rangle = \langle\varphi|\Pi B_+\Pi|\psi\rangle = \langle\varphi'|B_+|\psi'\rangle \tag{30}$$

with:

$$|\varphi'\rangle = \Pi|\varphi\rangle$$
$$|\psi'\rangle = \Pi|\psi\rangle \tag{31}$$

If one of the two kets, $|\varphi\rangle$ and $|\psi\rangle$, is even and the other odd ($|\varphi'\rangle = \pm|\varphi\rangle$, $|\psi'\rangle = \mp|\psi\rangle$), relation (30) yields:

$$\langle\varphi|B_+|\psi\rangle = -\langle\varphi|B_+|\psi\rangle = 0 \tag{32}$$

Hence the rule: the matrix elements of an even operator are zero between vectors of opposite parity.

If, now, B_- is odd, relation (30) becomes:

$$\langle\varphi|B_-|\psi\rangle = -\langle\varphi'|B_-|\psi'\rangle \tag{33}$$

which is zero when $|\varphi\rangle$ and $|\psi\rangle$ are both either even or odd. Hence the rule: the matrix elements of an odd operator are zero between vectors of the same parity. In particular, the diagonal matrix element $\langle\psi|B_-|\psi\rangle$ (the mean value of B_- in the state $|\psi\rangle$; *cf.* Chapter III, § C-4) is zero if $|\psi\rangle$ has a definite parity.

2-c. Examples

α. *The X, Y, Z operators*

In this case, we have:

$$\begin{aligned}\Pi X|\mathbf{r}\rangle &= \Pi X|x,y,z\rangle = x\Pi|x,y,z\rangle \\ &= x|-x,-y,-z\rangle = x|-\mathbf{r}\rangle\end{aligned} \tag{34}$$

and:

$$\begin{aligned}X\Pi|\mathbf{r}\rangle &= X|-\mathbf{r}\rangle = X|-x,-y,-z\rangle \\ &= -x|-x,-y,-z\rangle = -x|-\mathbf{r}\rangle\end{aligned} \tag{35}$$

Adding these two equations together, we obtain:

$$(\Pi X + X\Pi)|\mathbf{r}\rangle = 0 \tag{36}$$

or, since the vectors $|\mathbf{r}\rangle$ form a basis:

$$\Pi X + X\Pi = 0 \tag{37}$$

X is therefore odd.

The proofs are the same for Y and Z; $\mathbf{R}$ is therefore an odd operator.

β. The P_x, P_y, P_z operators

Let us calculate the ket $\Pi|\mathbf{p}\rangle$; we obtain:

$$\begin{aligned}\Pi|\mathbf{p}\rangle &= (2\pi\hbar)^{-3/2}\int \mathrm{d}^3r\,\mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\Pi|\mathbf{r}\rangle \\ &= (2\pi\hbar)^{-3/2}\int \mathrm{d}^3r\,\mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}/\hbar}|-\mathbf{r}\rangle \\ &= (2\pi\hbar)^{-3/2}\int \mathrm{d}^3r'\,\mathrm{e}^{-i\mathbf{p}\cdot\mathbf{r}'/\hbar}|\mathbf{r}'\rangle \\ &= |-\mathbf{p}\rangle \end{aligned} \tag{38}$$

We then have, using an argument analogous to the one developed in α:

$$\begin{aligned}\Pi P_x|\mathbf{p}\rangle &= p_x|-\mathbf{p}\rangle \\ P_x\Pi|\mathbf{p}\rangle &= -p_x|-\mathbf{p}\rangle\end{aligned} \tag{39}$$

and:

$$\Pi P_x + P_x\Pi = 0 \tag{40}$$

The $\mathbf{P}$ operator is odd.

γ. The parity operator

Π obviously commutes with itself; it is an even operator.

2-d. Functions of operators

Let B_+ be an even operator. Using relation (8a), we obtain:

$$\Pi B_+^n\Pi = \underbrace{(\Pi B_+\Pi)(\Pi B_+\Pi)\ldots(\Pi B_+\Pi)}_{n\text{ factors}} = B_+^n \tag{41}$$

An even operator raised to the nth power is even. Hence, any operator $F(B_+)$ is even.

Let B_- be an odd operator; let us calculate the operator $\Pi B_-^n\Pi$:

$$\Pi B_-^n\Pi = \underbrace{(\Pi B_-\Pi)(\Pi B_-\Pi)\ldots(\Pi B_-\Pi)}_{n\text{ factors}} = (-1)^n(B_-)^n \tag{42}$$

An odd operator raised to the nth power is even if n is even, odd if n is odd. Consider an operator $F(B_-)$; this operator is even if the corresponding function $F(z)$ is even, odd if it is odd. In general, $F(B_-)$ has no definite parity.

3. Eigenstates of an even observable B_+

Let us consider an arbitrary even observable B_+ and an eigenvector $|\varphi_b\rangle$ of B_+ with an eigenvalue b. Since B_+ is even, it commutes with Π. Applying the theorems of § D-3-a of Chapter II, we obtain the following results:

α. If b is a non-degenerate eigenvalue, $|\varphi_b\rangle$ is necessarily an eigenvector of Π; it is therefore either an even or an odd vector. The mean value $\langle\varphi_b|B_-|\varphi_b\rangle$ of any odd observable B_-, such as $\mathbf{R}$, $\mathbf{P}$, etc..., is then zero.

β. If b is a degenerate eigenvalue corresponding to the eigensubspace $\mathcal{E}_b$, the vectors of $\mathcal{E}_b$ do not all necessarily have a definite parity. $\Pi|\varphi_b\rangle$ may be a vector which is non-collinear with $|\varphi_b\rangle$; it is nevertheless a vector which has the same eigenvalue b. Moreover, it is possible to find a basis of eigenvectors common to Π and B_+ in every subspace $\mathcal{E}_b$.

4. Application to an important special case

We shall often need to find the eigenstates of a Hamiltonian operator H, acting in $\mathcal{E}_\mathbf{r}$, of the form:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R}) \tag{43}$$

Since the $\mathbf{P}$ operator is odd, the $\mathbf{P}^2$ operator is even. When, in addition, the function $V(\mathbf{r})$ is even ($V(\mathbf{r}) = V(-\mathbf{r})$), the operator H is even. According to what we have just seen, it is then possible to look for the eigenstates of H among the even or odd states. This often simplifies the calculations considerably.

We have already encountered a certain number of cases where the Hamiltonian H is even: the square well, the infinite well (*cf.* Complement $\mathrm{H_I}$). We shall study others: the harmonic oscillator, the hydrogen atom, etc... It is easy to verify in all these special cases the properties which we have derived.

Comment:

If H is even, and if one of its eigenstates $|\varphi_h\rangle$ which has no definite parity (i.e. $\Pi|\varphi_h\rangle$ is non-collinear to $|\varphi_h\rangle$) has been found, it can be asserted that the corresponding eigenvalue is degenerate: since Π commutes with H, $\Pi|\varphi_h\rangle$ is an eigenvector of H with the same eigenvalue as $|\varphi_h\rangle$.

References and suggestions for further reading:

Schiff (1.18), § 29; Roman (2.3), § 5-3 d; Feynman I (6.3), Chap. 52; Sakurai (2.7), Chap. 3; articles by Morrison (2.28), Feinberg and Goldhaber (2.29), Wigner (2.30).

Complement G_{II}

An application of the properties of the tensor product: the two-dimensional infinite well

1 Definition; eigenstates 201
2 Study of the energy levels 202
2-a Ground state 202
2-b First excited states 203
2-c Systematic and accidental degeneracies 203

In Complement H_I (§ 2-c) we have already studied, in a one-dimensional problem, the stationary states of a particle placed in an infinite potential well. By using the concept of a tensor product (*cf.* Chap. II, § F), we shall be able to generalize this discussion to the case of a two-dimensional infinite well (the introduction of a third dimension would not involve any additional theoretical difficulty).

1. Definition; eigenstates

We shall consider a particle of mass m, restricted to a plane xOy, inside a "square box" of edge a: its potential energy $V(x, y)$ becomes infinite when one of its coordinates x or y leaves the interval $[0, a]$:

$$V(x,y) = V_\infty(x) + V_\infty(y) \tag{1}$$

with:

$$\begin{aligned} V_\infty(u) &= 0 \qquad \text{if} \quad 0 \leq u \leq a \\ &= +\infty \qquad \text{if} \quad u < 0 \qquad \text{or} \quad u > a \end{aligned} \tag{2}$$

The Hamiltonian of the quantum particle is then (Chap. III, § B-5):

$$H = \frac{1}{2m}(P_x^2 + P_y^2) + V_\infty(X) + V_\infty(Y) \tag{3}$$

which can be written:

$$H = H_x + H_y \tag{4}$$

with:

$$\begin{aligned} H_x &= \frac{1}{2m}P_x^2 + V_\infty(X) \\ H_y &= \frac{1}{2m}P_y^2 + V_\infty(Y) \end{aligned} \tag{5}$$

We thus find ourselves in the important special case mentioned in Chapter II (§ F-4-a-β), and we can consider the eigenstates of H in the form:

$$|\Phi\rangle = |\varphi\rangle_x \, |\varphi\rangle_y \tag{6}$$

with:

$$\begin{aligned} &H_x|\varphi\rangle_x = E_x|\varphi\rangle_x \quad ; \quad |\varphi\rangle_x \in \mathcal{E}_x \\ &H_y|\varphi\rangle_y = E_y|\varphi\rangle_y \quad ; \quad |\varphi\rangle_y \in \mathcal{E}_y \end{aligned} \tag{7}$$

We then have:

$$H|\Phi\rangle = E|\Phi\rangle$$

with:

$$E = E_x + E_y \tag{8}$$

We have therefore reduced a two-dimensional problem to a one-dimensional problem, which, moreover, has already been solved (*cf.* Complement H_I). Applying the results of this complement, and formulas (7) and (8), we therefore see that:

– the eigenvalues of H are of the form:

$$E_{n,p} = \frac{1}{2ma^2}(n^2 + p^2)\pi^2\hbar^2 \tag{9}$$

where n and p are positive integers.

– to these energies correspond eigenstates $|\Phi_{n,\,p}\rangle$ which can be written in the form of tensor products:

$$|\Phi_{n,\,p}\rangle = |\varphi_n\rangle_x \, |\varphi_p\rangle_y \tag{10}$$

whose normalized wave function is:

$$\begin{aligned} \Phi_{n,\,p}(x,y) &= \varphi_n(x)\,\varphi_p(y) \\ &= \frac{2}{a}\sin\frac{n\pi x}{a}\sin\frac{p\pi y}{a} \end{aligned} \tag{11}$$

It is easy to verify that these wave functions vanish at the edges of the "square box" (x or $y = 0$ or a), where the potential energy becomes infinite.

2. Study of the energy levels

2-a. Ground state

n and p are strictly positive integers[1]. The ground state is therefore obtained when $n = 1$, $p = 1$. Its energy is:

$$E_{1,\,1} = \frac{\pi^2\hbar^2}{ma^2} \tag{12}$$

This value is attained only for $n = p = 1$. The ground state is therefore not degenerate.

[1] The values $n = 0$ or $p = 0$ are excluded as they give null wave functions (therefore impossible to normalize).

2-b. First excited states

The first excited state is obtained either for $n = 1$ and $p = 2$ or for $n = 2$ and $p = 1$. Its energy is:

$$E_{1,2} = E_{2,1} = \frac{5}{2}\frac{\pi^2\hbar^2}{ma^2} \tag{13}$$

This state is two-fold degenerate, since $|\Phi_{1,2}\rangle$ and $|\Phi_{2,1}\rangle$ are independent.

The second excited state corresponds to $n = p = 2$; it is not degenerate, and its energy is:

$$E_{2,2} = 4\frac{\pi^2\hbar^2}{ma^2} \tag{14}$$

The third excited state corresponds to $n = 1$, $p = 3$ and $n = 3$, $p = 1$, etc.

2-c. Systematic and accidental degeneracies

The general observation can be made that all levels for which $n \neq p$ are degenerate, since:

$$E_{n,p} = E_{p,n} \tag{15}$$

This degeneracy is related to a symmetry of the problem. The square well under consideration is symmetric with respect to the first bisectrix of the xOy plane. This is expressed by the fact that the Hamiltonian H is invariant under the exchange[2]:

$$X \leftrightarrow Y$$

$$P_x \leftrightarrow P_y \tag{16}$$

If an eigenstate of H is known whose wave function is $\Phi(x, y)$, the state which corresponds to $\Phi'(x, y) = \Phi(y, x)$ is also an eigenstate of H with the same eigenvalue. Consequently, if the function $\Phi(x, y)$ is not symmetric with respect to x and y, the eigenvalue associated with it is necessarily degenerate. This is the origin of the degeneracy (15): for $n \neq p$, $\Phi_{n,p}(x, y)$ is not symmetric with respect to x and y [formula (11)]. This interpretation is corroborated by the fact that if the symmetry is destroyed by choosing a well whose widths along Ox and along Oy are different (being equal to a and b respectively), the corresponding degeneracy disappears, and formula (9) becomes:

$$E_{n,p} = \frac{\pi^2\hbar^2}{2m}\left(\frac{n^2}{a^2} + \frac{p^2}{b^2}\right), \tag{17}$$

which implies:

$$E_{p,n} \neq E_{n,p} \tag{18}$$

Such degeneracies, whose origin lies in a symmetry of the problem, are called *systematic degeneracies.*

[2] In the state space, an operator could be defined to correspond to a reflection about the first bisectrix. It could then be shown that, in the present case, this operator commutes with H.

Comment:

The other symmetries of the two-dimensional square well do not create systematic degeneracies because the eigenstates of H are all invariant with respect to them. For example, for arbitrary n and p, $\Phi_{n,\,p}(x,y)$ is simply multiplied by a phase factor if x is replaced by $(a-x)$ and y by $(a-y)$ (symmetry with respect to the center of the well).

Degeneracies may also arise which are not directly related to the symmetry of the problem. They are called *accidental degeneracies*. For example, in the case which we have discussed, it so happens that $E_{5,\,5} = E_{7,\,1}$ and $E_{7,\,4} = E_{8,\,1}$.

Complement H_{II}

Exercises

Dirac notation. Commutators. Eigenvectors and eigenvalues

1. $|\varphi_n\rangle$ are the eigenstates of a Hermitian operator H (H is, for example, the Hamiltonian of some physical system). Assume that the states $|\varphi_n\rangle$ form a discrete orthonormal basis. The operator $U(m, n)$ is defined by:

$$U(m, n) = |\varphi_m\rangle \langle\varphi_n|$$

a. Calculate the adjoint $U^\dagger(m, n)$ of $U(m, n)$.

b. Calculate the commutator $[H, U(m, n)]$.

c. Prove the relation:

$$U(m, n)U^\dagger(p, q) = \delta_{nq}U(m, p)$$

d. Calculate Tr $\{U(m, n)\}$, the trace of the operator $U(m, n)$.

e. Let A be an operator, with matrix elements $A_{mn} = \langle\varphi_m|A|\varphi_n\rangle$. Prove the relation:

$$A = \sum_{m, n} A_{mn}\, U(m, n)$$

f. Show that $A_{pq} = \mathrm{Tr}\{AU^\dagger(p, q)\}$.

2. In a two-dimensional vector space, consider the operator whose matrix, in an orthonormal basis $\{|1\rangle, |2\rangle\}$, is written:

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

a. Is σ_y Hermitian? Calculate its eigenvalues and eigenvectors (giving their normalized expansion in terms of the $\{|1\rangle, |2\rangle\}$ basis).

b. Calculate the matrices that represent the projectors onto these eigenvectors. Then verify that they satisfy the orthogonality and closure relations.

c. Same questions for the matrices:

$$M = \begin{pmatrix} 2 & i\sqrt{2} \\ -i\sqrt{2} & 3 \end{pmatrix}$$

and, in a three-dimensional space

$$L_y = \frac{\hbar}{2i} \begin{pmatrix} 0 & \sqrt{2} & 0 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{2} & 0 \end{pmatrix}$$

3. The state space of a certain physical system is three-dimensional. Let $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ be an orthonormal basis of this space. The kets $|\psi_0\rangle$ and $|\psi_1\rangle$ are defined by:

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}|u_1\rangle + \frac{i}{2}|u_2\rangle + \frac{1}{2}|u_3\rangle$$

$$|\psi_1\rangle = \frac{1}{\sqrt{3}}|u_1\rangle + \frac{i}{\sqrt{3}}|u_3\rangle$$

a. Are these kets normalized?

b. Calculate the matrices ρ_0 and ρ_1 representing, in the $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ basis, the projection operators onto the state $|\psi_0\rangle$ and onto the state $|\psi_1\rangle$. Verify that these matrices are Hermitian.

4. Let K be the operator defined by $K = |\varphi\rangle\langle\psi|$, where $|\varphi\rangle$ and $|\psi\rangle$ are two vectors of the state space.

a. Under what condition is K Hermitian?

b. Calculate K^2. Under what condition is K a projector ?

c. Show that K can always be written in the form $K = \lambda P_1 P_2$ where λ is a constant to be calculated and P_1 and P_2 are projectors.

5. Let P_1 be the orthogonal projector onto the subspace $\mathcal{E}_1$, P_2 the orthogonal projector onto the subspace $\mathcal{E}_2$. Show that, for the product P_1P_2 to be an orthogonal projector as well, it is necessary and sufficient that P_1 and P_2 commute. In this case, what is the subspace onto which P_1P_2 projects?

6. The σ_x matrix is defined by:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Prove the relation:

$$e^{i\alpha\sigma_x} = I\cos\alpha + i\sigma_x \sin\alpha$$

where I is the 2×2 unit matrix.

7. Establish, for the σ_y matrix given in exercise 2, a relation analogous to the one proved for σ_x in the preceding exercise. Generalize for all matrices of the form:

$$\sigma_u = \lambda\sigma_x + \mu\sigma_y$$

with:

$$\lambda^2 + \mu^2 = 1$$

Calculate the matrices representing $e^{2i\sigma_x}$, $(e^{i\sigma_x})^2$ and $e^{i(\sigma_x+\sigma_y)}$. Is $e^{2i\sigma_x}$ equal to $(e^{i\sigma_x})^2$? $e^{i(\sigma_x+\sigma_y)}$ to $e^{i\sigma_x}\ e^{i\sigma_y}$?

8. Consider the Hamiltonian H of a particle in a one-dimensional problem, defined by:

$$H = \frac{1}{2m}P^2 + V(X)$$

where X and P are the operators defined in § E of Chapter II, which satisfy the relation: $[X, P] = i\hbar$. The eigenvectors of H are denoted by $|\varphi_n\rangle$: $H|\varphi_n\rangle = E_n|\varphi_n\rangle$, where n is a discrete index.

a. Show that:

$$\langle\varphi_n|P|\varphi_{n'}\rangle = \alpha\langle\varphi_n|X|\varphi_{n'}\rangle$$

where α is a coefficient which depends on the difference between E_n and $E_{n'}$. Calculate α (hint: consider the commutator $[X, H]$).

b. From this, deduce, using the closure relation, the equation:

$$\sum_{n'}(E_n - E_{n'})^2|\langle\varphi_n|X|\varphi_{n'}\rangle|^2 = \frac{\hbar^2}{m^2}\langle\varphi_n|P^2|\varphi_n\rangle$$

9. Let H be the Hamiltonian operator of a physical system. Denote by $|\varphi_n\rangle$ the eigenvectors of H, with eigenvalues E_n:

$H|\varphi_n\rangle = E_n|\varphi_n\rangle$

a. For an arbitrary operator A, prove the relation:

$$\langle\varphi_n|[A, H]|\varphi_n\rangle = 0$$

b. Consider a one-dimensional problem, where the physical system is a particle of mass m with potential energy $V(X)$. In this case, H is written:

$$H = \frac{1}{2m}P^2 + V(X)$$

α. In terms of P, X and $V(X)$, find the commutators: $[H, P]$, $[H, X]$ and $[H, XP]$.

β. Show that the matrix element $\langle\varphi_n|P|\varphi_n\rangle$ (which we shall interpret in Chapter III as the mean value of the momentum in the state $|\varphi_n\rangle$) is zero.

γ. Establish a relation between $E_k = \langle\varphi_n|\frac{P^2}{2m}|\varphi_n\rangle$ (the mean value of the kinetic energy in the state $|\varphi_n\rangle$) and $\langle\varphi_n|X\frac{dV}{dX}|\varphi_n\rangle$. Since the mean value of the potential energy in the state $|\varphi_n\rangle$ is $\langle\varphi_n|V(x)|\varphi_n\rangle$, how is it related to the mean value of the kinetic energy when:

$$V(X) = V_0\,X^s$$

$(s = 2, 4, 6\ldots;\ V_0 > 0)$?

10. Using the relation $\langle x|p\rangle = (2\pi\hbar)^{-1/2}\mathrm{e}^{ipx/\hbar}$, find the expressions of $\langle x|XP|\psi\rangle$ and $\langle x|PX|\psi\rangle$ in terms of $\psi(x)$. Can these results be found directly by using the fact that in the $\{|x\rangle\}$ representation, P acts like $\dfrac{\hbar}{i}\dfrac{\mathrm{d}}{\mathrm{d}x}$?

Complete sets of commuting observables, C.S.C.O.

11. Consider a physical system whose three-dimensional state space is spanned by the orthonormal basis formed by the three kets $|u_1\rangle$, $|u_2\rangle$, $|u_3\rangle$. In the basis of these three vectors, taken in this order, the two operators H and B are defined by:

$$H = \hbar\omega_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad B = b \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

where ω_0 and b are real constants.

a. Are H and B Hermitian?

b. Show that H and B commute. Give a basis of eigenvectors common to H and B.

c. Of the sets of operators: $\{H\}$, $\{B\}$, $\{H, B\}$, $\{H^2, B\}$, which form a C.S.C.O.?

12. In the same state space as that of the preceding exercise, consider two operators L_z and S defined by:

$$L_z|u_1\rangle = |u_1\rangle \qquad L_z|u_2\rangle = 0 \qquad L_z|u_3\rangle = -|u_3\rangle$$
$$S|u_1\rangle = |u_3\rangle \qquad S|u_2\rangle = |u_2\rangle \qquad S|u_3\rangle = |u_1\rangle$$

a. Write the matrices that represent, in the $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ basis, the operators L_z, L_z^2, S, S^2. Are these operators observables?

b. Give the form of the most general matrix which represents an operator which commutes with L_z. Same question for L_z^2, then for S^2.

c. Do L_z^2 and S form a C.S.C.O.? Give a basis of common eigenvectors.

Solution of exercise 11

a. H and B are Hermitian because the matrices which correspond to them are symmetric and real.

b. $|u_1\rangle$ is an eigenvector common to H and B. We therefore have $HB|u_1\rangle = BH|u_1\rangle$. We see, then, that for H and B to commute, it is sufficient that the restrictions of these operators to the subspace $\mathcal{E}_2$, spanned by $|u_2\rangle$ and $|u_3\rangle$, commute. Now, in this subspace, the matrix representing H is equal to $-\hbar\omega_0 I$ (where I is the 2×2 unit matrix), which commutes with all 2×2 matrices. H and B therefore commute

(this result could, of course, be obtained by calculating directly the matrices HB and BH). The restriction of B to $\mathcal{E}_2$ is written:

$$P_{\mathcal{E}_2} B P_{\mathcal{E}_2} = b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The normalized eigenvectors of this 2×2 matrix are easy to obtain; they are:

$$|p_2\rangle = \frac{1}{\sqrt{2}}[|u_2\rangle + |u_3\rangle] \quad \text{(eigenvalue + } b\text{)}$$

$$|p_3\rangle = \frac{1}{\sqrt{2}}[|u_2\rangle - |u_3\rangle] \quad \text{(eigenvalue – } b\text{)}$$

These vectors are automatically eigenvectors of H since $\mathcal{E}_2$ is the eigensubspace of H corresponding to the eigenvalue $-\hbar\omega_0$. To summarize, the eigenvectors common to H and B are given by:

	eigenvalue of H	eigenvalue of B
$\|p_1\rangle = \|u_1\rangle$	$\hbar\omega_0$	b
$\|p_2\rangle = \frac{1}{\sqrt{2}}[\|u_2\rangle + \|u_3\rangle]$	$-\hbar\omega_0$	b
$\|p_3\rangle = \frac{1}{\sqrt{2}}[\|u_2\rangle - \|u_3\rangle]$	$-\hbar\omega_0$	$-b$

These vectors are the only (to within, of course, a phase factor) normalized eigenvectors common to H and B.

c. It can be seen from the above table that H has a two-fold degenerate eigenvalue; it is therefore not a C.S.C.O. Similarly, B also has a two-fold degenerate eigenvalue and is therefore not a C.S.C.O.: an eigenvector of B with the eigenvalue b can be $|p_1\rangle$, or $|p_2\rangle$, or $\frac{1}{\sqrt{3}}|u_1\rangle + \frac{1}{\sqrt{3}}|u_2\rangle + \frac{1}{\sqrt{3}}|u_3\rangle$, for example. On the other hand, the set of the two operators H and B does constitute a C.S.C.O. We see from the above table that no two vectors $|p_j\rangle$ have the same eigenvalues for both H and B. This is why, as has already been pointed out, the system of normalized eigenvectors common to H and B is unique (to within phase factors). Note that within the eigensubspace $\mathcal{E}_2$ of H associated with the eigenvalue $-\hbar\omega_0$, the eigenvalues of B are distinct (b and $-b$). Similarly, in the eigensubspace of B spanned by $|p_1\rangle$ and $|p_2\rangle$, the eigenvalues of H are distinct ($\hbar\omega_0$ and $-\hbar\omega_0$).

H^2 has three eigenvectors with the eigenvalue $\hbar^2\omega_0^2$, $|p_1\rangle$, $|p_2\rangle$ and $|p_3\rangle$. It is easy to see that H^2 and B do not constitute a C.S.C.O., since two linearly independent eigenvectors $|p_1\rangle$ and $|p_2\rangle$ correspond to the pair of eigenvalues $\{\hbar^2\omega_0^2, b\}$.

Solution of exercise 12

a. Let us use the rule for constructing the matrix of an operator: "in the nth column of the matrix, write the components of the operator transform of the nth basis vector". We obtain easily:

$$L_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$L_z^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad S^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

These matrices are symmetric and real, and therefore Hermitian. Since the space is finite-dimensional, they can be diagonalized and therefore represent observables.

b. Let M be an operator that commutes with L_z. M cannot (*cf.* Chap. II, § D-3-a) have any matrix elements between $|u_1\rangle$ and $|u_2\rangle$, or between $|u_2\rangle$ and $|u_3\rangle$, or between $|u_1\rangle$ and $|u_3\rangle$ (eigenvectors of L_z with different eigenvalues). The matrix which represents M is therefore necessarily diagonal, that is, of the form:

$$[M, L_z] = 0 \qquad \Longleftrightarrow \qquad M = \begin{pmatrix} m_{11} & 0 & 0 \\ 0 & m_{22} & 0 \\ 0 & 0 & m_{33} \end{pmatrix}$$

Let N be an operator that commutes with L_z^2. The matrix representing N can have elements between $|u_1\rangle$ and $|u_3\rangle$ (eigenvectors of L_z^2 with the same eigenvalue), but none between $|u_2\rangle$ and $|u_1\rangle$ or $|u_3\rangle$. N is therefore written:

$$[N, L_z^2] = 0 \qquad \Longleftrightarrow \qquad N = \begin{pmatrix} n_{11} & 0 & n_{13} \\ 0 & n_{22} & 0 \\ n_{31} & 0 & n_{33} \end{pmatrix}$$

It is therefore less restrictive to impose the condition that an operator commute with L_z^2 than with L_z: N is not necessarily a diagonal matrix. It can only be said that N does not mix the vectors of the subspace $\mathcal{F}_2$ spanned by $|u_1\rangle$ and $|u_3\rangle$ with those of the one-dimensional subspace spanned by $|u_2\rangle$. This property, moreover, appears very clearly if the matrix N' which represents the operator N is written in the $\{|u_1\rangle, |u_3\rangle, |u_2\rangle\}$ basis (changing the order of the basis vectors):

$$N' = \begin{pmatrix} n_{11} & n_{13} & 0 \\ n_{31} & n_{33} & 0 \\ 0 & 0 & n_{22} \end{pmatrix}$$

Finally, since S^2 is the identity operator, any 3×3 matrix commutes with S^2, and its most general form is:

$$[P, S^2] = 0 \qquad \Longleftrightarrow \qquad P = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}$$

c. $|u_2\rangle$ is an eigenvector common to L_z^2 and S. In the subspace $\mathcal{F}_2$ spanned by $|u_1\rangle$ and $|u_3\rangle$, L_z^2 and S are written:

$$P_{\mathcal{F}_2} L_z^2 P_{\mathcal{F}_2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$P_{\mathcal{F}_2} S P_{\mathcal{F}_2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The eigenvectors of the latter matrix are:

$$|q_2\rangle = \frac{1}{\sqrt{2}}[|u_1\rangle + |u_3\rangle]$$

$$|q_3\rangle = \frac{1}{\sqrt{2}}[|u_1\rangle - |u_3\rangle]$$

and the basis of eigenvectors common to L_z^2 and S is:

vector	eigenvalue of L_z^2	eigenvalue of S
$\|q_1\rangle = \|u_2\rangle$	0	1
$\|q_2\rangle = \frac{1}{\sqrt{2}}[\|u_1\rangle + \|u_3\rangle]$	1	1
$\|q_3\rangle = \frac{1}{\sqrt{2}}[\|u_1\rangle - \|u_3\rangle]$	1	-1

No two lines are alike in the table of eigenvalues of L_z^2 and S: these two operators therefore form a C.S.C.O. (this is not, however, the case for either one of them taken alone).

Chapter III

The postulates of quantum mechanics

A Introduction 214

B Statement of the postulates 215

B-1 Description of the state of a system 215

B-2 Description of physical quantities 216

B-3 The measurement of physical quantities 216

B-4 Time evolution of systems 223

B-5 Quantization rules 223

C The physical interpretation of the postulates concerning observables and their measurement 226

C-1 The quantization rules are consistent with the probabilistic interpretation of the wave function 226

C-2 Quantization of certain physical quantities 227

C-3 The measurement process 227

C-4 Mean value of an observable in a given state 228

C-5 The root mean square deviation 230

C-6 Compatibility of observables 232

D The physical implications of the Schrödinger equation . . . 237

D-1 General properties of the Schrödinger equation 237

D-2 The case of conservative systems 245

E The superposition principle and physical predictions 253

E-1 Probability amplitudes and interference effects 253

E-2 Case in which several states can be associated with the same measurement result 259

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

A. Introduction

In classical mechanics, the motion of any physical system is determined if the position $\mathbf{r}(x, y, z)$ and velocity $\mathbf{v}(\dot{x}, \dot{y}, \dot{z})$ of each of its points are known as a function of time. In general (Appendix III), to describe such a system, one introduces generalized coordinates $q_i(t)$ $(i = 1, 2, ..., N)$, whose derivatives with respect to time, $\dot{q}_i(t)$, are the generalized velocities. Specifying the $q_i(t)$ and $\dot{q}_i(t)$ enables us to calculate, at any given instant, the position and velocity of any point of the system. Using the Lagrangian $\mathcal{L}(q_i, \dot{q}_i, t)$, one defines the conjugate momentum p_i of each of the generalized coordinates q_i:

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \tag{A-1}$$

The $q_i(t)$ and $p_i(t)$ $(i = 1, 2, ..., N)$ are called the fundamental dynamical variables. All the physical quantities associated with the system (energy, angular momentum, etc.) can be expressed in terms of the fundamental dynamical variables. For example, the total energy of the system is given by the classical Hamiltonian $\mathcal{H}$ (q_i, p_i, t). The motion of the system can be studied by using either Lagrange's equations or the Hamilton-Jacobi canonical equations, which are written:

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial p_i} \tag{A-2a}$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial q_i} \tag{A-2b}$$

In the special case of a system consisting of a single physical point of mass m, the q_i are simply the three coordinates of this point, and the $\dot{q}_i$ are the components of its velocity $\mathbf{v}$. If the forces acting on this particle can be derived from a scalar potential $V(\mathbf{r}, t)$, the three conjugate momenta of its position $\mathbf{r}$ (that is, the components of its linear momentum $\mathbf{p}$) are equal to the components of its mechanical momentum $m\mathbf{v}$. The total energy is then written:

$$E = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}, t) \tag{A-3}$$

and the angular momentum with respect to the origin:

$$\boldsymbol{\mathcal{L}} = \mathbf{r} \times \mathbf{p} \tag{A-4}$$

Since $\mathcal{H}(\mathbf{r}, \mathbf{p}, t) = (\mathbf{p}^2/2m) + V(\mathbf{r}, t)$, the Hamilton-Jacobi equations (A-2) take on the well-known form:

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\mathbf{p}}{m} \tag{A-5a}$$

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\boldsymbol{\nabla} V \tag{A-5b}$$

The classical description of a physical system can therefore be summarized as follows:

(*i*) The state of the system at a fixed time t_0 is defined by specifying N generalized coordinates $q_i(t_0)$ and their N conjugate momenta $p_i(t_0)$.

(*ii*) The value, at a given time, of the various physical quantities is completely determined when the state of the system at this time is known: knowing the state of the system, one can predict with certainty the result of any measurement performed at time t_0.

(*iii*) The time evolution of the state of the system is given by the Hamilton-Jacobi equations. Since these are first-order differential equations, their solution $\{q_i(t), p_i(t)\}$ is unique if the value of these functions at a given time t_0 is fixed, $\{q_i(t_0), p_i(t_0)\}$. The state of the system is known for all time if its initial state is known.

In this chapter, we shall study the postulates on which the quantum description of physical systems is based. We have already introduced them, in a qualitative and partial way, in Chapter I. Here we shall discuss them explicitly, within the framework of the formalism developed in Chapter II. These postulates will provide us with an answer to the following questions (which correspond to the three points enumerated above for the classical description):

(*i*) How is the state of a quantum system at a given time described mathematically?

(*ii*) Given this state, how can we predict the results of the measurement of various physical quantities?

(*iii*) How can the state of the system at an arbitrary time t be found when the state at time t_0 is known?

We shall begin by stating the postulates of quantum mechanics (§ B). Then we shall analyze their physical content and discuss their consequences (§§ C, D, E).

B. Statement of the postulates

B-1. Description of the state of a system

In Chapter I, we introduced the concept of the quantum state of a particle. We first characterized this state at a given time by a square-integrable wave function. Then, in Chapter II, we associated a ket of the state space $\mathcal{E}_\mathbf{r}$ with each wave function: choosing $|\psi\rangle$ belonging to $\mathcal{E}_\mathbf{r}$ is equivalent to choosing the corresponding function $\psi(\mathbf{r}) = \langle\mathbf{r}|\psi\rangle$. Therefore, the quantum state of a particle at a fixed time is characterized by a ket of the space $\mathcal{E}_\mathbf{r}$. In this form, the concept of a state can be generalized to any physical system.

> *First Postulate:* At a fixed time t_0, the state of an isolated physical system is defined by specifying a ket $|\psi(t_0)\rangle$ belonging to the state space $\mathcal{E}$.

It is important to note that, since $\mathcal{E}$ is a vector space, this first postulate implies a superposition principle: a linear combination of state vectors is a state vector. We shall discuss this fundamental point and its relations to the other postulates in § E.

B-2. Description of physical quantities

We have already used, in § D-1 of Chapter I, a differential operator H related to the total energy of a particle in a scalar potential. This is simply a special case of the second postulate.

> *Second Postulate:* Every measurable physical quantity $\mathcal{A}$ is described by an operator A acting in $\mathcal{E}$; this operator is an observable.

Comments:

(*i*) The fact that A is an observable (*cf.* Chap. II, § D-2) will be seen below (§ B-3) to be essential.

(*ii*) Unlike classical mechanics (*cf.* § A), quantum mechanics describes in a fundamentally different manner the state of a system and the associated physical quantities: a state is represented by a vector, a physical quantity by an operator.

B-3. The measurement of physical quantities

B-3-a. Possible results

The connection between the operator H and the total energy of the particle appeared in § D-1 of Chapter I in the following form: the only energies possible are the eigenvalues of the operator H. Here as well, this relation can be extended to all physical quantities.

> *Third Postulate:* The only possible result of the measurement of a physical quantity $\mathcal{A}$ is one of the eigenvalues of the corresponding observable A.

Comments:

(*i*) A measurement of $\mathcal{A}$ always gives a real value, since A is by definition Hermitian.

(*ii*) If the spectrum of A is discrete, the results that can be obtained by measuring $\mathcal{A}$ are quantized (§ C-2).

B-3-b. Principle of spectral decomposition

We are going to generalize and discuss in more detail the conclusions of § A-3 of Chapter I, where we analyzed a simple experiment performed on polarized photons.

Consider a system whose state is characterized, at a given time, by the ket $|\psi\rangle$, assumed to be normalized to 1:

$$\langle\psi|\psi\rangle = 1 \tag{B-1}$$

We want to predict the result of the measurement, at this time, of a physical quantity $\mathcal{A}$ associated with the observable A. This prediction, as we already know, is of a probabilistic sort. We are now going to give the rules that allow us to calculate the probability of obtaining any given eigenvalue of A.

α. *Case of a discrete spectrum*

First, let us assume that the spectrum of A is entirely discrete. If all the eigenvalues a_n of A are non-degenerate, there is associated with each of them a unique (to within a constant factor) eigenvector $|u_n\rangle$:

$$A|u_n\rangle = a_n|u_n\rangle \tag{B-2}$$

Since A is an observable, the set of the $|u_n\rangle$, which we shall take to be normalized, constitutes a basis in $\mathcal{E}$, and the state vector $|\psi\rangle$ can be written:

$$|\psi\rangle = \sum_n c_n|u_n\rangle \tag{B-3}$$

We postulate that the probability $\mathcal{P}(a_n)$ of finding a_n when $\mathcal{A}$ is measured is:

$$\mathcal{P}(a_n) = |c_n|^2 = |\langle u_n|\psi\rangle|^2 \tag{B-4}$$

> *Fourth Postulate (case of a discrete non-degenerate spectrum):* When the physical quantity $\mathcal{A}$ is measured on a system in the *normalized* state $|\psi\rangle$, the probability $\mathcal{P}(a_n)$ of obtaining the *non-degenerate eigenvalue* a_n of the corresponding observable A is:
>
> $$\mathcal{P}(a_n) = |\langle u_n|\psi\rangle|^2$$
>
> where $|u_n\rangle$ is the normalized eigenvector of A associated with the eigenvalue a_n.

If, now, some of the eigenvalues a_n are degenerate, several orthonormalized eigenvectors $|u_n^i\rangle$ correspond to them:

$$A|u_n^i\rangle = a_n|u_n^i\rangle\,; \quad i = 1, 2, \ldots g_n \tag{B-5}$$

$|\psi\rangle$ can still be expanded on the orthonormal basis $\{|u_n^i\rangle\}$:

$$|\psi\rangle = \sum_n \sum_{i=1}^{g_n} c_n^i|u_n^i\rangle \tag{B-6}$$

In this case, the probability $\mathcal{P}(a_n)$ becomes:

$$\mathcal{P}(a_n) = \sum_{i=1}^{g_n} |c_n^i|^2 = \sum_{i=1}^{g_n} |\langle u_n^i|\psi\rangle|^2 \tag{B-7}$$

(B-4) is then seen to be a special case of (B-7), which can therefore be considered to be the general formula.

Fourth Postulate (case of a discrete spectrum): When the physical quantity $\mathcal{A}$ is measured on a system in the *normalized* state $|\psi\rangle$, the probability $\mathcal{P}(a_n)$ of obtaining the eigenvalue a_n of the corresponding observable A is:

$$\mathcal{P}(a_n) = \sum_{i=1}^{g_n} |\langle u_n^i|\psi\rangle|^2$$

where g_n is the degree of degeneracy of a_n and $\{|u_n^i\rangle\}$ $(i = 1, 2, ..., g_n)$ is an orthonormal set of vectors which forms a basis in the eigensubspace $\mathcal{E}_n$ associated with the eigenvalue a_n of A.

For this postulate to make sense, it is obviously necessary that, if the eigenvalue a_n is degenerate, the probability $\mathcal{P}(a_n)$ be independent of the choice of the $\{|u_n^i\rangle\}$ basis in $\mathcal{E}_n$. To verify this, consider the vector:

$$|\psi_n\rangle = \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \tag{B-8}$$

where the coefficients c_n^i are the same as those appearing in the expansion (B-6) of $|\psi\rangle$:

$$c_n^i = \langle u_n^i|\psi\rangle \tag{B-9}$$

$|\psi_n\rangle$ is the part of $|\psi\rangle$ which belongs to $\mathcal{E}_n$, that is, the projection of $|\psi\rangle$ onto $\mathcal{E}_n$. This is, moreover, what we find when we substitute (B-9) into (B-8):

$$|\psi_n\rangle = \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i|\psi\rangle = P_n|\psi\rangle \tag{B-10}$$

where:

$$P_n = \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i| \tag{B-11}$$

is the projector onto $\mathcal{E}_n$ (§ B-3-b of Chapter II). Let us now calculate the square of the norm of $|\psi_n\rangle$. From (B-8):

$$\langle\psi_n|\psi_n\rangle = \sum_{i=1}^{g_n} |c_n^i|^2 \tag{B-12}$$

Therefore, $\mathcal{P}(a_n)$ *is the square of the norm of* $|\psi_n\rangle = P_n|\psi\rangle$, the projection of $|\psi\rangle$ onto $\mathcal{E}_n$. From this expression, it is clear that a change in the basis in $\mathcal{E}_n$ does not affect $\mathcal{P}(a_n)$. This probability is written:

$$\mathcal{P}(a_n) = \langle\psi|P_n^\dagger P_n|\psi\rangle \tag{B-13}$$

or, using the fact that P_n is Hermitian $(P_n^\dagger = P_n)$ and that it is a projector $(P_n^2 = P_n)$:

$$\mathcal{P}(a_n) = \langle\psi|P_n|\psi\rangle \tag{B-14}$$

β. Case of a continuous spectrum

Now let us assume that the spectrum of A is continuous and, for the sake of simplicity, non-degenerate. The system, orthonormal in the extended sense, of eigenvectors $|v_\alpha\rangle$ of A:

$$A|v_\alpha\rangle = \alpha|v_\alpha\rangle \tag{B-15}$$

forms a continuous basis in $\mathcal{E}$, in terms of which $|\psi\rangle$ can be expanded:

$$|\psi\rangle = \int \mathrm{d}\alpha\, c(\alpha)|v_\alpha\rangle \tag{B-16}$$

Since the possible results of a measurement of $\mathcal{A}$ form a continuous set, we must define a probability density, just as we did for the interpretation of the wave function of a particle (§ B-2 of Chapter I). The probability $\mathrm{d}\mathcal{P}(\alpha)$ of obtaining a value included between α and $\alpha + \mathrm{d}\alpha$ is given by:

$$\mathrm{d}\mathcal{P}(\alpha) = \rho(\alpha)\,\mathrm{d}\alpha$$

with:

$$\rho(\alpha) = |c(\alpha)|^2 = |\langle v_\alpha|\psi\rangle|^2 \tag{B-17}$$

Fourth Postulate (case of a continuous non-degenerate spectrum): When the physical quantity $\mathcal{A}$ is measured on a system in the *normalized* state $|\psi\rangle$, the probability $\mathrm{d}\mathcal{P}(\alpha)$ of obtaining a result included between α and $\alpha + \mathrm{d}\alpha$ is equal to:

$$\mathrm{d}\mathcal{P}(\alpha) = |\langle v_\alpha|\psi\rangle|^2\,\mathrm{d}\alpha$$

where $|v_\alpha\rangle$ is the eigenvector corresponding to the eigenvalue α of the observable A associated with $\mathcal{A}$.

Comments:

(*i*) It can be verified explicitly, in each of the cases considered above, that the total probability is equal to 1. For example, starting with formula (B-7), we find:

$$\sum_n \mathcal{P}(a_n) = \sum_n \sum_{i=1}^{g_n} |c_n^i|^2 = \langle\psi|\psi\rangle = 1 \tag{B-18}$$

since $|\psi\rangle$ is normalized. This last condition is therefore indispensable if the statements we have made are to be coherent. Nevertheless, it is not essential: if it is not fulfilled, it suffices to replace (B-7) and (B-17), respectively, by:

$$\mathcal{P}(a_n) = \frac{1}{\langle\psi|\psi\rangle} \sum_{i=1}^{g_n} |c_n^i|^2 \tag{B-19}$$

and:

$$\rho(\alpha) = \frac{1}{\langle\psi|\psi\rangle}|c(\alpha)|^2 \tag{B-20}$$

(*ii*) For the fourth postulate to be coherent, it is necessary for the operator A associated with any physical quantity to be an observable: it must be possible to expand any state on the eigenvectors of A.

(*iii*) We have not given the fourth postulate in its most general form. Starting with the discussion of the cases we have envisaged, it is simple to extend the principle of spectral decomposition to any situation (continuous degenerate spectrum, partially continuous and partially discrete spectrum, etc...). In § E, and later in Chapter IV, we shall apply this fourth postulate to a certain number of examples, pointing out certain implications of the superposition principle mentioned in § B-1.

γ. An important consequence

Consider two kets $|\psi\rangle$ and $|\psi'\rangle$ such that:

$$|\psi'\rangle = \mathrm{e}^{i\theta}|\psi\rangle \tag{B-21}$$

where θ is a real number. If $|\psi\rangle$ is normalized, so is $|\psi'\rangle$:

$$\langle\psi'|\psi'\rangle = \langle\psi|\mathrm{e}^{-i\theta}\,\mathrm{e}^{i\theta}|\psi\rangle = \langle\psi|\psi\rangle \tag{B-22}$$

The probabilities predicted for an arbitrary measurement are the same for $|\psi\rangle$ and $|\psi'\rangle$ since, for any $|u_n^i\rangle$:

$$|\langle u_n^i|\psi'\rangle|^2 = |\mathrm{e}^{i\theta}\langle u_n^i|\psi\rangle|^2 = |\langle u_n^i|\psi\rangle|^2 \tag{B-23}$$

Similarly, we can multiply $|\psi\rangle$ by a constant factor:

$$|\psi''\rangle = a\,\mathrm{e}^{i\theta}|\psi\rangle \tag{B-24}$$

without changing any of the physical results: since each coefficient c_n^i, or $c(\alpha)$, is multipled by the same factor in both the numerator and denominator of (B-19) and (B-20), there appear two factors of $|a|^2$ that cancel each other. Therefore, *two proportional state vectors represent the same physical state.*

Care must be taken to interpret this result correctly. For example, let us assume that:

$$|\psi\rangle = \lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle \tag{B-25}$$

where λ_1 and λ_2 are complex numbers. It is true that $\mathrm{e}^{i\theta_1}|\psi_1\rangle$ represents, for all real θ_1, the same physical state as $|\psi_1\rangle$, and $\mathrm{e}^{i\theta_2}|\psi_2\rangle$ represents the same state as $|\psi_2\rangle$. But, in general:

$$|\varphi\rangle = \lambda_1\,\mathrm{e}^{i\theta_1}|\psi_1\rangle + \lambda_2\,\mathrm{e}^{i\theta_2}|\psi_2\rangle \tag{B-26}$$

does not describe the same state as $|\psi\rangle$ (we shall see in § E-1 that the *relative* phases of the expansion coefficients of the state vector play an important role). This is not true for the special case where $\theta_1 = \theta_2 + 2n\pi$, that is, where:

$$|\varphi\rangle = \mathrm{e}^{i\theta_1}[\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle] = \mathrm{e}^{i\theta_1}|\psi\rangle \tag{B-27}$$

In other words: *a global phase factor does not affect the physical predictions, but the relative phases of the coefficients of an expansion are significant.*

B-3-c. Reduction of the wave packet

We have already introduced this concept in speaking of the measurement of the polarization of photons in the experiment described in § A-3 of Chapter I. We are now going to generalize it, confining ourselves, nevertheless, to the case of a discrete spectrum (we shall take up the case of a continuous spectrum in § E).

Assume that we want to measure, at a given time, the physical quantity $\mathcal{A}$. If the ket $|\psi\rangle$, which represents the state of the system immediately before the measurement, is known, the fourth postulate allows us to predict the probabilities of obtaining the various possible results. But when the measurement is actually performed, it is obvious that only one of these possible results is obtained. Immediately after this measurement, we cannot speak of the "probability of having obtained" this or that value: we know which one was actually obtained. We therefore possess additional information, and it is understandable that the state of the system after the measurement, which must incorporate this information, should be different from $|\psi\rangle$.

Let us first consider the case where the measurement of $\mathcal{A}$ yields a simple eigenvalue a_n of the observable A. We then postulate that the state of the system immediately after this measurement is the eigenvector $|u_n\rangle$ associated with a_n:

$$|\psi\rangle \overset{(a_n)}{\Longrightarrow} |u_n\rangle \tag{B-28}$$

Comments:

(*i*) We have been speaking about states "immediately before" the measurement ($|\psi\rangle$) and "immediately after" ($|u_n\rangle$). The precise meaning of these expressions is the following: assume that the measurement takes place at the time $t_0 > 0$, and that we know the state $|\psi(0)\rangle$ of the system at the time $t = 0$. The sixth postulate (see § B-4) describes how the system evolves over time, that is, enables us to calculate from $|\psi(0)\rangle$ the state $|\psi(t_0)\rangle$ "immediately before" the measurement. If the measurement has yielded the non-degenerate eigenvalue a_n, the state $|\psi'(t_1)\rangle$ at a time $t_1 > t_0$ must be calculated from $|\psi'(t_0)\rangle = |u_n\rangle$, the state "immediately after" the measurement, using the sixth postulate to determine the evolution of the state vector between the times t_0 and t_1 (Fig. 1).

(*ii*) If we perform a second measurement of $\mathcal{A}$ immediately after the first one (that is, before the system has had time to evolve), we shall always find the same result a_n, since the state of the system immediately before the second measurement is $|u_n\rangle$, and no longer $|\psi\rangle$.

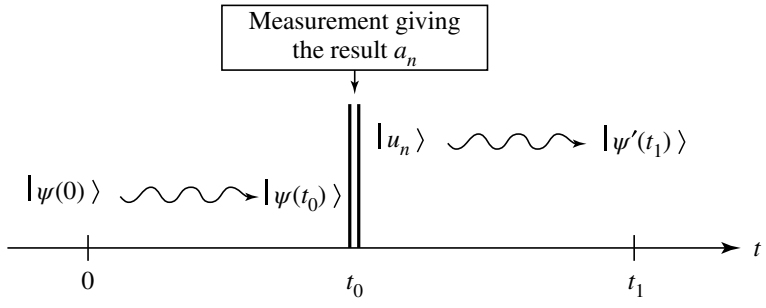


Figure 1: When a measurement at time t_0 of the observable A gives the result a_n, the state vector of the system undergoes an abrupt modification and becomes $|u_n\rangle$. This new initial state then evolves.

When the eigenvalue a_n given by the measurement is degenerate, postulate (B-28) can be generalized as follows. If the expansion of the state $|\psi\rangle$ immediately before the measurement is written, with the same notation as in section B-3-b:

$$|\psi\rangle = \sum_n \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \tag{B-29}$$

the modification of the state vector due to the measurement is written:

$$|\psi\rangle \overset{(a_n)}{\Longrightarrow} \frac{1}{\sqrt{\sum_{i=1}^{g_n} |c_n^i|^2}} \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \tag{B-30}$$

$\sum_{i=1}^{g_n} c_n^i |u_n^i\rangle$ is the vector $|\psi_n\rangle$ defined above [formula (B-8)], that is, the projection of $|\psi\rangle$ onto the eigensubspace associated with a_n. In (B-30), we normalized this vector since it is always more convenient to use state vectors of norm 1 [comment (i) of § B-3-b above]. With the notation of (B-10) and (B-11), we can therefore write (B-30) in the form:

$$|\psi\rangle \overset{(a_n)}{\Longrightarrow} \frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}} \tag{B-31}$$

Fifth Postulate: If the measurement of the physical quantity $\mathcal{A}$ on the system in the state $|\psi\rangle$ gives the result a_n, the state of the system immediately after the measurement is the normalized projection, $\dfrac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}$, of $|\psi\rangle$ onto the eigensubspace associated with a_n.

The state of the system immediately after the measurement is therefore always an eigenvector of A with the eigenvalue a_n. We stress the fact, however, that *it is not*

an arbitrary ket of the subspace $\mathcal{E}_n$, but the part of $|\psi\rangle$ that belongs to $\mathcal{E}_n$ (suitably normalized, for convenience). In the light of § B-3-b-γ above, equation (B-28) can be seen to be a special case of (B-30). When $g_n = 1$, the summation over i disappears from (B-30), which becomes:

$$\frac{1}{|c_n|} c_n |u_n\rangle = \mathrm{e}^{i\mathrm{Arg}(c_n)} |u_n\rangle \tag{B-32}$$

This ket indeed describes the same physical state as $|u_n\rangle$.

B-4. Time evolution of systems

We have already presented, in § B-2 of Chapter I, the Schrödinger equation for one particle. Here we shall write it in the general case.

Sixth Postulate: The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H(t) |\psi(t)\rangle$$

where $H(t)$ is the observable associated with the total energy of the system.

H is called the *Hamiltonian operator* of the system, as it is obtained from the classical Hamiltonian (Appendix III and § B-5 below).

B-5. Quantization rules

We are finally going to discuss how to construct, for a physical quantity $\mathcal{A}$ already defined in classical mechanics, the operator A which describes it in quantum mechanics.

B-5-a. Statement

Let us first consider a system composed of a single particle, without spin, subjected to a scalar potential. In this case:

With the position $\mathbf{r}(x, y, z)$ of the particle is associated the observable $\mathbf{R}(X, Y, Z)$. With the momentum $\mathbf{p}(p_x, p_y, p_z)$ of the particle is associated the observable $\mathbf{P}(P_x, P_y, P_z)$.

Recall that the components of $\mathbf{R}$ and $\mathbf{P}$ satisfy the canonical commutation relations [Chap. II, equations (E-30)]:

$$[R_i, R_j] = [P_i, P_j] = 0$$

$$[R_i, P_j] = i\hbar\, \delta_{ij} \tag{B-33}$$

Any physical quantity $\mathcal{A}$ related to this particle is expressed in terms of the fundamental dynamical variables $\mathbf{r}$ and $\mathbf{p}$: $\mathcal{A}(\mathbf{r}, \mathbf{p}, t)$. To obtain the corresponding observable

A, one could simply replace[1], in the expression for $\mathcal{A}(\mathbf{r}, \mathbf{p}, t)$, the variables $\mathbf{r}$ and $\mathbf{p}$ by the observables $\mathbf{R}$ and $\mathbf{P}$:

$$A(t) = \mathcal{A}(\mathbf{R}, \mathbf{P}, t) \tag{B-34}$$

However, this mode of action would be, in general, ambiguous. Assume, for example, that in $\mathcal{A}(\mathbf{r}, \mathbf{p}, t)$ there appears a term of the form:

$$\mathbf{r} \cdot \mathbf{p} = xp_x + yp_y + zp_z \tag{B-35}$$

In classical mechanics, the scalar product $\mathbf{r} \cdot \mathbf{p}$ is commutative, and one can just as well write:

$$\mathbf{p} \cdot \mathbf{r} = p_x x + p_y y + p_z z \tag{B-36}$$

But when $\mathbf{r}$ and $\mathbf{p}$ are replaced by the corresponding observables $\mathbf{R}$ and $\mathbf{P}$, the operators obtained from (B-35) and (B-36) are not identical [see relations (B-33)]:

$$\mathbf{R} \cdot \mathbf{P} \neq \mathbf{P} \cdot \mathbf{R} \tag{B-37}$$

Moreover, neither $\mathbf{R} \cdot \mathbf{P}$ nor $\mathbf{P} \cdot \mathbf{R}$ is Hermitian:

$$(\mathbf{R} \cdot \mathbf{P})^\dagger = (XP_x + YP_y + ZP_z)^\dagger = \mathbf{P} \cdot \mathbf{R} \tag{B-38}$$

To the preceding postulates, therefore, must be added a symmetrization rule. For example, the observable associated with $\mathbf{r} \cdot \mathbf{p}$ will be:

$$\frac{1}{2}(\mathbf{R} \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{R}) \tag{B-39}$$

which is indeed Hermitian. For an observable which is more complicated than $\mathbf{R} \cdot \mathbf{P}$, an analogous symmetrization is to be performed.

> The observable A which describes a classically defined physical quantity $\mathcal{A}$ is obtained by replacing, in the suitably symmetrized expression for $\mathcal{A}$, $\mathbf{r}$ and $\mathbf{p}$ by the observables $\mathbf{R}$ and $\mathbf{P}$ respectively.

We shall see, however, that there exist quantum physical quantities that have no classical equivalent and which are therefore defined directly by the corresponding observables (this is the case, for example, for particle spin).

Comment:

The preceding rules, and commutation rules (B-33) in particular, are valid only in cartesian coordinates. It would be possible to generalize them to other coordinate systems; however, they would no longer have the same simple form as they do above.

[1] See, in Complement B_{II}, the definition of a function of an operator.

B-5-b. Important examples

α. The Hamiltonian of a particle in a scalar potential

Consider a (spinless) particle of charge q and mass m, placed in an electric field derived from a scalar potential $U(\mathbf{r})$. The potential energy of the particle is therefore $V(\mathbf{r}) = qU(\mathbf{r})$, and the corresponding classical Hamiltonian is written [Appendix III, formula (29)]:

$$\mathcal{H}(\mathbf{r},\mathbf{p}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \tag{B-40}$$

with:

$$\mathbf{p} = m\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = m\mathbf{v} \tag{B-41}$$

where $\mathbf{v}$ is the particle's velocity.

No difficulties are presented by the construction of the quantum operator H corresponding to $\mathcal{H}$. No symmetrization is necessary, since neither $\mathbf{P}^2 = P_x^2 + P_y^2 + P_z^2$ nor $V(\mathbf{R})$ involves products of non-commuting operators. We therefore have:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R}) \tag{B-42}$$

$V(\mathbf{R})$ is the operator obtained by replacing $\mathbf{r}$ by $\mathbf{R}$ in $V(\mathbf{r})$ (*cf.* Complement $\mathrm{B_{II}}$, § 4).

In this particular case, the Schrödinger equation, given in the sixth postulate, becomes:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \left[\frac{\mathbf{P}^2}{2m} + V(\mathbf{R})\right]|\psi(t)\rangle \tag{B-43}$$

β. The Hamiltonian of a particle in a vector potential

If the particle is now placed in an arbitrary electromagnetic field, the classical Hamiltonian becomes [Appendix III, relation (66)]:

$$\mathcal{H}(\mathbf{r},\mathbf{p}) = \frac{1}{2m}\left[\mathbf{p} - q\mathbf{A}(\mathbf{r},t)\right]^2 + qU(\mathbf{r},t) \tag{B-44}$$

where $U(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ are the scalar and vector potentials which describe the electromagnetic field, and where $\mathbf{p}$ is given by:

$$\mathbf{p} = m\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} + q\mathbf{A}(\mathbf{r},t) = m\mathbf{v} + q\mathbf{A}(\mathbf{r},t) \tag{B-45}$$

Once again, since $\mathbf{A}(\mathbf{r},t)$ depends only on $\mathbf{r}$ and the parameter t (and not on $\mathbf{p}$), construction of the corresponding quantum operator $\mathbf{A}(\mathbf{R},t)$ presents no problem. The Hamiltonian operator H is then given by:

$$H(t) = \frac{1}{2m}\left[\mathbf{P} - q\mathbf{A}(\mathbf{R},t)\right]^2 + V(\mathbf{R},t) \tag{B-46}$$

with:

$$V(\mathbf{R},t) = qU(\mathbf{R},t) \tag{B-47}$$

and the Schrödinger equation is written:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \left\{\frac{1}{2m}\left[\mathbf{P} - q\mathbf{A}(\mathbf{R},t)\right]^2 + V(\mathbf{R},t)\right\}|\psi(t)\rangle \tag{B-48}$$

Comment:

Care must be taken not to confuse **p** (the momentum of the particle, also called the conjugate momentum of **r**) with $m\mathbf{v}$ (the mechanical momentum of the particle): the difference between these two quantities appears clearly in (B-45). In quantum mechanics, there of course exists an operator associated with the velocity of the particle which is written here:

$$\boldsymbol{\mathcal{V}} = \frac{1}{m}(\mathbf{P} - q\mathbf{A}) \tag{B-49}$$

H is then given by:

$$H(t) = \frac{1}{2}m\boldsymbol{\mathcal{V}}^2 + V(\mathbf{R},t) \tag{B-50}$$

It is the sum of two terms, one corresponding to the kinetic energy and the other to the potential energy of the particle.

However, it is the conjugate momentum **p** and not the mechanical momentum $m\mathbf{v}$ that becomes in quantum mechanics the operator **P** satisfying the canonical commutation relations (B-33).

C. The physical interpretation of the postulates concerning observables and their measurement

C-1. The quantization rules are consistent with the probabilistic interpretation of the wave function

It is natural to associate the observables **R** and **P**, whose action was defined in § E of Chapter II, with the position and momentum of a particle. First of all, each of the observables X, Y, Z and P_x, P_y, P_z possesses a continuous spectrum, and experiments indeed show that all real values are possible for the six position and momentum variables. Moreover, we shall see that applying the fourth postulate to the case of these observables enables us to re-derive the probabilistic interpretation of the wave function as well as that of its Fourier transform (see §§ B-2 and C-3 of Chapter I).

Let us consider, for simplicity, the one-dimensional problem. If the particle is in the normalized state $|\psi\rangle$, the probability that a measurement of its position will yield a result included between x and $x + \mathrm{d}x$ is equal to [formula (B-17)]:

$$\mathrm{d}\mathcal{P}(x) = |\langle x|\psi\rangle|^2\mathrm{d}x \tag{C-1}$$

where $|x\rangle$ is the eigenket of X with the eigenvalue x. We again find that the square of the modulus of the wave function $\psi(x) = \langle x|\psi\rangle$ is the particle's position probability density. Now, to the eigenvector $|p\rangle$ of the observable P corresponds the plane wave:

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\mathrm{e}^{\frac{ipx}{\hbar}} \tag{C-2}$$

and we have seen (§ C-3 of Chapter I) that the de Broglie relations associate with this wave a well-defined momentum which is precisely p. In addition, the probability of finding, for a particle in the state $|\psi\rangle$, a momentum between p and $p + \mathrm{d}p$ is:

$$\mathrm{d}\mathcal{P}(p) = |\langle p|\psi\rangle|^2 \mathrm{d}p = |\overline{\psi}(p)|^2 \mathrm{d}p \tag{C-3}$$

This is indeed what we found in § C-3 of Chapter I.

C-2. Quantization of certain physical quantities

As we have already pointed out, the third postulate enables us to explain the quantization observed for certain quantities, such as the energy of atoms. But it does not imply that all quantities are quantized, since observables exist whose spectrum is continuous. The physical predictions based on the third postulate are therefore not at all obvious *a priori*. For example, when we study the hydrogen atom (Chap. VII), we shall start from the total energy of the electron in the Coulomb potential of the proton, from which we shall deduce the Hamiltonian operator. Solving its eigenvalue equation, we shall find that the bound states of the system can only correspond to certain discrete energies which we shall calculate. Thus we shall not only explain the quantization of the levels of the hydrogen atom, but also predict the possible energy values, which can be measured experimentally. We stress the fact that these results will be obtained using the same fundamental interaction law used in classical mechanics in the macroscopic domain.

C-3. The measurement process

The fourth and fifth postulates pose a certain number of fundamental problems which we shall not consider here (see section 5 of the bibliography of volumes I and II, or for instance *Do we really understand quantum mechanics?* by F. Laloë, Cambridge University Press, 2019). There is, in particular, the question of the "fundamental" perturbation involved in the observation of a quantum system (*cf.* Chap. I, §§ A-2 and A-3). The origin of these problems lies in the fact that the system under study is treated independently from the measurement device, although their interaction is essential to the observation process. One should actually consider the system and the measurement device together as a whole. This raises delicate questions concerning the details of the measurement process.

We shall content ourselves with pointing out that the nondeterministic formulation of the fourth and fifth postulates is related to the problems that we have just mentioned. For example, the abrupt change from one state vector to another due to the measurement corresponds to the fundamental perturbation of which we have spoken. But it is impossible to predict what this perturbation will be, since it depends on the measurement result, which is not known with certainty in advance[2].

We shall consider here only ideal measurements. To understand this concept, let us return, for example, to the experiment of § A-3 of Chapter I on polarized photons. It is clear that when we grant that all photons polarized in a certain direction traverse the analyzer, we assume that the analyzer is perfect. In practice, obviously, it also absorbs some of the photons that it should let through. We shall therefore make the hypothesis,

[2]Except, obviously, in the case where one is sure of the result that will be found (probability equal to 1: the measurement does not modify the state of the system).

in the general case, that the measurement devices used are perfect: this amounts to assuming that the perturbation they provoke is due only to the quantum mechanical aspect of the measurement. Of course, real devices always present imperfections that affect the measurement and the system; but one can, in principle, constantly ameliorate them and thus approach the ideal limit defined by the postulates which we have stated.

C-4. Mean value of an observable in a given state

The predictions deduced from the fourth postulate are expressed in terms of probabilities. To verify them, it would be necessary to perform a large number of measurements under identical conditions. This means measuring the same quantity in a large number of systems which are all in the same quantum state. If these predictions are correct, the proportion of N identical experiments resulting in a given event will approach, as $N \longrightarrow \infty$, the theoretically predicted probability $\mathcal{P}$ of this event. Such a verification can only be carried out in the limit where $N \longrightarrow \infty$; in practice, N is of course finite, and statistical techniques must be used to interpret the results.

The mean value of the observable[3] A in the state $|\psi\rangle$, which we shall denote by $\langle A\rangle_\psi$, or more simply by $\langle A\rangle$, is defined as the average of the results obtained when a large number N of measurements of this observable are performed on systems which are all in the state $|\psi\rangle$. When $|\psi\rangle$ is given, the probabilities of finding all the possible results are known. The mean value $\langle A\rangle_\psi$ can therefore be predicted. We shall show that if $|\psi\rangle$ is normalized, $\langle A\rangle_\psi$ is given by the formula:

$$\boxed{\langle A\rangle_\psi = \langle\psi|A|\psi\rangle} \tag{C-4}$$

First consider the case where the entire spectrum of A is discrete. Out of N measurements of A (the system being in the state $|\psi\rangle$ each time), the eigenvalue a_n will be obtained $\mathcal{N}(a_n)$ times, with:

$$\frac{\mathcal{N}(a_n)}{N} \underset{N\longrightarrow\infty}{\longrightarrow} \mathcal{P}(a_n) \tag{C-5}$$

and:

$$\sum_n \mathcal{N}(a_n) = N \tag{C-6}$$

The mean value of the results of these N experiments is the sum of the values found divided by N (when $\mathcal{N}$ experiments have yielded the same result, this result will clearly appear $\mathcal{N}$ times in this sum). It is therefore equal to:

$$\frac{1}{N}\sum_n a_n\,\mathcal{N}(a_n) \tag{C-7}$$

Using (C-5), we see that when $N \longrightarrow \infty$, this mean value approaches:

$$\langle A\rangle_\psi = \sum_n a_n\,\mathcal{P}(a_n) \tag{C-8}$$

[3] We shall henceforth use the word "observable" to designate a physical quantity as well as the associated operator.

Now substitute into this formula expression (B-7) for $\mathcal{P}(a_n)$:

$$\langle A\rangle_\psi = \sum_n a_n \sum_{i=1}^{g_n} \langle\psi|u_n^i\rangle \langle u_n^i|\psi\rangle \tag{C-9}$$

Since:

$$A|u_n^i\rangle = a_n|u_n^i\rangle \tag{C-10}$$

(C-9) can be written in the form:

$$\begin{aligned}\langle A\rangle_\psi &= \sum_n \sum_{i=1}^{g_n} \langle\psi|A|u_n^i\rangle \langle u_n^i|\psi\rangle \\ &= \langle\psi|A\left[\sum_n \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i|\right]|\psi\rangle\end{aligned} \tag{C-11}$$

Since the $\{|u_n^i\rangle\}$ form an orthonormal basis of $\mathcal{E}$, the expression in brackets is equal to the identity operator (closure relation), and we obtain formula (C-4).

The argument is completely analogous for the case where the spectrum of A is continuous (for simplicity, we shall continue to assume it to be non-degenerate). Consider N identical experiments, and call $\mathrm{d}\mathcal{N}(\alpha)$ the number of experiments which have yielded a result included between α and $\alpha + \mathrm{d}\alpha$. We have, similarly:

$$\frac{\mathrm{d}\mathcal{N}(\alpha)}{N} \underset{N\longrightarrow\infty}{\longrightarrow} \mathrm{d}\mathcal{P}(\alpha) \tag{C-12}$$

The mean value of the results obtained is $\dfrac{1}{N}\displaystyle\int \alpha\ \mathrm{d}\mathcal{N}(\alpha)$, which, when $N \longrightarrow \infty$, approaches:

$$\langle A\rangle_\psi = \int \alpha\ \mathrm{d}\mathcal{P}(\alpha) \tag{C-13}$$

Substitute into (C-13) the expression for $\mathrm{d}\mathcal{P}(\alpha)$ given by (B-17):

$$\langle A\rangle_\psi = \int \alpha\ \langle\psi|v_\alpha\rangle \langle v_\alpha|\psi\rangle\ \mathrm{d}\alpha \tag{C-14}$$

We can use the equation:

$$A|v_\alpha\rangle = \alpha|v_\alpha\rangle \tag{C-15}$$

to transform (C-14) into:

$$\begin{aligned}\langle A\rangle_\psi &= \int \langle\psi|A|v_\alpha\rangle \langle v_\alpha|\psi\rangle\ \mathrm{d}\alpha \\ &= \langle\psi|A\left[\int \mathrm{d}\alpha\ |v_\alpha\rangle \langle v_\alpha|\right]|\psi\rangle\end{aligned} \tag{C-16}$$

Using the closure relation satisfied by the states $|v_\alpha\rangle$, we again find formula (C-4).

Comments:

(*i*) $\langle A \rangle$, the average over a set of identical measurements, must not be confused with the time averages sometimes taken when dealing with time-dependent phenomena.

(*ii*) If the ket $|\psi\rangle$ representing the state of the system is not normalized, formula (C-4) becomes [*cf.* comment (*i*) of § B-3-b]:

$$\langle A \rangle_\psi = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle} \tag{C-17}$$

(*iii*) In practice, to calculate $\langle A \rangle_\psi$ explicitly, one often places oneself in a particular representation. For example:

$$\begin{aligned} \langle X \rangle_\psi &= \langle \psi | X | \psi \rangle \\ &= \int \mathrm{d}^3 r \ \langle \psi | \mathbf{r} \rangle \langle \mathbf{r} | X | \psi \rangle \\ &= \int \mathrm{d}^3 r \ \psi^*(\mathbf{r}) \ x \ \psi(\mathbf{r}) \end{aligned} \tag{C-18}$$

using the definition of the X operator [*cf.* Chap. II, relations (E-22)]. Similarly:

$$\begin{aligned} \langle P_x \rangle_\psi &= \langle \psi | P_x | \psi \rangle \\ &= \int \mathrm{d}^3 p \ \overline{\psi}^*(\mathbf{p}) \ p_x \ \overline{\psi}(\mathbf{p}) \end{aligned} \tag{C-19}$$

or, using the $\{|\mathbf{r}\rangle\}$ representation:

$$\begin{aligned} \langle P_x \rangle_\psi &= \int \mathrm{d}^3 r \ \langle \psi | \mathbf{r} \rangle \langle \mathbf{r} | P_x | \psi \rangle \\ &= \int \mathrm{d}^3 r \ \psi^*(\mathbf{r}) \left[\frac{\hbar}{i} \frac{\partial}{\partial x} \psi(\mathbf{r}) \right] \end{aligned} \tag{C-20}$$

since $\mathbf{P}$ is then represented by $\dfrac{\hbar}{i} \boldsymbol{\nabla}$ [formula (E-26) of Chapter II].

C-5. The root mean square deviation

$\langle A \rangle$ indicates the order of magnitude of the values of the observable A when the system is in the state $|\psi\rangle$. However, this mean value does not give any idea of the dispersion of the results we expect when measuring A. Assume, for example, that the spectrum of A is continuous and that, for a given state $|\psi\rangle$, the curve representing the variation with respect to α of the probability density $\rho(\alpha) = |\langle v_\alpha | \psi \rangle|^2$ has the shape shown in Figure 2. For a system in the state $|\psi\rangle$, nearly all the values that can be found when A is measured are included in an interval of width δA containing $\langle A \rangle$, where

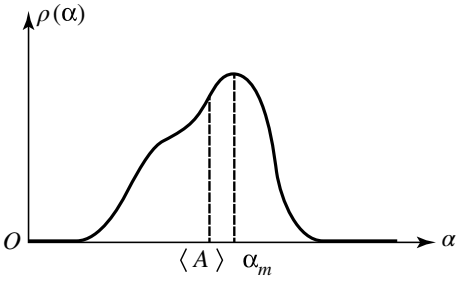


Figure 2: Variation with respect to α of the probability density $\rho(\alpha)$. The mean value $\langle A\rangle$ is the abscissa of the center of gravity of the area under the curve (it does not necessarily coincide with the abscissa α_m of the maximum of the function).

the quantity δA characterizes the width of the curve: the smaller δA, the more the measurement results are concentrated about $\langle A\rangle$.

How can we define, in a general way, a quantity which characterizes the dispersion of the measurement results about $\langle A\rangle$? We might envisage the following method: for each measurement, take the difference between the value obtained and $\langle A\rangle$; then calculate the average of these deviations, dividing their sum by the number N of experiments. It is easy to see, however, that the result obtained would be zero; we have, obviously:

$$\langle A - \langle A\rangle\rangle = \langle A\rangle - \langle A\rangle = 0 \tag{C-21}$$

By the very definition of $\langle A\rangle$, the average of the negative deviations balances exactly the average of the positive ones.

To avoid this compensation, it suffices to define ΔA such that $(\Delta A)^2$ is the mean of the squares of the deviations:

$$(\Delta A)^2 = \langle (A - \langle A\rangle)^2\rangle \tag{C-22}$$

By definition, we therefore introduce the *root mean square deviation* ΔA by setting:

$$\boxed{\Delta A = \sqrt{\langle (A - \langle A\rangle)^2\rangle}} \tag{C-23}$$

Using the expression for the mean value given in (C-4), we then have:

$$\Delta A = \sqrt{\langle \psi | (A - \langle A\rangle)^2 | \psi\rangle} \tag{C-24}$$

This relation can also be written in a slightly different way:

$$\begin{aligned}\langle (A - \langle A\rangle)^2\rangle &= \langle (A^2 - 2\langle A\rangle A + \langle A\rangle^2)\rangle \\ &= \langle A^2\rangle - 2\langle A\rangle^2 + \langle A\rangle^2 \\ &= \langle A^2\rangle - \langle A\rangle^2 \end{aligned} \tag{C-25}$$

The root-mean-square deviation ΔA is therefore also given by:

$$\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2} \tag{C-26}$$

For example, in the case of the continuous spectrum of the observable A considered above, ΔA is given by:

$$(\Delta A)^2 = \int_{-\infty}^{+\infty} [\alpha - \langle A \rangle]^2 \, \rho(\alpha) \, \mathrm{d}\alpha$$

$$= \int_{-\infty}^{+\infty} \alpha^2 \rho(\alpha) \, \mathrm{d}\alpha - \left[\int_{-\infty}^{+\infty} \alpha \, \rho(\alpha) \, \mathrm{d}\alpha \right]^2 \tag{C-27}$$

If definition (C-23) is applied to the observables $\mathbf{R}$ and $\mathbf{P}$, it can be shown (Complement $\mathrm{C_{III}}$), using their commutation relations, that for any state $|\psi\rangle$, one has:

$$\begin{cases} \Delta X . \Delta P_x \geq \hbar/2 \\ \Delta Y . \Delta P_y \geq \hbar/2 \\ \Delta Z . \Delta P_z \geq \hbar/2 \end{cases} \tag{C-28}$$

In other words, we find the Heisenberg relations (§ C-3 of Chapter I) again, but with a precise lower limit, which arises from the precise definition of the uncertainties.

C-6. Compatibility of observables

C-6-a. Compatibility and commutation rules

Consider two observables A and B which commute:

$$[A, B] = 0 \tag{C-29}$$

We shall assume for simplicity that both of their spectra are discrete. According to the theorem proved in § D-3-a of Chapter II, there exists a basis of the state space composed of eigenkets common to A and B, which we shall denote by $|a_n, b_p, i\rangle$:

$$A|a_n, b_p, i\rangle = a_n \, |a_n, b_p, i\rangle$$

$$B|a_n, b_p, i\rangle = b_p \, |a_n, b_p, i\rangle \tag{C-30}$$

(the index i allows us to distinguish, if necessary, between the different vectors corresponding to one pair of eigenvalues). Therefore, for any a_n and b_p (chosen, respectively, in the spectra of A and B), there exists at least one state $|a_n, b_p, i\rangle$ for which a measurement of A will always give a_n and a measurement of B will always give b_p. Two such observables A and B which can be simultaneously determined are said to be compatible.

On the other hand, if A and B do not commute, a state cannot in general[4] be a simultaneous eigenvector of these two observables. They are said to be incompatible.

Let us examine more closely the measurement of two compatible observables on a system which is initially in an arbitrary (normalized) state $|\psi\rangle$. This state can always be written:

$$|\psi\rangle = \sum_{n,p,i} c_{n,p,i} |a_n, b_p, i\rangle \tag{C-31}$$

[4]Some kets may be simultaneous eigenvectors of A and B. But there would not be a sufficient number of them to form a basis, as would be the case if A and B commuted.

First assume that we measure A and then, immediately afterwards, B (before the system has had time to evolve). Let us calculate the probability $\mathcal{P}(a_n, b_p)$ of obtaining a_n in the first measurement and b_p in the second one. We begin by measuring A in the state $|\psi\rangle$; the probability of finding a_n is therefore:

$$\mathcal{P}(a_n) = \sum_{p,i} |c_{n,p,i}|^2 \tag{C-32}$$

When we then measure B, the system is no longer in the state $|\psi\rangle$ but, if we have found a_n, in the state $|\psi'_n\rangle$:

$$|\psi'_n\rangle = \frac{1}{\sqrt{\sum_{p,i} |c_{n,p,i}|^2}} \sum_{p,i} c_{n,p,i} |a_n, b_p, i\rangle \tag{C-33}$$

The probability of obtaining b_p when it is known that the first measurement has yielded a_n is therefore equal to:

$$\mathcal{P}_{a_n}(b_p) = \frac{1}{\sum_{p,i} |c_{n,p,i}|^2} \sum_{i} |c_{n,p,i}|^2 \tag{C-34}$$

The probability $\mathcal{P}(a_n, b_p)$ sought corresponds to a "composite event": to be in a favorable case, we must first find a_n and then, having satisfied this first condition, find b_p. Therefore:

$$\mathcal{P}(a_n, b_p) = \mathcal{P}(a_n) \times \mathcal{P}_{a_n}(b_p) \tag{C-35}$$

Substituting into this formula expressions (C-32) and (C-34), we obtain:

$$\mathcal{P}(a_n, b_p) = \sum_{i} |c_{n,p,i}|^2 \tag{C-36}$$

Moreover, the state of the system becomes, immediately after the second measurement:

$$|\psi''_{n,p}\rangle \frac{1}{\sqrt{\sum_{i} |c_{n,p,i}|^2}} \sum_{i} c_{n,p,i} |a_n, b_p, i\rangle \tag{C-37}$$

Therefore, if we decide to measure either A or B again, we are sure of the result (a_n or b_p): $|\psi''_{n,p}\rangle$ is an eigenvector common to A and B with the eigenvalues a_n and b_p respectively.

Let us now return to the system in the state $|\psi\rangle$, and let us measure the two observables in the opposite order (B, then A). What is the probability $\mathcal{P}(b_p, a_n)$ of obtaining the same results as before? The reasoning is the same. We have here:

$$\mathcal{P}(b_p, a_n) = \mathcal{P}(b_p) \times \mathcal{P}_{b_p}(a_n) \tag{C-38}$$

From (C-31), we see that:

$$\mathcal{P}(b_p) = \sum_{n,i} |c_{n,p,i}|^2 \tag{C-39}$$

and that, after a measurement of B which yields b_p, the state of the system becomes:

$$|\varphi'_p\rangle = \frac{1}{\sqrt{\sum_{n,i} |c_{n,p,i}|^2}} \sum_{n,i} c_{n,p,i} |a_n, b_p, i\rangle \tag{C-40}$$

Therefore:

$$\mathcal{P}_{b_p}(a_n) = \frac{1}{\sum\limits_{n,i} |c_{n,p,i}|^2} \sum_i |c_{n,p,i}|^2 \tag{C-41}$$

and finally:

$$\mathcal{P}(b_p, a_n) = \sum_i |c_{n,p,i}|^2 \tag{C-42}$$

If we have indeed found b_p and then a_n, the system has gone into the state:

$$|\varphi''_{p,n}\rangle = \frac{1}{\sqrt{\sum\limits_i |c_{n,p,i}|^2}} \sum_i |c_{n,p,i}| a_n, b_p, i\rangle \tag{C-43}$$

When two observables are compatible, the physical predictions are the same, whatever the order of performing the two measurements (provided that the time interval which separates them is sufficiently small). The probabilities of obtaining either a_n then b_p or b_p then a_n are identical:

$$\mathcal{P}(a_n, b_p) = \mathcal{P}(b_p, a_n) = \sum_i |c_{n,p,i}|^2 = \sum_i |\langle a_n, b_p, i|\psi\rangle|^2 \tag{C-44}$$

Moreover, the state of the system immediately after the two measurements is in both cases (if the results are a_n and b_p for A and B respectively):

$$|\psi''_{n,p}\rangle = |\varphi''_{p,n}\rangle = \frac{1}{\sqrt{\sum\limits_i |c_{n,p,i}|^2}} \sum_i c_{n,p,i} |a_n, b_p, i\rangle \tag{C-45}$$

New measurements of A or B will yield the same values again without fail.

The preceding discussion thus leads to the following result: when two observables A and B are compatible, the measurement of B does not cause any loss of information previously obtained from a measurement of A (and vice versa) but, on the contrary, adds to it. Moreover, the order of measuring the two observables A and B is of no importance. This last point, furthermore, enables us to envisage the *simultaneous measurement* of A and B. The fourth and fifth postulates can be generalized to the case of such a simultaneous measurement, as can be seen from formulas (C-44) and (C-45). To the result $\{a_n, b_p\}$ correspond the orthonormal eigenvectors $|a_n, b_p, i\rangle$. From this, (C-44) and (C-45) can be seen to be applications of postulates (B-7) and (B-30).

On the other hand, if A and B do not commute, the preceding arguments are no longer valid. To understand this in a simple way, imagine that the state space $\mathcal{E}$ is replaced by the two-dimensional space of real vectors. The vectors $|u_1\rangle$ and $|u_2\rangle$ in Figure 3 are eigenvectors of A with eigenvalues a_1 and a_2 respectively; $|v_1\rangle$ and $|v_2\rangle$ are eigenvectors of B with eigenvalues b_1 and b_2 respectively. Each of the two sets $\{|u_1\rangle, |u_2\rangle\}$ and $\{|v_1\rangle, |v_2\rangle\}$ forms an orthonormal basis in $\mathcal{E}$. We shall therefore represent them in Figure 3 by two pairs of perpendicular unit vectors. The fact that A and B do not commute implies that these two pairs do not coincide. The physical system under study is initially in the normalized state $|\psi\rangle$, which is represented in the figure by an arbitrary

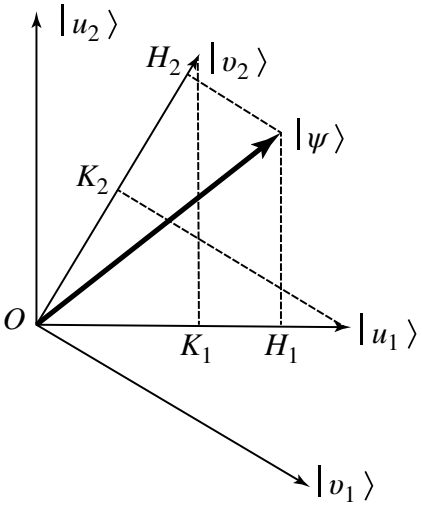


Figure 3: Diagram associated with the successive measurement of two non-compatible observables A and B. The state vector of the system is $|\psi\rangle$. The eigenvectors of A are $|u_1\rangle$ and $u_2\rangle$ (eigenvalues a_1 and a_2), which are different from those of B, $|v_1\rangle$ and $|v_2\rangle$ (eigenvalues b_1 and b_2).

unit vector. We measure A and find, for example, a_1; the system goes into the state $|u_1\rangle$. We then measure B and find, for example, b_2; the state of the system becomes $|v_2\rangle$:

$$|\psi\rangle \overset{(a_1)}{\Longrightarrow} |u_1\rangle \overset{(b_2)}{\Longrightarrow} |v_2\rangle \tag{C-46}$$

If, on the other hand, we perform the measurements in the opposite order, obtaining the same results:

$$|\psi\rangle \overset{(b_2)}{\Longrightarrow} |v_2\rangle \overset{(a_1)}{\Longrightarrow} |u_1\rangle \tag{C-47}$$

The final state of the system is not the same in both cases. We also see from Figure 3 that:

$$\begin{aligned} \mathcal{P}(a_1, b_2) &= |OH_1|^2 \times |OK_2|^2 \\ \mathcal{P}(b_2, a_1) &= |OH_2|^2 \times |OK_1|^2 \end{aligned} \tag{C-48}$$

Although $|OK_1| = |OK_2|$, in general $|OH_1| \neq |OH_2|$ and:

$$\mathcal{P}(b_2, a_1) \neq \mathcal{P}(a_1, b_2) \tag{C-49}$$

Therefore: *two incompatible observables cannot be measured simultaneously.* It can be seen from (C-46) and (C-47) that *the second measurement causes the information supplied by the first one to be lost.* If, for example, after the sequence represented in (C-46), we measure A again, we can no longer be sure of the result since $|v_2\rangle$ is not an eigenvector of A. All that was gained by the first measurement of A is thus lost.

C-6-b. Preparation of a state

Let us consider a physical system in the state $|\psi\rangle$ and measure the observable A (whose spectrum we assume to be discrete).

If the measurement yields a non-degenerate eigenvalue a_n, the state of the system immediately after this measurement is the corresponding eigenvector $|u_n\rangle$. In this case, it suffices to know the result of this measurement to be able to determine unambiguously the state of the system after this measurement, as it does not depend on the initial ket. As we have already noted, at the end of § B-3-c, this is due to the fact that $\frac{c_n}{|c_n|}|u_n\rangle$ represents the same physical state as $|u_n\rangle$ itself.

The same does not hold true when the eigenvalue a_n found in the measurement is degenerate. In:

$$|\psi'_n\rangle = \frac{1}{\sqrt{\sum_i |c_n^i|^2}} \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \tag{C-50}$$

the absolute values of the coefficients c_n^i and their relative phases are significant (§ B-3-b-γ). Since the c_n^i are fixed when the initial state $|\psi\rangle$ is specified, the state $|\psi'_n\rangle$ after the measurement therefore depends on $|\psi\rangle$.

However, we saw in § C-6-a that two compatible observables A and B can be measured simultaneously. If the result (a_n, b_p) of this combined measurement corresponds to only one eigenvector $|a_n, b_p\rangle$ common to A and B, there is no summation over i in formula (C-37), which becomes:

$$|\psi''_{n,p}\rangle = \frac{c_{n,p}}{|c_{n,p}|}|a_n, b_p\rangle \tag{C-51}$$

This state is physically equivalent to $|a_n, b_p\rangle$. Again, specifying the result of the measurement uniquely determines the final state of the system, which is therefore independent of the initial ket $|\psi\rangle$.

If there are associated with (a_n, b_p) several eigenvectors $|a_n, b_p, i\rangle$ of A and B, we can go back to the first argument, measuring, at the same time as A and B, a third observable C which is compatible with both of them. We then arrive at the following conclusion: *for the state of the system after a measurement to be completely defined uniquely by the result obtained, this measurement must be made on a complete set of commuting observables* (§ D-3-b of Chapter II). This is the property which justifies physically the introduction of the concept of a C.S.C.O.

The methods that can be used to *prepare a system in a well-defined quantum state* are analogous, in principle, to those used to obtain polarized light. When a polarizer is placed in the path of a light beam, the outgoing light is polarized along a direction which is characteristic of the polarizer and therefore independent of the state of polarization of the incoming light. Similarly, we can construct devices, intended to prepare a quantum system, in such a way that they only allow the passage of one state, corresponding to a particular eigenvalue for each of the observables of the complete set chosen. We shall study a concrete example of the preparation of a quantum system in Chapter IV (§ B-1).

Comment:

The measurement of a C.S.C.O. enables us to prepare only one of the basis states associated with this C.S.C.O. However, it is obvious that changing the set of observables allows us to obtain other states of the system. We shall see explicitly in a concrete example, in § B-1 of Chapter IV, that we can prepare in this way any state of the space $\mathcal{E}$.

D. The physical implications of the Schrödinger equation

The Schrödinger equation plays a fundamental role in quantum mechanics since, according to the sixth postulate stated above, it is the equation that governs the time evolution of the physical system. In this section, we shall study in detail the most important properties of this equation.

D-1. General properties of the Schrödinger equation

D-1-a. Determinism in the evolution of physical systems

The Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle \tag{D-1}$$

is of first order in t. Consequently, given the initial state $|\psi(t_0)\rangle$, the state $|\psi(t)\rangle$ at any subsequent time t is determined. There is no indeterminacy in the time evolution of a quantum system. Indeterminacy appears only when a physical quantity is measured, the state vector then undergoing an unpredictable modification (*cf.* fifth postulate). However, between two measurements, the state vector evolves in a perfectly deterministic way, in accordance with equation (D-1).

D-1-b. The superposition principle

Equation (D-1) is linear and homogeneous. It follows that its solutions are linearly superposable.

Let $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$ be two solutions of (D-1). If the initial state of the system is $|\psi(t_0)\rangle = \lambda_1|\psi_1(t_0)\rangle + \lambda_2|\psi_2(t_0)\rangle$ (where λ_1 and λ_2 are two complex constants), to it corresponds, at time t, the state $|\psi(t)\rangle = \lambda_1|\psi_1(t)\rangle + \lambda_2|\psi_2(t)\rangle$. The correspondence between $|\psi(t_0)\rangle$ and $|\psi(t)\rangle$ is therefore linear. We shall later study (Complement $\mathrm{F_{III}}$) the properties of the linear operator $U(t,t_0)$ which transforms $|\psi(t_0)\rangle$ into $|\psi(t)\rangle$:

$$|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle \tag{D-2}$$

D-1-c. Conservation of probability

α. The norm of the state vector remains constant

Since the Hamiltonian operator $H(t)$ that appears in (D-1) is Hermitian, the square of the norm of the state vector, $\langle\psi(t)|\psi(t)\rangle$, does not depend on t as we shall now show:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\psi(t)\rangle = \left[\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\right]|\psi(t)\rangle + \langle\psi(t)|\left[\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle\right] \tag{D-3}$$

According to (D-1), we can write:

$$\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \frac{1}{i\hbar}H(t)|\psi(t)\rangle \tag{D-4}$$

Taking the Hermitian conjugates of both sides of (D-4), we find:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)| = -\frac{1}{i\hbar}\langle\psi(t)|H^\dagger(t) = -\frac{1}{i\hbar}\langle\psi(t)|H(t) \tag{D-5}$$

since $H(t)$ is Hermitian (it is an observable). Substituting (D-4) and (D-5) into (D-3), we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\psi(t)\rangle = -\frac{1}{i\hbar}\langle\psi(t)|H(t)|\psi(t)\rangle + \frac{1}{i\hbar}\langle\psi(t)|H(t)|\psi(t)\rangle$$
$$= 0 \tag{D-6}$$

The property of the norm conservation is very useful in quantum mechanics. For example, it becomes indispensable when we interpret the square of the modulus $|\psi(\mathbf{r},t)|^2$ of the wave function of a spinless particle as being the position probability density. The fact that the state $|\psi(t_0)\rangle$ of the particle is normalized at time t_0 is expressed by the relation:

$$\langle\psi(t_0)|\psi(t_0)\rangle = \int \mathrm{d}^3r\ |\psi(\mathbf{r},t_0)|^2 = 1 \tag{D-7}$$

where $\psi(\mathbf{r},t_0) = \langle\mathbf{r}|\psi(t_0)\rangle$ is the wave function associated with $|\psi(t_0)\rangle$. Equation (D-7) means that the total probability of finding the particle somewhere in all space is equal to 1. The property of conservation of the norm which we have just proved is expressed by the equation:

$$\langle\psi(t)|\psi(t)\rangle = \int \mathrm{d}^3r\ |\psi(\mathbf{r},t)|^2 = \langle\psi(t_0)|\psi(t_0)\rangle = 1 \tag{D-8}$$

where $|\psi(t)\rangle$ is the solution of (D-1) corresponding to the initial state $|\psi(t_0)\rangle$. In other words, time evolution does not modify the global probability of finding the particle in all space, which always remains equal to 1. Thus $|\psi(\mathbf{r},t)|^2$ can be interpreted as a probability density.

β. Local conservation of probability. Probability densities and probability currents

In this paragraph, we shall confine ourselves to the case of a physical system composed of *only one* (spinless) *particle*.

In this case, if $\psi(\mathbf{r},t)$ is normalized,

$$\rho(\mathbf{r},t) = |\psi(\mathbf{r},t)|^2 \tag{D-9}$$

is a *probability density*: the probability $\mathrm{d}\mathcal{P}(\mathbf{r},t)$ of finding, at time t, the particle in an infinitesimal volume d^3r located at point $\mathbf{r}$ is equal to:

$$\mathrm{d}\mathcal{P}(\mathbf{r},t) = \rho(\mathbf{r},t)\ \mathrm{d}^3r \tag{D-10}$$

We have just shown that the integral of $\rho(\mathbf{r},t)$ over all space remains constant for all time (and equal to 1 if ψ is normalized). This does not mean that $\rho(\mathbf{r},t)$ must be independent of t at every point $\mathbf{r}$. The situation is analogous to the one encountered in electromagnetism. If, in an isolated physical system, there is a charge distributed in space with the volume density $\rho(\mathbf{r},t)$, the total charge [the integral of $\rho(\mathbf{r},t)$ over all space] is conserved in time. However, within the system, the spatial distribution of this charge may vary, giving rise to electric currents.

In fact, this analogy can be carried further. Global conservation of electrical charge is based on local conservation. If the charge Q contained within a fixed volume V varies over time, the closed surface S which limits V must be traversed by an electric current. More precisely, the variation $\mathrm{d}Q$ during a time $\mathrm{d}t$ of the charge contained within V is equal to $-I\ \mathrm{d}t$, where I is the intensity of the current traversing S, that is, the flux of the vector current density $\mathbf{J}(\mathbf{r},t)$ leaving S. Using classical vector analysis, we can express local conservation of electrical charge in the form:

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) + \text{div}\ \mathbf{J}(\mathbf{r},t) = 0 \tag{D-11}$$

We are going to show that it is possible to find a vector $\mathbf{J}(\mathbf{r},t)$, a *probability current*, which satisfies an equation identical to (D-11): there is then *local* conservation of probability. It is as if we were dealing with a "probability fluid" whose density and motion were described by $\rho(\mathbf{r},t)$ and $\mathbf{J}(\mathbf{r},t)$. If the probability of finding the particle in the (fixed) volume d^3r about $\mathbf{r}$ varies over time, it means that the probability current has a non-zero flux accross the surface which limits this volume element.

First of all, let us assume that the particle under study is subjected to a scalar potential $V(\mathbf{r},t)$. Its Hamiltonian is then:

$$H = \frac{\mathbf{P^2}}{2m} + V(\mathbf{R},t) \tag{D-12}$$

and the Schrödinger equation is written, in the $\{|\mathbf{r}\rangle\}$ representation (see Complement $\mathrm{D_{II}}$):

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = -\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r},t) + V(\mathbf{r},t)\,\psi(\mathbf{r},t) \tag{D-13}$$

$V(\mathbf{r},t)$ must be real for H to be Hermitian. The complex conjugate equation of (D-13) is therefore:

$$-i\hbar\frac{\partial}{\partial t}\psi^*(\mathbf{r},t) = -\frac{\hbar^2}{2m}\Delta\psi^*(\mathbf{r},t) + V(\mathbf{r},t)\,\psi^*(\mathbf{r},t) \tag{D-14}$$

Multiply both sides of (D-13) by $\psi^*(\mathbf{r},t)$ and both sides of (D-14) by $-\psi(\mathbf{r},t)$. Add the two equations thus obtained. It follows that:

$$i\hbar\frac{\partial}{\partial t}\left[\psi^*(\mathbf{r},t)\psi(\mathbf{r},t)\right] = -\frac{\hbar^2}{2m}\left[\psi^*\Delta\psi - \psi\,\Delta\psi^*\right] \tag{D-15}$$

that is:

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) + \frac{\hbar}{2mi}\left[\psi^*(\mathbf{r},t)\Delta\psi(\mathbf{r},t) - \psi(\mathbf{r},t)\,\Delta\psi^*(\mathbf{r},t)\right] = 0 \tag{D-16}$$

If we set:

$$\begin{aligned}\mathbf{J}(\mathbf{r},t) &= \frac{\hbar}{2mi}\left[\psi^*\boldsymbol{\nabla}\psi - \psi\,\boldsymbol{\nabla}\psi^*\right] \\ &= \frac{1}{m}\mathrm{Re}\left[\psi^*\left(\frac{\hbar}{i}\boldsymbol{\nabla}\psi\right)\right]\end{aligned} \tag{D-17}$$

equation (D-16) can be put into the form of (D-11) since:

$$\begin{aligned}\operatorname{div} \mathbf{J}(\mathbf{r},t) &= \boldsymbol{\nabla}\cdot\mathbf{J} \\ &= \frac{\hbar}{2mi}\left[(\boldsymbol{\nabla}\psi^*)\cdot(\boldsymbol{\nabla}\psi)+\psi^*\left(\boldsymbol{\nabla}^2\psi\right)-(\boldsymbol{\nabla}\psi)\cdot(\boldsymbol{\nabla}\psi^*)-\psi\left(\boldsymbol{\nabla}^2\psi^*\right)\right] \\ &= \frac{\hbar}{2mi}\left[\psi^*\Delta\psi-\psi\,\Delta\psi^*\right] \qquad \text{(D-18)}\end{aligned}$$

We have therefore proved the equation of local conservation of probability and have found the expression for the probability current in terms of the normalized wave function $\psi(\mathbf{r},t)$.

Comment:

The form of the probability current (D-17) can be interpreted as follows. $\mathbf{J}(\mathbf{r},t)$ appears as the mean value, in the state $|\psi(t)\rangle$, of an operator $\mathbf{K}(\mathbf{r})$ given by:

$$\mathbf{K}(\mathbf{r}) = \frac{1}{2m}\left[|\mathbf{r}\rangle\langle\mathbf{r}|\mathbf{P}+\mathbf{P}|\mathbf{r}\rangle\langle\mathbf{r}|\right] \qquad \text{(D-19)}$$

Now the mean value of the operator $|\mathbf{r}\rangle\langle\mathbf{r}|$ is $|\psi(\mathbf{r},t)|^2$, that is, the probability density $\rho(\mathbf{r},t)$, and $\dfrac{\mathbf{P}}{m}$ is the velocity operator $\boldsymbol{\mathcal{V}}$. Therefore, $\mathbf{K}$ is the quantum operator constructed, with the help of an appropriate symmetrization, from the product of the probability density and the velocity of the particle. This indeed corresponds to the vector current density of a classical fluid (it is well known, for example, that the electrical current density associated with a fluid of charged particles is equal to the product of the charge volume density and the velocity of the particles).

If the particle is placed in an electromagnetic field described by the potentials $U(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$, we can use the preceding argument, starting with the Hamiltonian (B-46). We then find, in this case:

$$\mathbf{J}(\mathbf{r},t) = \frac{1}{m}\operatorname{Re}\left\{\psi^*\left[\frac{\hbar}{i}\boldsymbol{\nabla}-q\mathbf{A}\right]\psi\right\} \qquad \text{(D-20)}$$

We see that this expression can be obtained from (D-17), using the same rule as was used for the Hamiltonian: $\mathbf{P}$ is simply replaced by $\mathbf{P}-q\mathbf{A}$.

Example of a plane wave. Consider a wave function of the form:

$$\psi(\mathbf{r},t) = A\,\mathrm{e}^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \qquad \text{(D-21)}$$

with: $\hbar\omega = \dfrac{\hbar^2\mathbf{k}^2}{2m}$. The corresponding probability density:

$$\rho(\mathbf{r},t) = |\psi(\mathbf{r},t)|^2 = |A|^2 \qquad \text{(D-22)}$$

is uniform throughout all space and does not depend on time. The calculation of $\mathbf{J}(\mathbf{r},t)$ from (D-17) presents no difficulties and leads to:

$$\mathbf{J}(\mathbf{r},t) = |A|^2\frac{\hbar\mathbf{k}}{m} = \rho(\mathbf{r},t)\,\mathbf{v}_G \qquad \text{(D-23)}$$

where $\mathbf{v}_G = \dfrac{\hbar\mathbf{k}}{m}$ is the group velocity associated with the momentum $\hbar\mathbf{k}$ (Chap. I, § C-4). We see that the probability current is indeed equal to the product of the probability density and the group velocity of the particle. In this case, ρ and $\mathbf{J}$ are time-independent: the flow of the probability fluid associated with a plane wave is in a *steady state* condition (since ρ and $\mathbf{J}$ do not depend on $\mathbf{r}$ either, this state is also homogeneous and uniform).

D-1-d. Evolution of the mean value of an observable; relationship with classical mechanics

Let A be an observable. If the state $|\psi(t)\rangle$ of the system is normalized (and we have just seen that this normalization is conserved for all t), the mean value of the observable A at the instant t is equal to[5]:

$$\langle A\rangle(t) = \langle\psi(t)|A|\psi(t)\rangle \tag{D-24}$$

We see that $\langle A\rangle(t)$ depends on t through $|\psi(t)\rangle$ [and $\langle\psi(t)|$], which evolve over time according to the Schrödinger equation (D-4) [and (D-5)]. Moreover, the observable A may depend explicitly on time, causing an additional variation of the mean value $\langle A\rangle(t)$ with respect to t.

We intend to study, in this section, the evolution of $\langle A\rangle(t)$ and to show how this enables us to relate classical mechanics to quantum mechanics.

α. General formula

Differentiating (D-24) with respect to t, we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|A(t)|\psi(t)\rangle = \left[\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\right]A(t)|\psi(t)\rangle + \langle\psi(t)|\,A(t)\left[\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle\right] + \langle\psi(t)|\frac{\partial A}{\partial t}|\psi(t)\rangle \tag{D-25}$$

Using (D-4) and (D-5) for $\dfrac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle$ and $\dfrac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|$, we find:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|A(t)|\psi(t)\rangle = \frac{1}{i\hbar}\langle\psi(t)|\left[A(t)\,H(t) - H(t)\,A(t)\right]|\psi(t)\rangle + \langle\psi(t)|\frac{\partial A}{\partial t}|\psi(t)\rangle \tag{D-26}$$

that is:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle = \frac{1}{i\hbar}\langle[A, H(t)]\rangle + \langle\frac{\partial A}{\partial t}\rangle} \tag{D-27}$$

Comment:

The mean value $\langle A\rangle$ is a number which depends only on t. It is essential to understand how this dependence arises. For example, consider the case of a spinless

[5] The notation $\langle A\rangle(t)$ means that the mean value $\langle A\rangle$ of A is a number which depends on t.

particle. Let $\mathcal{A}(\mathbf{r}, \mathbf{p}, t)$ be a classical quantity. In classical mechanics, $\mathbf{r}$ and $\mathbf{p}$ depend on time (they evolve according to the Hamilton equations), so that $\mathcal{A}(\mathbf{r}, \mathbf{p}, t)$ depends on t explicitly, and implicitly through $\mathbf{r}$ and $\mathbf{p}$. To the classical quantity $\mathcal{A}(\mathbf{r}, \mathbf{p}, t)$ corresponds the Hermitian operator $A = \mathcal{A}(\mathbf{R}, \mathbf{P}, t)$, obtained by replacing, in $\mathcal{A}$, $\mathbf{r}$ and $\mathbf{p}$ by the operators $\mathbf{R}$ and $\mathbf{P}$, and by symmetrizing the operators if necessary (quantization rules, see § B-5). The eigenstates and eigenvalues of $\mathbf{R}$ and $\mathbf{P}$ and, consequently, these observables themselves, no longer depend on t. *The time dependence of* $\mathbf{r}$ *and* $\mathbf{p}$, which characterizes the time evolution of the classical state, *no longer appears in* $\mathbf{R}$ *and* $\mathbf{P}$, *but in the quantum state vector* $|\psi(t)\rangle$, associated in the $\{|\mathbf{r}\rangle\}$ representation with the wave function $\psi(\mathbf{r}, t) = \langle \mathbf{r}|\psi(t)\rangle$. In this representation, the mean value of A is written:

$$\langle A \rangle = \int \mathrm{d}^3 r \; \psi^*(\mathbf{r}, t) \, \mathcal{A}\left(\mathbf{r}, \frac{\hbar}{i}\boldsymbol{\nabla}, t\right) \psi(\mathbf{r}, t) \tag{D-28}$$

It is clear that integration over $\mathbf{r}$ leads to a number that only depends on t. With regard to classical mechanics, it is this number [and not the operator $\mathcal{A}(\mathbf{r}, \frac{\hbar}{i}\boldsymbol{\nabla}, t)$] that must be compared with the value taken on by the classical quantity $\mathcal{A}(\mathbf{r}, \mathbf{p}, t)$ at time t (*cf.* § D-1-d-γ below).

β. *Application to the observables* $\mathbf{R}$ *and* $\mathbf{P}$ *(Ehrenfest's theorem)*

Now let us apply the general formula (D-27) to the observables $\mathbf{R}$ and $\mathbf{P}$. We shall consider, for simplicity, the case of a spinless particle in a scalar stationary potential $V(\mathbf{r})$. We then have:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R}) \tag{D-29}$$

so that we can write:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{R} \rangle = \frac{1}{i\hbar}\langle [\mathbf{R}, H] \rangle = \frac{1}{i\hbar}\langle \left[\mathbf{R}, \frac{\mathbf{P}^2}{2m}\right] \rangle \tag{D-30}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{P} \rangle = \frac{1}{i\hbar}\langle [\mathbf{P}, H] \rangle = \frac{1}{i\hbar}\langle [\mathbf{P}, V(\mathbf{R})] \rangle \tag{D-31}$$

The commutator that appears in (D-30) can easily be calculated from the canonical commutation relations; we obtain:

$$\left[\mathbf{R}, \frac{\mathbf{P}^2}{2m}\right] = \frac{i\hbar}{m}\mathbf{P} \tag{D-32}$$

For the one in formula (D-31), the following generalization of formula (B-33) must be used [*cf.* Complement $\mathrm{B_{II}}$, formula (48)]:

$$[\mathbf{P}, V(\mathbf{R})] = -i\hbar \, \boldsymbol{\nabla} V(\mathbf{R}) \tag{D-33}$$

where $\boldsymbol{\nabla} V(\mathbf{R})$ denotes the set of three operators obtained by replacing $\mathbf{r}$ by $\mathbf{R}$ in the three components of the gradient of the function $V(\mathbf{r})$. Therefore:

$$\boxed{\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{R} \rangle &= \frac{1}{m}\langle \mathbf{P} \rangle \\ \frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{P} \rangle &= -\langle \boldsymbol{\nabla} V(\mathbf{R}) \rangle \end{aligned}} \qquad \begin{matrix} \text{(D-34)} \\ \text{(D-35)} \end{matrix}$$

These two equations express *Ehrenfest's theorem*. Their form recalls that of the classical Hamilton-Jacobi equations for a particle (Appendix III, § 3):

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{r} = \frac{1}{m}\mathbf{p} \tag{D-36a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p} = -\boldsymbol{\nabla} V(\mathbf{r}) \tag{D-36b}$$

which reduce, in this simple case, to Newton's well-known equation:

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = -\boldsymbol{\nabla} V(\mathbf{r}) \tag{D-37}$$

γ. *Discussion of Ehrenfest's theorem; classical limit*

Let us analyze the physical meaning of Ehrenfest's theorem, that is, equations (D-34) and (D-35). We shall assume that the wave function $\psi(\mathbf{r}, t)$ describing the state of the particle is a wave packet like those studied in Chapter I. $\langle \mathbf{R} \rangle$ then represents a set of three time-dependent numbers $\{\langle X \rangle, \langle Y \rangle, \langle Z \rangle\}$. We shall call the point $\langle \mathbf{R} \rangle(t)$ the *center of the wave packet*[6] at the instant t. The set of points corresponding to various values of t constitutes the *trajectory of the center of the wave packet*. Recall, however, that one can never rigorously speak of the trajectory of the particle itself, whose state is described by the wave packet as a whole, which inevitably has a certain spatial extension. We see, nevertheless, that if this extension is much smaller than the other distances involved in the problem, we can approximate the wave packet by its center. In this limiting case, there is no appreciable difference between the quantum and classical descriptions of the particle.

It is therefore important to know the answer to the following question: does the motion of the center of the wave packet obey the laws of classical mechanics? This answer is supplied by Ehrenfest's theorem. Equation (D-34) expresses the fact that the velocity of the center of the wave packet is equal to the average momentum of this wave packet divided by m. Consequently, the left-hand side of (D-35) can be written $m\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}\langle \mathbf{R} \rangle$, so that the answer to the preceding question will be affirmative if the right-hand side of (D-35) is equal to the classical force $\mathbf{F}_{cl}$ at the point where the center of the wave packet is situated:

$$\mathbf{F}_{cl} = [-\boldsymbol{\nabla} V(\mathbf{r})]_{\mathbf{r}=\langle \mathbf{R} \rangle} \tag{D-38}$$

[6] The center and the maximum of a wave packet are, in general, distinct. They coincide, however, if the wave packet has a symmetrical shape (§ C-5, Fig. 2).

In fact, the right-hand side of (D-35) is equal to the average of the force over the whole wave packet, and, in general:

$$\langle \boldsymbol{\nabla} V(\mathbf{R}) \rangle \neq [\boldsymbol{\nabla} V(\mathbf{r})]_{\mathbf{r}=\langle \mathbf{R} \rangle} \tag{D-39}$$

(in other words, the mean value of a function is not equal to its value for the mean value of the variable). Strictly speaking, the answer to the question we asked is therefore negative.

Comment:

It is easy to convince ourselves of (D-39) if we consider a concrete example. Let us choose, for simplicity, a one-dimensional model, and assume that:

$$V(x) = \lambda x^n \tag{D-40}$$

where λ is a real constant and n, a positive integer. From this we deduce the operator associated with $V(x)$:

$$V(X) = \lambda X^n \tag{D-41}$$

The left-hand side of (D-39) can be written (replacing $\boldsymbol{\nabla}$ by $\dfrac{\mathrm{d}}{\mathrm{d}x}$) $\lambda n \langle X^{n-1} \rangle$. As for the right-hand side, it is equal to:

$$\left[\frac{\mathrm{d}V}{\mathrm{d}x}\right]_{x=\langle X \rangle} = [\lambda n\, x^{n-1}]_{x=\langle X \rangle} = \lambda n \langle X \rangle^{n-1} \tag{D-42}$$

Now we know that in general $\langle X^{n-1} \rangle \neq \langle X \rangle^{n-1}$; for example, for $n = 3$, we have $\langle X^2 \rangle \neq \langle X \rangle^2$ (since the difference between these two quantities enters into the calculation of the root mean square deviation ΔX).

Note however that for $n = 1$ or 2, $\langle X^{n-1} \rangle = \langle X \rangle^{n-1}$. The two sides of (D-39) are then equal. The same holds true, moreover, for $n = 0$, in which case both sides are equal to zero. For a free particle ($n = 0$), or a particle placed in a uniform force field ($n = 1$) or in a parabolic potential well ($n = 2$, the case of a harmonic oscillator), the motion of the center of the wave packet therefore rigorously obeys the laws of classical mechanics. We have already established this result for the free particle ($n = 0$) (*cf.* Chapter I, § C-4).

Although the two sides of (D-39) are not, in general, equal, there exist situations (called quasi-classical) where the difference between these two quantities is negligible: this is the case when the wave packet is sufficiently localized. To see this, let us write explicitly, in the $\{|\mathbf{r}\rangle\}$ representation, the left-hand side of this equation:

$$\begin{aligned} \langle \boldsymbol{\nabla} V(\mathbf{R}) \rangle &= \int \mathrm{d}^3 r \; \psi^*(\mathbf{r}, t) \; [\boldsymbol{\nabla} V(\mathbf{r})] \; \psi(\mathbf{r}, t) \\ &= \int \mathrm{d}^3 r \; |\psi(\mathbf{r}, t)|^2 \boldsymbol{\nabla} V(\mathbf{r}) \end{aligned} \tag{D-43}$$

Let us assume the wave packet to be highly localized: more precisely, $|\psi(\mathbf{r}, t)|^2$ takes on non-negligible values only within a domain whose dimensions are much smaller than the distances over which $V(\mathbf{r})$ varies appreciably. Then, within this domain, centered about

$\langle \mathbf{R} \rangle$, $\boldsymbol{\nabla} V(\mathbf{r})$ is practically constant. Therefore, in (D-43), $\boldsymbol{\nabla} V(\mathbf{r})$ can be replaced by its value for $\mathbf{r} = \langle \mathbf{R} \rangle$ and taken outside the integral, which is then equal to 1, since $\psi(\mathbf{r}, t)$ is normalized. Thus we find that for sufficiently localized wave packets:

$$\langle \boldsymbol{\nabla} V(\mathbf{R}) \rangle \simeq [\boldsymbol{\nabla} V(\mathbf{r})]_{\mathbf{r}=\langle \mathbf{R} \rangle} \tag{D-44}$$

In the macroscopic limit (where the de Broglie wavelengths are much smaller than the distances over which the potential varies[7]), wave packets can be made sufficiently small to satisfy (D-44) while retaining a good degree of definition for the momentum. The motion of the wave packet is then practically that of a classical particle of mass m placed in the potential $V(\mathbf{r})$. The result that we have thus established is very important since it enables us to show that the equations of classical mechanics follow from the Schrödinger equation in certain limiting conditions satisfied, in particular, by most macroscopic systems.

D-2. The case of conservative systems

When the Hamiltonian of a physical system does not depend explicitly on time, the system is said to be *conservative*. In classical mechanics, the most important consequence of such a situation is the *conservation of energy* over time. It can also be said that the total energy of the system is a *constant of the motion*. We shall see in this section that in quantum mechanics as well, conservative systems possess important special properties in addition to the general properties of the preceding section.

D-2-a. Solution of the Schrödinger equation

First, let us consider the eigenvalue equation for H:

$$H|\varphi_{n,\tau}\rangle = E_n|\varphi_{n,\tau}\rangle \tag{D-45}$$

For simplicity, we assume the spectrum of H to be discrete; τ denotes the set of indices other than n that are necessary for characterizing a unique vector $|\varphi_{n,\tau}\rangle$ (in general, these indices will determine the eigenvalues of operators forming a C.S.C.O. with H). Since, by hypothesis, H does not depend explicitly on time, neither the eigenvalue E_n nor the eigenket $|\varphi_{n,\tau}\rangle$ is t dependent.

First, we are going to show that, given the E_n and the $|\varphi_{n,\tau}\rangle$, it is very simple to solve the Schrödinger equation, that is, to determine the time evolution of any state. Since the $|\varphi_{n,\tau}\rangle$ form a basis (H is an observable), it is always possible, for every value of t, to expand any state $|\psi(t)\rangle$ of the system in terms of the $|\varphi_{n,\tau}\rangle$:

$$|\psi(t)\rangle = \sum_{n,\tau} c_{n,\tau}(t)|\varphi_{n,\tau}\rangle \tag{D-46}$$

with:

$$c_{n,\tau}(t) = \langle \varphi_{n,\tau}|\psi(t)\rangle \tag{D-47}$$

Since the $|\varphi_{n,\tau}\rangle$ do not depend on t, all the time dependence of $|\psi(t)\rangle$ is contained within the $c_{n,\tau}(t)$. To calculate the $c_{n,\tau}(t)$, let us project the Schrödinger equation onto each

[7] See the order of magnitude of the de Broglie wavelengths associated with a macroscopic system in Complement A_I.

of the states $|\varphi_{n,\tau}\rangle$. This yields[8]:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle\varphi_{n,\tau}|\psi(t)\rangle = \langle\varphi_{n,\tau}|H|\psi(t)\rangle \tag{D-48}$$

Since H is Hermitian, it can be deduced from (D-45) that:

$$\langle\varphi_{n,\tau}|H = E_n\langle\varphi_{n,\tau}| \tag{D-49}$$

so that (D-48) can be written in the form:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}c_{n,\tau}(t) = E_n\, c_{n,\tau}(t) \tag{D-50}$$

This equation can be integrated directly to give:

$$c_{n,\tau}(t) = c_{n,\tau}(t_0)\,\mathrm{e}^{-iE_n(t-t_0)/\hbar} \tag{D-51}$$

When H does not depend explicitly on time, to find $|\psi(t)\rangle$, given $|\psi(t_0)\rangle$, proceed as follows:

(i) Expand $|\psi(t_0)\rangle$ in terms of a basis of eigenstates of H:

$$|\psi(t_0)\rangle = \sum_n\sum_\tau c_{n,\tau}(t_0)|\varphi_{n,\tau}\rangle \tag{D-52}$$

$c_{n,\tau}(t_0)$ is given by the usual formula:

$$c_{n,\tau}(t_0) = \langle\varphi_{n,\tau}|\psi(t_0)\rangle \tag{D-53}$$

(ii) Now, to obtain $|\psi(t)\rangle$ for arbitrary t, multiply each coefficient $c_{n,\tau}(t_0)$ of the expansion (D-52) by $\mathrm{e}^{-iE_n(t-t_0)/\hbar}$, where E_n is the eigenvalue of H associated with the state $|\varphi_{n,\tau}\rangle$:

$$|\psi(t)\rangle = \sum_n\sum_\tau c_{n,\tau}(t_0)\,\mathrm{e}^{-iE_n(t-t_0)/\hbar}|\varphi_{n,\tau}\rangle \tag{D-54}$$

The preceding argument can easily be generalized to the case where the spectrum of H is continuous; formula (D-54) then becomes, with obvious notation:

$$|\psi(t)\rangle = \sum_\tau\int \mathrm{d}E\; c_\tau(E,t_0)\;\mathrm{e}^{-iE(t-t_0)/\hbar}\;|\varphi_{E,\tau}\rangle \tag{D-55}$$

D-2-b. Stationary states

An important special case is that in which $|\psi(t_0)\rangle$ is itself an eigenstate of H. Expansion (D-52) of $|\psi(t_0)\rangle$ then involves only eigenstates of H with the same eigenvalue (for example, E_n):

$$|\psi(t_0)\rangle = \sum_\tau c_{n,\tau}(t_0)\;|\varphi_{n,\tau}\rangle \tag{D-56}$$

[8] In (D-48), $\langle\varphi_{n,\tau}|$ can be placed to the right of $\frac{\mathrm{d}}{\mathrm{d}t}$, since $\langle\varphi_{n,\tau}|$ does not depend on t.

In formula (D-56), there is no summation over n, and the passage from $|\psi(t_0)\rangle$ to $|\psi(t)\rangle$ involves only one factor $\mathrm{e}^{-iE_n(t-t_0)/\hbar}$, which can be taken outside the summation over τ:

$$\begin{aligned} |\psi(t)\rangle &= \sum_{\tau} c_{n,\tau}(t_0)\ \mathrm{e}^{-iE_n(t-t_0)/\hbar}\ |\varphi_{n,\tau}\rangle \\ &= \mathrm{e}^{-iE_n(t-t_0)/\hbar} \sum_{\tau} c_{n,\tau}(t_0)\ |\varphi_{n,\tau}\rangle \\ &= \mathrm{e}^{-iE_n(t-t_0)/\hbar}\ |\psi(t_0)\rangle \end{aligned} \tag{D-57}$$

$|\psi(t)\rangle$ and $|\psi(t_0)\rangle$ therefore differ only by the *global* phase factor $\mathrm{e}^{-iE_n(t-t_0)/\hbar}$. These two states are physically indistinguishable (*cf.* discussion in § B-3-b-γ). From this we conclude that all the physical properties of a system which is in an eigenstate of H do not vary over time; the eigenstates of H are called, for this reason, *stationary states.*

It is also interesting to see how conservation of energy in a conservative system appears in quantum mechanics. Let us assume that, at time t_0, we measure the energy of such a system and we find, for example: E_k. Immediately after the measurement, the system is in an eigenstate of H, with an eigenvalue of E_k (the postulate of the reduction of the wave packet). We have just seen that the eigenstates of H are stationary states. Therefore, the state of the system will no longer evolve after the first measurement and will always remain an eigenstate of H with an eigenvalue of E_k. It follows that a second measurement of the energy of the system, at any subsequent time t, will always yield the same result E_k as the first one.

Comment:

One passes from (D-52) to (D-54) by multiplying each coefficient $c_{n,\tau}(t_0)$ of (D-52) by $\mathrm{e}^{-iE_n(t-t_0)/\hbar}$. The fact that $\mathrm{e}^{-iE_n(t-t_0)/\hbar}$ is a phase factor should not lead us to believe that $|\psi(t)\rangle$ and $|\psi(t_0)\rangle$ are always physically indistinguishable. Actually, expansion (D-52) involves, in general, *several eigenstates of H with different eigenvalues.* To these different possible values of E_n correspond *different phase factors.* This modifies the *relative phases* of the expansion coefficients of the state vector and leads, consequently, to a state $|\psi(t)\rangle$ which is physically distinct from $|\psi(t_0)\rangle$.

Only in the case where a single value of n enters into (D-52) [the case where $|\psi(t_0)\rangle$ is an eigenstate of H] is the time evolution described by a single phase factor, which is then a global one, of no physical importance. In other words, there is physical evolution over time only if the energy of the initial state is not known with certainty. We shall come back later to the relation between time evolution and energy uncertainty (*cf.* § D-2-e).

D-2-c. Constants of the motion

By definition, a constant of the motion is an observable A which does not depend explicitly on time and which commutes with H:

$$\begin{cases} \dfrac{\partial A}{\partial t} = 0 \\ \\ [A, H] = 0 \end{cases} \tag{D-58}$$

For a conservative system, H is therefore itself a constant of the motion.

Constants of the motion possess important properties which we are now going to derive.

(i) If we substitute (D-58) into the general formula (D-27), we find:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|A|\psi(t)\rangle = 0 \tag{D-59}$$

Whatever the state $|\psi(t)\rangle$ of the physical system, *the mean value of* A in this state *does not evolve over time* (hence the term "constant of the motion").

(ii) Since A and H are two observables which commute, we can always find for them a system of common eigenvectors, which we shall denote by $\{|\varphi_{n,p,\tau}\rangle\}$:

$$\begin{aligned} H|\varphi_{n,p,\tau}\rangle &= E_n\ |\varphi_{n,p,\tau}\rangle \\ A|\varphi_{n,p,\tau}\rangle &= a_p\ |\varphi_{n,p,\tau}\rangle \end{aligned} \tag{D-60}$$

We shall assume for simplicity that the spectra of H and A are discrete. The index τ fixes the eigenvalues of observables which form a C.S.C.O. with H and A. Since the states $|\varphi_{n,p,\tau}\rangle$ are eigenstates of H, they are stationary states. If the system is in the state $|\varphi_{n,p,\tau}\rangle$ at the initial instant, it will therefore remain there indefinitely (to within a global phase factor). But the state $|\varphi_{n,p,\tau}\rangle$ is an eigenstate of A as well. Consequently, when A is a constant of the motion, there exist stationary states of the physical system (the states $|\varphi_{n,p,\tau}\rangle$) that always remain, for all t, eigenstates of A with the same eigenvalue a_p. The eigenvalues of A are called, for this reason, *good quantum numbers*.

(iii) Finally, let us show that for an arbitrary state $|\psi(t)\rangle$, *the probability of finding the eigenvalue* a_p, when the constant of the motion A is measured, *is not time-dependent*. $|\psi(t_0)\rangle$ can always be expanded on the $\{|\varphi_{n,p,\tau}\rangle\}$ basis introduced above:

$$|\psi(t_0)\rangle = \sum_n \sum_p \sum_\tau c_{n,p,\tau}(t_0)\ |\varphi_{n,p,\tau}\rangle \tag{D-61}$$

From this we directly deduce:

$$|\psi(t)\rangle = \sum_n \sum_p \sum_\tau c_{n,p,\tau}(t)\ |\varphi_{n,p,\tau}\rangle \tag{D-62}$$

with:

$$c_{n,p,\tau}(t) = c_{n,p,\tau}(t_0)\ \mathrm{e}^{-iE_n(t-t_0)/\hbar} \tag{D-63}$$

According to the postulate of spectral decomposition, the probability $\mathcal{P}(a_p, t_0)$ of finding a_p when A is measured at time t_0, on the system in the state $|\psi(t_0)\rangle$, is equal to:

$$\mathcal{P}(a_p, t_0) = \sum_n \sum_\tau |c_{n,p,\tau}(t_0)|^2 \tag{D-64}$$

Similarly:

$$\mathcal{P}(a_p, t) = \sum_n \sum_\tau |c_{n,p,\tau}(t)|^2 \tag{D-65}$$

Now we see from (D-63) that $c_{n,p,\tau}(t)$ and $c_{n,p,\tau}(t_0)$ have the same modulus. Therefore, $\mathcal{P}(a_p, t) = \mathcal{P}(a_p, t_0)$, which proves the property stated above.

Comment:

If all but one of the probabilities $\mathcal{P}(a_p, t_0)$ are zero [leaving for example $\mathcal{P}(a_k, t_0)$ non-zero and, moreover, necessarily equal to 1], the physical system at time t_0 is in an eigenstate of A with an eigenvalue of a_k. Since the $\mathcal{P}(a_p, t)$ do not depend on t, the state of the system at time t remains an eigenstate of A with an eigenvalue of a_k.

D-2-d. Bohr frequencies of a system. Selection rules

Let B be an arbitrary observable of the system under consideration (it does not necessarily commute with H). Formula (D-27) enables us to calculate the derivative $\dfrac{\mathrm{d}}{\mathrm{d}t}\langle B\rangle$ of the mean value of B:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle B\rangle = \frac{1}{i\hbar}\langle [B, H]\rangle + \langle \frac{\partial B}{\partial t}\rangle \tag{D-66}$$

For a conservative system, we know the general form (D-54) of $|\psi(t)\rangle$. Therefore, in this case, we can calculate explicitly $\langle\psi(t)|B|\psi(t)\rangle$ (and not merely $\dfrac{\mathrm{d}}{\mathrm{d}t}\langle B\rangle$).

The Hermitian conjugate expression of (D-54) is written (changing the summation indices):

$$\langle\psi(t)| = \sum_{n'} \sum_{\tau'} c^*_{n',\tau'}(t_0)\, \mathrm{e}^{iE_{n'}(t-t_0)/\hbar}\, \langle\varphi_{n',\tau'}| \tag{D-67}$$

We can then, in $\langle\psi(t)|B|\psi(t)\rangle$, replace $|\psi(t)\rangle$ and $\langle\psi(t)|$ by expansions (D-54) and (D-67), respectively. Thus we obtain:

$$\begin{aligned}\langle\psi(t)|B|\psi(t)\rangle &= \langle B\rangle(t) \\ &= \sum_n \sum_\tau \sum_{n'} \sum_{\tau'} c^*_{n',\tau'}(t_0)\, c_{n,\tau}(t_0) \langle\varphi_{n',\tau'}|B|\varphi_{n,\tau}\rangle\, \mathrm{e}^{i(E_{n'}-E_n)(t-t_0)/\hbar}\end{aligned} \tag{D-68}$$

From now on, we shall assume that B does not depend explicitly on time: the matrix elements $\langle\varphi_{n',\tau'}|B|\varphi_{n,\tau}\rangle$ are therefore constant. Formula (D-68) then shows that the evolution of $\langle B\rangle(t)$ is described by a series of *oscillating terms*, whose frequencies $\frac{1}{2\pi}\frac{|E_{n'}-E_n|}{\hbar} = \left|\frac{E_{n'}-E_n}{h}\right| = \nu_{n'n}$ are characteristic of the system under consideration but independent of B and of the initial state of the system. The frequencies $\nu_{n'n}$ are called the *Bohr frequencies* of the system. Thus, for an atom, the mean values of all the atomic quantities (electric and magnetic dipole moments, etc...) oscillate at the various Bohr frequencies of the atom. It is reasonable to imagine that only these frequencies can be radiated or absorbed by the atom. This remark allows us to understand intuitively the Bohr relation between the spectral frequencies emitted or absorbed and the differences in atomic energies.

It can also be seen from (D-68) that, while the frequencies involved in the motion of $\langle B\rangle(t)$ are independent of B, the same does not hold true for the respective *weights* of these frequencies in the variation of $\langle B\rangle$. The importance of each frequency $\nu_{n'n}$ depends on the matrix elements $\langle\varphi_{n',\tau'}|B|\varphi_{n,\tau}\rangle$. In particular, if these matrix elements are zero for certain values of n and n', the corresponding frequencies $\nu_{n'n}$ are absent from the expansion of $\langle B\rangle(t)$, whatever the initial state of the system. This is the origin of the *selection rules* which indicate what frequencies can be emitted or absorbed under given conditions. To establish these rules, one must study the non-diagonal matrix elements ($n \neq n'$) of the various atomic operators such as the electric and magnetic dipoles, etc...

Finally, the weights of the various Bohr frequencies also depend on the initial state, via $c^*_{n',\tau'}(t_0)\, c_{n,\tau}(t_0)$. In particular, if the initial state is a stationary state of energy E_k, the expansion of $|\psi(t_0)\rangle$ contains only one value of n ($n = k$) and $c^*_{n',\tau'}(t_0)\, c_{n,\tau}(t_0)$ can be non-zero only for $n = n' = k$. In this case, $\langle B\rangle$ is not time-dependent.

Comment:

It can be directly verified, using (D-68), that the mean value of a constant of the motion is always time-independent. We see that if B commutes with H, the matrix elements of B are zero between two eigenstates of H that correspond to different eigenvalues (*cf.* Chap. II, § D-3-a). It follows that $\langle\varphi_{n',\tau'}|B|\varphi_{n,\tau}\rangle$ is zero for $n' \neq n$. The only terms of $\langle B\rangle$ that are non-zero are thus constant.

D-2-e. The time-energy uncertainty relation

We shall now see that for a conservative system, the greater the energy uncertainty, the more rapid the time evolution. More precisely, if Δt is a time interval at the end of which the system has evolved to an appreciable extent, and if ΔE denotes the energy uncertainty, Δt and ΔE satisfy the relation:

$$\boxed{\Delta t \cdot \Delta E \gtrsim h} \tag{D-69}$$

First, if the system is in an eigenstate of H, its energy is perfectly well-defined: $\Delta E = 0$. But we have seen that such a state is stationary, meaning that it does not evolve in time; in a sense, its evolution time Δt is infinite [relation (D-69) indicates that when $\Delta E = 0$, Δt must be infinite].

Now let us assume that $|\psi(t_0)\rangle$ is a linear superposition of two eigenstates of H, $|\varphi_1\rangle$ and $|\varphi_2\rangle$, with different eigenvalues E_1 and E_2:

$$|\psi(t_0)\rangle = c_1|\varphi_1\rangle + c_2|\varphi_2\rangle \tag{D-70}$$

Then:

$$|\psi(t)\rangle = c_1\,\mathrm{e}^{-iE_1(t-t_0)/\hbar}|\varphi_1\rangle + c_2\,\mathrm{e}^{-iE_2(t-t_0)/\hbar}|\varphi_2\rangle \tag{D-71}$$

If we measure the energy, we find either E_1 or E_2. The uncertainty of E is therefore of the order of:

$$\Delta E \simeq |E_2 - E_1| \tag{D-72}$$

Now consider an arbitrary observable B which does not commute with H. The probability of finding, in a measurement of B at time t, the eigenvalue b_m associated with the eigenvector $|u_m\rangle$ (we assume, for simplicity, b_m to be non-degenerate) is given by:

$$\begin{aligned}\mathcal{P}(b_m,t) = |\langle u_m|\psi(t)\rangle|^2 &= |c_1|^2|\langle u_m|\varphi_1\rangle|^2 + |c_2|^2|\langle u_m|\varphi_2\rangle|^2 \\ &+ 2\,\mathrm{Re}\left[c_2^* c_1\,\mathrm{e}^{i(E_2-E_1)(t-t_0)/\hbar}\,\langle u_m|\varphi_2\rangle^*\langle u_m|\varphi_1\rangle\right]\end{aligned} \tag{D-73}$$

This equation shows that $\mathcal{P}(b_m,t)$ oscillates between two extreme values, with the Bohr frequency $\nu_{21} = \dfrac{|E_2-E_1|}{h}$. The characteristic evolution time of the system is therefore:

$$\Delta t \simeq \frac{h}{|E_2 - E_1|} \tag{D-74}$$

and comparison with (D-72) shows that: $\Delta E \cdot \Delta t \simeq h$.

Let us now assume that the spectrum of H is continuous (and non-degenerate). The most general state $|\psi(t_0)\rangle$ can be written:

$$|\psi(t_0)\rangle = \int \mathrm{d}E\; c(E)\;|\varphi_E\rangle \tag{D-75}$$

where $|\varphi_E\rangle$ is the eigenstate of H with the eigenvalue E. Let us assume that $|c(E)|^2$ has non-negligible values only in a domain of width ΔE about E_0 (Fig. 4). ΔE then represents the energy uncertainty of the system. $|\psi(t)\rangle$ is obtained by using (D-55):

$$|\psi(t)\rangle = \int \mathrm{d}E\; c(E)\;\mathrm{e}^{-iE(t-t_0)/\hbar}\,|\varphi_E\rangle \tag{D-76}$$

The quantity $\mathcal{P}(b_m,t)$ introduced above, which represents the probability of finding the eigenvalue b_m when the observable B is measured on the system in the state $|\psi(t)\rangle$, is here equal to:

$$\mathcal{P}(b_m,t) = |\langle u_m|\psi(t)\rangle|^2 = \left|\int \mathrm{d}E\; c(E)\;\mathrm{e}^{-iE(t-t_0)/\hbar}\,\langle u_m|\varphi_E\rangle\right|^2 \tag{D-77}$$

In general, $\langle u_m|\varphi_E\rangle$ does not vary rapidly with E when E varies about E_0. If ΔE is sufficiently small, the variation of $\langle u_m|\varphi_E\rangle$, in integral (D-77), can be neglected relative

to that of $c(E)$. One can then replace $\langle u_m|\varphi_E\rangle$ by $\langle u_m|\varphi_{E_0}\rangle$ and take this quantity outside integral (D-77):

$$\mathcal{P}(b_m,t) \simeq |\langle u_m|\varphi_{E_0}\rangle|^2 \left| \int \mathrm{d}E\ c(E)\ \mathrm{e}^{-iE(t-t_0)/\hbar} \right|^2 \tag{D-78}$$

If this approximation is valid, we thus see that $\mathcal{P}(b_m,t)$ is, to within a coefficient, the square of the modulus of the Fourier transform of $c(E)$. According to the properties of the Fourier transform (*cf.* Appendix I, § 2-b), the width in t of $\mathcal{P}(b_m,t)$, that is, Δt, is therefore related to the width ΔE of $|c(E)|^2$ by relation (D-69).

Comment:

(D-69) can be established directly for a free one-dimensional wave packet. One can associate with the momentum uncertainty Δp of this wave packet an energy uncertainty $\Delta E = \dfrac{\mathrm{d}E}{\mathrm{d}p}\Delta p$. Since $E = \hbar\omega$ and $p = \hbar k$, we have $\dfrac{\mathrm{d}E}{\mathrm{d}p} = \dfrac{\mathrm{d}\omega}{\mathrm{d}k} = v_G$, where v_G is the group velocity of the wave packet (Chap. I, § C-4). Consequently:

$$\Delta E = v_G\,\Delta p \tag{D-79}$$

Now the characteristic evolution time Δt is the time taken by this wave packet, travelling at the velocity v_G, to "pass" a point in space. If Δx is the spatial extension of the wave packet, we therefore have:

$$\Delta t \simeq \frac{\Delta x}{v_G} \tag{D-80}$$

From this we deduce, combining (D-79) and (D-80):

$$\Delta E.\Delta t\ \simeq \Delta x.\Delta p \gtrsim \hbar \tag{D-81}$$

Relation (D-69) is often called *the fourth Heisenberg uncertainty relation.* It is clearly different, however, from the other three uncertainty relations which relate to the three components of $\mathbf{R}$ and $\mathbf{P}$ [formulas (14) of Complement $\mathrm{F_I}$]. In (D-69), only the energy is a physical quantity like $\mathbf{R}$ and $\mathbf{P}$; t, on the other hand, *is a parameter*, with which no quantum mechanical operator is associated.

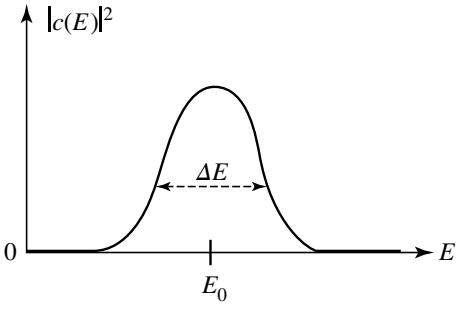


Figure 4: By superposing the stationary states $|\varphi_E\rangle$ with the coefficients $c(E)$, we obtain a state $|\psi\rangle$ of the system where the energy is not perfectly well-defined. The corresponding uncertainty ΔE is given by the width of the curve that represents $|c(E)|^2$. According to the fourth uncertainty relation, the evolution of the state $|\psi(t)\rangle$ will be significant after a time Δt such that $\Delta E \cdot \Delta t \gtrsim \hbar$.

E. The superposition principle and physical predictions

The physical meaning of the first postulate remains to be examined. According to this postulate, the states of a physical system belong to a vector space and are, consequently, linearly superposable.

One of the important consequences of the first postulate, when it is combined with the others, is the appearance of interference effects such as those which led us to wave-particle duality (Chap. I). Our understanding of these phenomena is based on the concept of probability amplitudes, which we shall examine here with the aid of some simple examples.

E-1. Probability amplitudes and interference effects

E-1-a. The physical meaning of a linear superposition of states

α. The difference between a linear superposition and a statistical mixture

Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be two orthogonal normalized states:

$$\begin{aligned} \langle\psi_1|\psi_1\rangle = \langle\psi_2|\psi_2\rangle = 1 \\ \langle\psi_1|\psi_2\rangle = 0 \end{aligned} \tag{E-1}$$

($|\psi_1\rangle$ and $|\psi_2\rangle$ could be, for example, two eigenstates of the same observable B associated with two different eigenvalues b_1 and b_2).

If the system is in the state $|\psi_1\rangle$, we can calculate all the probabilities concerning the measurement results for a given observable A. For example, if $|u_n\rangle$ is the (normalized) eigenvector of A which corresponds to the eigenvalue a_n (assumed to be non-degenerate), the probability $\mathcal{P}_1(a_n)$ of finding a_n when A is measured on the system in the state $|\psi_1\rangle$ is:

$$\mathcal{P}_1(a_n) = |\langle u_n|\psi_1\rangle|^2 \tag{E-2}$$

An analogous quantity $\mathcal{P}_2(a_n)$ can be defined for the state $|\psi_2\rangle$:

$$\mathcal{P}_2(a_n) = |\langle u_n|\psi_2\rangle|^2 \tag{E-3}$$

Now consider a normalized state $|\psi\rangle$ which is a *linear superposition* of $|\psi_1\rangle$ and $|\psi_2\rangle$:

$$\begin{aligned} |\psi\rangle = \lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle \\ |\lambda_1|^2 + |\lambda_2|^2 = 1 \end{aligned} \tag{E-4}$$

It is often said that, when the system is in the state $|\psi\rangle$, one has a probability $|\lambda_1|^2$ of finding it in the state $|\psi_1\rangle$ and a probability $|\lambda_2|^2$ of finding it in the state $|\psi_2\rangle$. The exact meaning of this manner of speaking is the following: if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two eigenvectors (here assumed to be normalized) of the observable B corresponding to different eigenvalues b_1 and b_2, the probability of finding b_1 when B is measured is $|\lambda_1|^2$ and that of finding b_2 is $|\lambda_2|^2$.

This could lead us to believe (wrongly, as we shall see), that a state such as (E-4) is a *statistical mixture* of the states $|\psi_1\rangle$ and $|\psi_2\rangle$ with the weights $|\lambda_1|^2$ and $|\lambda_2|^2$. In

other words, if we consider a large number N of identical systems, all in the state (E-4), we might imagine that this set of N systems in the state $|\psi\rangle$ was completely equivalent to another set composed of $N|\lambda_1|^2$ systems in the state $|\psi_1\rangle$ and $N|\lambda_2|^2$ systems in the state $|\psi_2\rangle$. Such an interpretation of the state $|\psi\rangle$ is *erroneous* and leads to inaccurate physical predictions, as we shall see.

Assume that we are actually trying to calculate the probability $\mathcal{P}(a_n)$ of finding the eigenvalue a_n when the observable A is measured on the system in the state$|\psi\rangle$ given by (E-4). If we interpret the state $|\psi\rangle$ as being a statistical mixture of the states $|\psi_1\rangle$ and $|\psi_2\rangle$ with the weights $|\lambda_1|^2$ and $|\lambda_2|^2$, then we can obtain $\mathcal{P}(a_n)$ by taking the weighted sum of the probabilities $\mathcal{P}_1(a_n)$ and $\mathcal{P}_2(a_n)$ calculated above [formulas (E-2) and (E-3)]:

$$\mathcal{P}(a_n) \stackrel{?}{=} |\lambda_1|^2 \mathcal{P}_1(a_n) + |\lambda_2|^2 \mathcal{P}_2(a_n) \tag{E-5}$$

Actually, the postulates of quantum mechanics unambiguously indicate how to calculate $\mathcal{P}(a_n)$. The correct expression for this probability is:

$$\mathcal{P}(a_n) = |\langle u_n|\psi\rangle|^2 \tag{E-6}$$

$\mathcal{P}(a_n)$ is therefore the square of the modulus of the *probability amplitude* $\langle u_n|\psi\rangle$. We see from (E-4) that this amplitude is the sum of two terms:

$$\langle u_n|\psi\rangle = \lambda_1 \langle u_n|\psi_1\rangle + \lambda_2 \langle u_n|\psi_2\rangle \tag{E-7}$$

Thus we obtain:

$$\begin{aligned} \mathcal{P}(a_n) &= |\lambda_1 \langle u_n|\psi_1\rangle + \lambda_2 \langle u_n|\psi_2\rangle|^2 \\ &= |\lambda_1|^2 |\langle u_n|\psi_1\rangle|^2 + |\lambda_2|^2 |\langle u_n|\psi_2\rangle|^2 \\ &\quad + 2\,\mathrm{Re}\left\{\lambda_1 \lambda_2^* \langle u_n|\psi_1\rangle \langle u_n|\psi_2\rangle^*\right\} \end{aligned} \tag{E-8}$$

Taking (E-2) and (E-3) into account, we find that the correct expression for $\mathcal{P}(a_n)$ is therefore written:

$$\mathcal{P}(a_n) = |\lambda_1|^2 \mathcal{P}_1(a_n) + |\lambda_2|^2 \mathcal{P}_2(a_n) + 2\,\mathrm{Re}\left\{\lambda_1 \lambda_2^* \langle u_n|\psi_1\rangle \langle u_n|\psi_2\rangle^*\right\} \tag{E-9}$$

This result is different from that of formula (E-5).

It is therefore wrong to consider $|\psi\rangle$ to be a statistical mixture of states. Such an interpretation eliminates all the *interference effects* contained in the double product of formula (E-9). We see that it is not only the moduli of λ_1 and λ_2 that play a role; the relative phase[9] of λ_1 and λ_2 is just as important, since it enters explicitly into the physical predictions, through the intermediary of $\lambda_1 \lambda_2^*$.

β. A concrete example

Consider photons propagating along Oz whose polarization state is represented by the unit vector (Fig. 5):

$$\mathbf{e} = \frac{1}{\sqrt{2}}(\mathbf{e}_x + \mathbf{e}_y) \tag{E-10}$$

[9] Multiplying $|\psi\rangle$ by a *global* phase factor $e^{i\theta}$ is equivalent to changing λ_1 and λ_2 to $\lambda_1 e^{i\theta}$ and $\lambda_2 e^{i\theta}$. It can be verified from (E-9) that such an operation does not modify the physical predictions, which depend only on $|\lambda_1|^2$, $|\lambda_2|^2$ and $\lambda_1 \lambda_2^*$.

This state is a linear superposition of two orthogonal polarization states $\mathbf{e}_x$ and $\mathbf{e}_y$. It represents light which is linearly polarized at an angle of 45° with respect to $\mathbf{e}_x$ and $\mathbf{e}_y$. It would be absurd to assume that N photons in the state $\mathbf{e}$ are equivalent to $N \times \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{N}{2}$ photons in the state $\mathbf{e}_x$ and $N \times \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{N}{2}$ photons in the state $\mathbf{e}_y$. If we place in the beam's trajectory an analyzer whose axis $\mathbf{e}'$ is perpendicular to $\mathbf{e}$, we know that *none* of the N photons in the state $\mathbf{e}$ will pass through this analyzer. But, for the statistical mixture $\left\{\frac{N}{2}$ photons in the state $\mathbf{e}_x$, $\frac{N}{2}$ photons in the state $\mathbf{e}_y\right\}$, half the photons will pass through the analyzer. In this concrete example, it is clear that a linear superposition

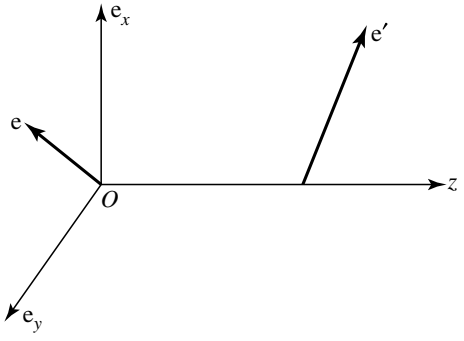


Figure 5: A simple experiment which illustrates the difference between a linear superposition and a statistical mixture of states. If all the incident photons are in the polarization state:

$$\mathbf{e} = \frac{1}{\sqrt{2}}(\mathbf{e}_x + \mathbf{e}_y)$$

none of them will pass through an analyzer whose axis $\mathbf{e}'$ *is perpendicular to* $\mathbf{e}$*. If we had, on the contrary, a statistical mixture of photons polarized either along* $\mathbf{e}_x$ *or along* $\mathbf{e}_y$ *(in equal proportions; i.e., natural light), half of them would pass through the analyzer.*

such as (E-10), associated with *light polarized* at an angle of 45° with respect to $\mathbf{e}_x$ and $\mathbf{e}_y$, is physically different from a statistical mixture of equal proportions of the states $\mathbf{e}_x$ and $\mathbf{e}_y$ associated with *natural light* (an unpolarized beam).

We can also understand the importance of the relative phase of the expansion

coefficients of the state vector, by considering the four states:

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}}(\mathbf{e}_x + \mathbf{e}_y) \tag{E-11}$$

$$\mathbf{e}_2 = \frac{1}{\sqrt{2}}(\mathbf{e}_x - \mathbf{e}_y) \tag{E-12}$$

$$\mathbf{e}_3 = \frac{1}{\sqrt{2}}(\mathbf{e}_x + i\mathbf{e}_y) \tag{E-13}$$

$$\mathbf{e}_4 = \frac{1}{\sqrt{2}}(\mathbf{e}_x - i\mathbf{e}_y) \tag{E-14}$$

which differ only by the relative phase of the coefficients (this phase being equal to 0, π, $+\frac{\pi}{2}$ and $-\frac{\pi}{2}$, respectively). These four states are physically quite different: the first two represent light which is polarized linearly along the bisectors of $(\mathbf{e}_x, \mathbf{e}_y)$; the second two represent circularly polarized light (right and left respectively).

E-1-b. Summation over the intermediate states

α. Prediction of measurement results in two simple experiments

(*i*) *Experiment 1.* Imagine that the observable A has been measured, at a given time, on a physical system, and that the non-degenerate eigenvalue a has been found. If $|u_a\rangle$ is the eigenvector associated with a, the physical system, immediately after the measurement, is in the state $|u_a\rangle$.

Before the system has had time to evolve, we measure another observable C that does not commute with A. Using the notation introduced in § C-6-a, we denote by $\mathcal{P}_a(c)$ the probability that this second measurement will yield the result c. Immediately before the measurement of C, the system is in the state $|u_a\rangle$. Therefore, if $|v_c\rangle$ is the eigenvector of C associated with the eigenvalue c (assumed to be non-degenerate), the postulates of quantum mechanics lead to:

$$\mathcal{P}_a(c) = |\langle v_c|u_a\rangle|^2 \tag{E-15}$$

(ii) Experiment 2. We now imagine *another* experiment, in which we measure successively and very rapidly three observables A, B, C, which do not commute with each other (the time separating two measurements is too short for the system to evolve). Denote by $\mathcal{P}_a(b, c)$ the probability, given that the result of the first measurement is a, that the results of the second and third will be b and c respectively. $\mathcal{P}_a(b, c)$ is equal to the product of $\mathcal{P}_a(b)$ (the probability that, the measurement of A having yielded a, that of B will yield b) and $\mathcal{P}_b(c)$ (the probability that, the measurement of B having yielded b, that of C will yield c):

$$\mathcal{P}_a(b, c) = \mathcal{P}_a(b)\ \mathcal{P}_b(c) \tag{E-16}$$

If all the eigenvalues of B are assumed to be non-degenerate and if $|w_b\rangle$ denotes the corresponding eigenvectors, it follows that [using for $\mathcal{P}_a(b)$ and $\mathcal{P}_b(c)$ formulas analogous to (E-15)]:

$$\mathcal{P}_a(b, c) = |\langle v_c|w_b\rangle|^2\ \langle w_b|u_a\rangle|^2 \tag{E-17}$$

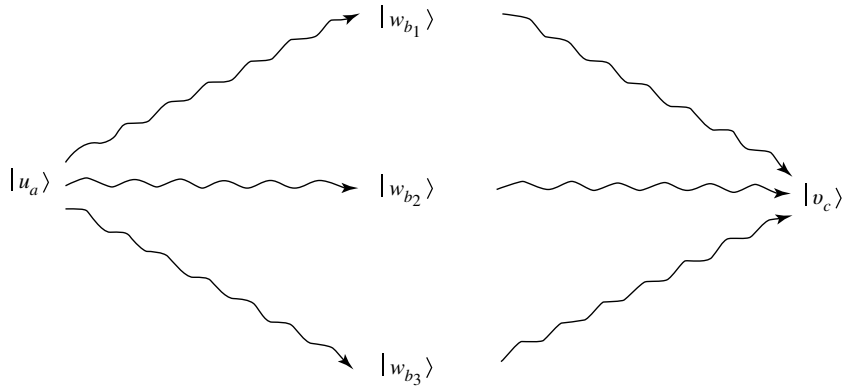


Figure 6: Different possible "paths" for the state vector of the system when the system is allowed to evolve freely (without undergoing any measurement) between the initial state $|u_a\rangle$ and the final state $|v_c\rangle$. In this case, we must add together the probability amplitudes associated with these different paths, and not their probabilities.

β. The fundamental difference between these two experiments

In both of these experiments, the state of the system after the measurement of the observable A is $|u_a\rangle$ (the role of this measurement being to fix this initial state). It then becomes $|v_c\rangle$ after the last measurement, that of the observable C (for this reason, $|v_c\rangle$ will be called the "final state"). It is possible in both cases to decompose the state of the system just before the measurement of C in terms of the eigenvectors $|w_b\rangle$ of B, and to say that between the state $|u_a\rangle$ and the state $|v_c\rangle$, the system "can pass" through several different "intermediate states" $|w_b\rangle$. Each of these intermediate states defines a possible "path" between the initial state $|u_a\rangle$ and the final state $|v_c\rangle$ (Fig. 6).

The difference between the two experiments described above is the following. In the first one, the path that the system has taken between the state $|u_a\rangle$ and the state $|v_c\rangle$ is not determined experimentally [we measure only the probability $\mathcal{P}_a(c)$ that, starting from $|u_a\rangle$, it ends up in $|v_c\rangle$]. On the other hand, in the second experiment, this path is determined, by measuring the observable B [thus enabling us to obtain the probability $\mathcal{P}_a(b,c)$ that the system, starting from $|u_a\rangle$, passes through a given intermediate state $|w_b\rangle$ and ends up finally in $|v_c\rangle$].

We could then be tempted, in order to relate $\mathcal{P}_a(c)$ to $\mathcal{P}_a(b,c)$, to use the following argument: in experiment 1, the system is "free to pass" through all intermediate states $|w_b\rangle$; it would then seem that the global probability $\mathcal{P}_a(c)$ should be equal to the sum of all the probabilities $\mathcal{P}_a(b,c)$ associated with each of the possible "paths". Can we not then write:

$$\mathcal{P}_a(c) \stackrel{?}{=} \sum_b \mathcal{P}_a(b,c) \tag{E-18}$$

As we shall see, *this formula is wrong.* Let us go back to the exact formula (E-15) for $\mathcal{P}_a(c)$; this formula brings in the probability amplitude $\langle v_c|u_a\rangle$ which we can write,

using the closure relation for the states $|w_b\rangle$:

$$\langle v_c|u_a\rangle = \sum_b \langle v_c|w_b\rangle \langle w_b|u_a\rangle \tag{E-19}$$

Substitute this expression into (E-15):

$$\begin{aligned}\mathcal{P}_a(c) &= \left|\sum_b \langle v_c|w_b\rangle \langle w_b|u_a\rangle\right|^2 \\ &= \sum_b |\langle v_c|w_b\rangle|^2 \ |\langle w_b|u_a\rangle|^2 \\ &\quad + \sum_b \sum_{b'\neq b} \langle v_c|w_b\rangle \langle w_b|u_a\rangle \langle v_c|w_{b'}\rangle^* \langle w_{b'}|u_a\rangle^*\end{aligned} \tag{E-20}$$

Using (E-17), we therefore obtain:

$$\mathcal{P}_a(c) = \sum_b \mathcal{P}_a(b,c) + \sum_b \sum_{b'\neq b} \langle v_c|w_b\rangle \langle w_b|u_a\rangle \langle v_c|w_{b'}\rangle^* \langle w_{b'}|u_a\rangle^* \tag{E-21}$$

This equation enables us to understand why formula (E-18) is wrong: all the "cross terms" that appear in the square of the modulus of sum (E-19) are absent in (E-18). All the *interference effects between the different possible paths* are thus missing in (E-18).

If, therefore, we want to establish a relation between these two experiments, we see that it is necessary to reason in terms of probability amplitudes. *When the intermediate states of the system are not determined experimentally, it is the probability amplitudes, and not the probabilities, which must be summed.*

The error in the reasoning which led to the wrong relation (E-18) is obvious, moreover, if we remember the fifth postulate (reduction of the wave packet). In the second experiment, the measurement of the observable B must, in fact, involve a perturbation of the system under study: during the measurement its state vector undergoes an abrupt change (projection onto one of the states $|w_b\rangle$). It is this unavoidable perturbation which is responsible for the disappearance of interference effects. In the first experiment, on the other hand, it is incorrect to say that the physical system "passes through one or another of the states $|w_b\rangle$"; it would be more accurate to say that it passes through *all* the states $|w_b\rangle$.

Comments:

(i) The preceding discussion resembles in every respect that of § A-2-a of Chapter I concerning Young's double-slit experiment. To determine the probability that a photon emitted by the source will arrive at a given point M of the screen, one must first calculate the total electric field at M. In this problem, the electric field plays the role of a probability amplitude. When one is not trying to determine through which slit the photon passes, it is the electric fields radiated by the two slits, and not their intensities, that must be added together to obtain the total field at M (whose square yields the desired probability). In other words, the field radiated by one of the slits at point M is

analogous to the probability amplitude for a photon, emitted by the source, to pass through this slit and to arrive at M.

(*ii*) It is not necessary to retain the assumption that the measurements of A and C in experiment 1 and of A, B, C in experiment 2 are performed very close together in time. If the system has had time to evolve between two of these measurements, we can use the Schrödinger equation to determine the modification of the state of the system due to this evolution [*cf.* Complement $\mathrm{F_{III}}$, comment *(ii)* of § 2].

E-1-c. Conclusion: The importance of the concept of probability amplitudes

The two examples studied in §§ E-1-a and E-1-b demonstrate the importance of the concept of probability amplitudes. Formulas (E-5) and (E-18), as well as the arguments that lead to them, are incorrect since they represent an attempt to calculate a probability directly without first considering the corresponding probability amplitude. In both cases, the correct expression (E-8) or (E-20) has the form of a *square of a sum* (more precisely, the square of the modulus of this sum), while the incorrect formula (E-5) or (E-18) contains only a *sum of squares* (all the cross terms, responsible for interference effects, being omitted).

From the preceding discussion, we shall therefore retain the following ideas:

(*i*) The probabilistic predictions of quantum theory are always obtained by squaring the modulus of a *probability amplitude*.

(*ii*) When, in a particular experiment, no measurement is made at an intermediate stage, one must never reason in terms of the probabilities of the various results that could have been obtained in such a measurement, but rather in terms of their probability amplitudes.

(*iii*) The fact that the states of a physical system are linearly superposable means that a probability amplitude often presents the form of a sum of partial amplitudes. The corresponding probability is then equal to the square of the modulus of a sum of terms, and *the various partial amplitudes interfere with each other*.

E-2. Case in which several states can be associated with the same measurement result

In the preceding section, we stressed and illustrated the fact that, in certain cases, the probability of an event is given by the postulates of quantum mechanics in the form of a *square of a sum* of terms (more precisely, the square of the modulus of such a sum). Now the statement of the fourth postulate [formula (B-7)] involves a *sum of squares* (the sum of the squares of the moduli) when the measurement result whose probability is sought is associated with a degenerate eigenvalue. It is important to understand that these two rules are not contradictory but, on the contrary, complementary: each term of the sum of squares (B-7) can itself be the square of a sum. This is the first point on which we shall focus our attention in this section. Furthermore, this discussion will enable us to complete the statement of the postulates: we shall consider measurement devices whose accurary is limited (as is always, of course, the case) and see how to predict theoretically the possible results. Finally, we shall extend to the case of continuous spectra the fifth postulate of reduction of the wave packet.

E-2-a. Degenerate eigenvalues

In the examples treated in § E-1, we always assumed that the results of the various measurements envisaged were simple, i.e. non-degenerate, eigenvalues of the corresponding observables. This hypothesis was intended to simplify these examples so that the origin of the interference effects appeared as clearly as possible.

Now let us consider a degenerate eigenvalue a_n of an observable A. The eigenstates associated with a_n form a vector subspace of dimension g_n, in which an orthonormal basis $\{|u_n^i\rangle; i = 1, 2, ..., g_n\}$ can be chosen.

The discussion of § C-6-b shows that knowing a measurement of A has yielded a_n is not sufficient to determine the state of the physical system after this measurement. We shall say that *several final states can be associated with the same result* a_n: if the initial state (the state before the measurement) is given, the final state after the measurement is perfectly well-defined; but if the initial state is changed, the final state is, in general, different (for the same measurement result a_n). All final states associated with a_n are linear combinations of the g_n orthonormal vectors $|u_n^i\rangle$, with $i = 1, 2, ..., g_n$.

Formula (B-7) indicates unambiguously how to find the probability $\mathcal{P}(a_n)$ that a measurement of A on a system in the state $|\psi\rangle$ will yield the result a_n. One chooses an orthonormal basis, for example $\{|u_n^i\rangle; i = 1, 2, ..., g_n\}$, in the eigensubspace which corresponds to a_n; one calculates the probability $|\langle u_n^i|\psi\rangle|^2$ of finding the system in each of the states of this basis; $\mathcal{P}(a_n)$ is then the sum of these g_n probabilities. However, it must not be forgotten that each probability $|\langle u_n^i|\psi\rangle|^2$ can be the square of the modulus of a sum of terms. Consider, for example, the case envisaged in § E-1-a-α and assume now that the eigenvalue a_n of the observable A, whose probability $\mathcal{P}(a_n)$ is to be calculated, is g_n-fold degenerate. Formula (E-6) is then replaced by:

$$\mathcal{P}(a_n) = \sum_{i=1}^{g_n} |\langle u_n^i|\psi\rangle|^2 \tag{E-22}$$

with:

$$\langle u_n^i|\psi\rangle = \lambda_1\langle u_n^i|\psi_1\rangle + \lambda_2\langle u_n^i|\psi_2\rangle \tag{E-23}$$

The discussion of § E-1-a-α remains valid for each of the terms of formula (E-22): $|\langle u_n^i|\psi\rangle|^2$, which is obtained from (E-23), is the square of a sum; $\mathcal{P}(a_n)$ is then the sum of these squares. § E-1-b can similarly be generalized to the case where the eigenvalues of the observables measured are degenerate.

Before summarizing the preceding discussions, we are going to study another important situation where several final states are associated with the same measurement result.

E-2-b. Insufficiently selective measurement devices

α. *Definition*

Assume that, in order to measure the observable A in a given physical system, we have at our disposal a device which works in the following way:

(i) This device can give only two different answers[10], which we shall denote, for convenience, by "yes" and "no".

[10] The following arguments can easily be generalized to cases where the device can give several different answers having characteristics similar to those described in (ii) and (iii).

(*ii*) If the system is in an eigenstate of A whose eigenvalue is included in a given interval Δ of the real axis, the answer is always "yes"; this is also the case when the state of the system is any linear combination of eigenstates of A associated with eigenvalues which are all included in Δ.

(*iii*) If the state of the system is an eigenstate of A whose eigenvalue falls outside Δ, or any linear combination of such eigenstates, the answer is always "no".

Δ therefore characterizes the resolving power of the measurement device under consideration. If there exists only one eigenvalue a_n of A in the interval Δ, the resolving power is infinite: when the system is in an arbitrary state, the probability $\mathcal{P}(\text{yes})$ of obtaining the answer "yes" is equal to the probability of finding a_n in a measurement of A; the probability $\mathcal{P}(\text{no})$ of obtaining "no" is obviously equal to $1 - \mathcal{P}(\text{yes})$. If, on the other hand, Δ contains several eigenvalues of A, the device does not have a sufficient resolution to distinguish between these various eigenvalues: we shall say that it is *insufficiently selective*. We shall see how to calculate $\mathcal{P}(\text{yes})$ and $\mathcal{P}(\text{no})$ in this case.

To be able to study the perturbation created by such a measurement on the state of the system, we are going to add the following hypothesis: the device transmits without perturbation the eigenstates of A associated with the eigenvalues of the interval Δ (as well as any linear combination of these eigenstates), while it "blocks" the eigenstates of A associated with the eigenvalues outside Δ (as well as all their linear combinations). The device thus behaves like a perfect filter for all states associated with Δ.

β. *Example*

Most measurement devices used in practice are insufficiently selective. For example, to measure the abscissa of an electron propagating parallel to the Oz axis, one can (Fig. 7) place in the xOy plane (Oy is perpendicular to the plane of the figure) a plate with a slit whose axis is parallel to Oy, the abscissas of the edges being x_1 and x_2. It can then be seen that any wave packet which is entirely included between the $x = x_1$ and $x = x_2$ planes (a superposition of eigenstates of X having eigenvalues x contained within the interval $[x_1, x_2]$) will enter the region to the right of the slit ("yes" answer); in this case, it will not undergo any modification. On the other hand, any wave packet situated below the $x = x_1$ plane or above the $x = x_2$ plane will be blocked by the plate and will not pass to the right ("no" answer).

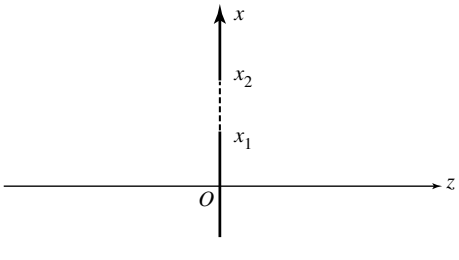


Figure 7: Schematic drawing of a device for measuring the abscissa x of a particle. Since the interval $\{x_1, x_2\}$ is necessarily non-zero, such a device is always imperfectly selective.

γ. *Quantum description*

For such an insufficiently selective device, several final states are possible after a measurement which has yielded the answer yes. They can be, for example, the various eigenstates of A that correspond to the eigenvalues contained in the interval Δ.

The physical problem posed by such devices, and which we are now going to consider, consists of predicting the answer which will be obtained when a system in an arbitrary state

enters the device. For example, for the apparatus of Figure 7, what happens when we are dealing with a wave packet which is neither entirely contained between the $x = x_1$ and $x = x_2$ planes (in which case the answer is certainly yes) nor entirely situated outside this region (in which case the answer is certainly no)? We shall see that this is equivalent to measuring an observable whose spectrum is degenerate.

Consider the subspace $\mathcal{E}_\Delta$ spanned by all the eigenstates of A whose eigenvalues a_n are contained within the interval Δ. The projector P_Δ onto this subspace is written (*cf.* § B-3-b-γ of Chap. II):

$$P_\Delta = \sum_{a_n \in \Delta} \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i| \tag{E-24}$$

(the eigenvalues a_n of the interval Δ can be degenerate, hence the additional index i; the vectors $|u_n^i\rangle$ are assumed to be orthonormal). $\mathcal{E}_\Delta$ is the subspace formed by all the possible states of the system after a measurement which has given the result yes.

Referring to the definition of the measurement device, we see that the response will certainly be yes for any state belonging to $\mathcal{E}_\Delta$, that is, for any eigenstate of P_Δ with the eigenvalue of +1. The answer will certainly be no for any state belonging to the supplement of $\mathcal{E}_\Delta$, that is, for any eigenstates of P_Δ with the eigenvalue of 0. The yes and no answers which can be furnished by the measurement device therefore correspond to the eigenvalues +1 and 0 of the observable P_Δ: it could be said that the device is actually measuring the observable P_Δ rather than A.

In the light of this interpretation, the case of an insufficiently selective measurement device can be treated in the framework of the postulates which we have stated. The probability $\mathcal{P}$(yes) of obtaining the answer yes is equal to the probability of finding the (degenerate) eigenvalue +1 of P_Δ. Now an orthonormal basis in the corresponding eigensubspace is constituted by the set of states $|u_n^i\rangle$ which are eigenstates of A with eigenvalues contained within the interval Δ. Applying formula (B-7) to the eigenvalue + 1 of the observable P_Δ, we therefore obtain (for a system in the state $|\psi\rangle$):

$$\mathcal{P}(\text{yes}) = \sum_{a_n \in \Delta} \sum_{i=1}^{g_n} |\langle u_n^i|\psi\rangle|^2 \tag{E-25}$$

Since there are only two possible answers, we obviously have:

$$\mathcal{P}(\text{no}) = 1 - \mathcal{P}(\text{yes}) \tag{E-26}$$

The projector onto the eigensubspace associated with the eigenvalue +1 of the observable P_Δ is P_Δ itself; formula (B-14) therefore gives here:

$$\mathcal{P}(\text{yes}) = \langle\psi|P_\Delta|\psi\rangle \tag{E-27}$$

[this formula is equivalent to (E-25)].

Similarly, since the device does not perturb states belonging to $\mathcal{E}_\Delta$ and blocks those of the supplement of $\mathcal{E}_\Delta$, we find that the state of the system after a measurement which has given the result yes is:

$$|\psi'\rangle = \frac{1}{\sqrt{\sum_{a_n \in \Delta} \sum_{i=1}^{g_n} |\langle u_n^i|\psi\rangle|^2}} \sum_{a_n \in \Delta} \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i|\psi\rangle \tag{E-28}$$

that is:

$$|\psi'\rangle = \frac{1}{\sqrt{\langle\psi|P_\Delta|\psi\rangle}} P_\Delta|\psi\rangle \tag{E-29}$$

When Δ contains only one eigenvalue a_n, P_Δ reduces to P_n: formulas (B-14) and (B-31) are then seen to be special cases of formulas (E-27) and (E-29).

E-2-c. Recapitulation: must one sum the amplitudes or the probabilities?

There are therefore cases (§ E-1) where, to calculate a probability, one takes the square of a sum, because several probability amplitudes must be added together. In other cases (§ E-2), one takes a sum of squares, because several probabilities must be added together. It is clearly important not to confuse these different cases and to know, in a given situation, if it is the probability amplitudes or the probabilities themselves which must be summed.

Young's double-slit experiment will again furnish us with a very convenient physical example which will enable us to illustrate and summarize the preceding discussions. Imagine that we want to calculate the probability for a particular photon to strike the plate anywhere between two points M_1 and M_2 having abscissas of x_1 and x_2 (Fig. 8). This probability is proportional to the total light intensity received by this portion of the plate. It is therefore a "sum of squares"; more precisely, it is the integral of the intensity $I(x)$ between x_1 and x_2. But each term $I(x)$ of this sum is obtained by squaring the electric field $\mathscr{E}(x)$ at x, which is equal to the sum of the electric fields $\mathscr{E}_A(x)$ and $\mathscr{E}_B(x)$ radiated at M by the slits A and B. $I(x)$ is therefore proportional to $|\mathscr{E}_A(x) + \mathscr{E}_B(x)|^2$, that is, to the square of a sum. $\mathscr{E}_A(x)$ and $\mathscr{E}_B(x)$ are the amplitudes associated with the two possible paths SAM and SBM which end at the same point M; they are added to obtain the amplitude at M since one is not trying to determine through which slit the photon passes. Then, to calculate the total light intensity received by the interval M_1M_2, one adds the intensities which arrive at the various points of this interval.

To sum up, the fundamental idea to be retained from the discussions of this section can be expressed schematically in the following way:

> Add the amplitudes corresponding to the same final state, then the probabilities corresponding to orthogonal final states.

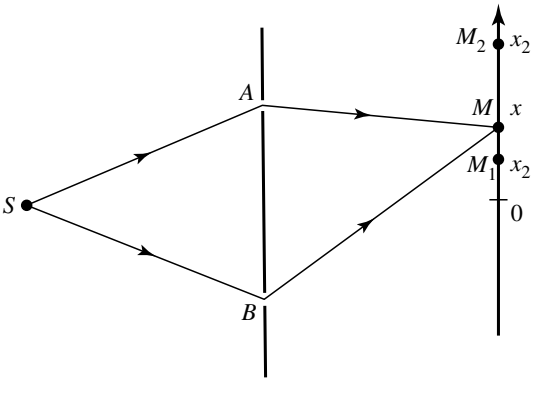


Figure 8: Young's double-slit experiment. To calculate the probability density for detecting a photon at the point M, it is necessary to add the electric fields radiated by the slits A and B, then to square the field thus obtained ("square of the sum"). The probability of finding a photon in the interval $[x_1, x_2]$ is now obtained by summing this probability density between x_1 and x_2 ("sum of squares").

E-2-d. Application to the treatment of continuous spectra

When the observable we want to measure has a continuous spectrum, only insufficiently selective devices can be used: it is impossible to imagine a physical device that could isolate a single eigenvalue belonging to a continuous set. We now show how the study of § E-2-b enables us to be more precise and complete in our treatment of observables with continuous spectra.

α. Example: measurement of the position of a particle

Let $\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle$ be the wave function of a (spinless) particle. What is the probability of finding the abscissa of this particle within the interval $[x_1, x_2]$ of the x-axis, using, for example, a measurement device like the one in Figure 7?

The subspace $\mathcal{E}_\Delta$ associated with this measurement result is the space spanned by the kets $|\mathbf{r}\rangle = |x, y, z\rangle$ which are such that: $x_1 \leq x \leq x_2$. Since these kets are orthonormal in the extended sense, application of the rule stated in § E-2-c above yields:

$$\begin{aligned}\mathcal{P}(x_1 \leq x \leq x_2) &= \int_{x_1}^{x_2} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}z \ |\langle x, y, z|\psi\rangle|^2 \\ &= \int_{x_1}^{x_2} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}z \ |\psi(\mathbf{r})|^2 \end{aligned} \tag{E-30}$$

Formula (E-27) obviously leads to the same result, since the projector P_Δ is written here:

$$P_\Delta = \int_{x_1}^{x_2} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}z \ |x, y, z\rangle \langle x, y, z| \tag{E-31}$$

and we therefore have:

$$\begin{aligned}\mathcal{P}(x_1 \leq x \leq x_2) &= \langle\psi|P_\Delta|\psi\rangle \\ &= \int_{x_1}^{x_2} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}z \ \langle\psi|x, y, z\rangle \langle x, y, z|\psi\rangle \end{aligned} \tag{E-32}$$

To know the state $|\psi'\rangle$ of the particle after such a measurement, which has yielded the result yes, it suffices to apply formula (E-29):

$$\begin{aligned}|\psi'\rangle &= \frac{1}{N} \ P_\Delta|\psi\rangle \\ &= \frac{1}{N} \int_{x_1}^{x_2} \mathrm{d}x' \int_{-\infty}^{+\infty} \mathrm{d}y' \int_{-\infty}^{+\infty} \mathrm{d}z' \ |x', y', z'\rangle \langle x', y', z'|\psi\rangle \end{aligned} \tag{E-33}$$

where the normalization factor $N = \sqrt{\langle\psi|P|\psi\rangle}$ is known [formula (E-32)]. Let us calculate the wave function $\psi'(\mathbf{r}) = \langle \mathbf{r}|\psi'\rangle$ associated with the ket $|\psi'\rangle$:

$$\langle \mathbf{r}|\psi'\rangle = \frac{1}{N} \int_{x_1}^{x_2} \mathrm{d}x' \int_{-\infty}^{+\infty} \mathrm{d}y' \int_{-\infty}^{+\infty} \mathrm{d}z' \ \langle \mathbf{r}|\mathbf{r}'\rangle \psi(\mathbf{r}') \tag{E-34}$$

Now $\langle \mathbf{r}|\mathbf{r}'\rangle = \delta(\mathbf{r}-\mathbf{r}') = \delta(x-x') \ \delta(y-y') \ \delta(z-z')$. The integrations over y' and z' can therefore be performed immediately: they amount to replacing y' and z' by y and z in the function to be integrated. Equation (E-34) thus becomes:

$$\psi'(x, y, z) = \frac{1}{N} \int_{x_1}^{x_2} \mathrm{d}x' \ \delta(x - x') \ \psi(x', y, z) \tag{E-35}$$

If the point x is situated inside the interval of integration $[x_1, x_2]$, the result is the same as if we were integrating x' from $-\infty$ to $+\infty$:

$$\psi'(x,y,z) = \frac{1}{N}\ \psi(x,y,z) \qquad \text{for} \quad x_1 \leq x \leq x_2 \tag{E-36}$$

On the other hand, if x falls outside the interval of integration, $\delta(x-x')$ is zero for all values of x' included in this interval, and:

$$\psi'(x,y,z) = 0 \qquad \text{for} \quad x > x_2 \quad \text{and} \quad x < x_1 \tag{E-37}$$

The part of $\psi(\mathbf{r})$ that corresponds to the interval accepted by the measurement device therefore persists, undeformed, immediately after the measurement [the factor $1/N$ simply ensures that $\psi'(\mathbf{r})$ remains normalized]; the rest is suppressed by the measurement. The wave packet $\psi(\mathbf{r})$ representing the initial state of the particle is, as it were, "truncated" by the edges of the slit.

Comments:

(*i*) This example clearly reveals the concrete meaning of the "reduction of the wave packet".

(*ii*) If a large number of particles, all in the same state $|\psi\rangle$, enter the device successively, the result will sometimes be yes and sometimes be no [with the probabilities $\mathcal{P}$(yes) and $\mathcal{P}$(no)]. If the result is yes, the particle continues on its way, starting from the "truncated" state $|\psi'\rangle$; if the result is no, the particle is absorbed by the screen.

In the example we are considering here, the measurement device becomes all the more selective as $x_2 - x_1$ becomes smaller. We see, however, that it is impossible to make it perfectly selective because the spectrum of X is continuous: however narrow the slit may be, the interval $[x_1, x_2]$ which it defines always contains an infinity of eigenvalues. Nevertheless, in the limiting case of a slit of an infinitely small width Δx, we find the equivalent of formula (B-17), which was the expression of the fourth postulate in the case of a continuous spectrum.

Let us choose $x_1 = x_0 - \dfrac{\Delta x}{2}$ and $x_2 = x_0 + \dfrac{\Delta x}{2}$ (a slit of width Δx centered at x_0), and assume that the wave function $\psi(\mathbf{r})$ varies very little within the interval Δx. Then, in (E-30), we can replace $|\psi(\mathbf{r})|^2$ by $|\psi(x_0,y,z)|^2$ and perform the integration over x:

$$\mathcal{P}\left(x_0 - \frac{\Delta x}{2} \leq x \leq x_0 + \frac{\Delta x}{2}\right) \simeq \Delta x \int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}z\ |\psi(x_0,y,z)|^2 \tag{E-38}$$

We indeed find a probability equal to the product of Δx and a positive quantity which plays the role of a probability density at the point x_0. The difference with formula (B-17) lies in the fact that the latter applies to the case of a continuous but non-degenerate spectrum, while here the eigenvalues of X are infinitely degenerate in $\mathcal{E}_{\mathbf{r}}$; this is the origin of the integrals over y and z that appear in (E-38) (summation over the indices associated with the degeneracy).

β. Postulate of reduction of wave packets in the case of a continuous spectrum

In § B-3-c, we limited ourselves, in the statement of the fifth postulate, to the case of a discrete spectrum. Formula (E-33) and its accompanying discussion enable us to understand the form assumed by this postulate when a continuous spectrum is considered: one simply applies the results of § E-2-b concerning insufficiently selective devices.

Let A be an observable with a continuous spectrum (assumed, for simplicity, to be non-degenerate). The notation is the same as in § B-3-b-β.

If a measurement of A on a system in the state $|\psi\rangle$ has yielded the result α_0 to within $\Delta\alpha$, the state of the system immediately after this measurement is described by:

$$|\psi'\rangle = \frac{1}{\sqrt{\langle\psi|P_{\Delta\alpha}(\alpha_0)\ |\psi\rangle}}\ P_{\Delta\alpha}(\alpha_0)|\psi\rangle \tag{E-39}$$

with:

$$P_{\Delta\alpha}(\alpha_0) = \int_{\alpha_0-\frac{\Delta\alpha}{2}}^{\alpha_0+\frac{\Delta\alpha}{2}} \mathrm{d}\alpha\ |v_\alpha\rangle\,\langle v_\alpha| \tag{E-40}$$

Figures 9-a and 9-b illustrate this statement. If the function $\langle v_a|\psi\rangle$, representing $|\psi\rangle$ in the $\{|v_a\rangle\}$ basis, has the form indicated in Figure 9-a, the state of the system immediately after the measurement is represented, to within a normalization factor, by the function of Figure 9b [the calculation is analogous in all respects to the one which derives (E-36) and (E-37) from (E-33)].

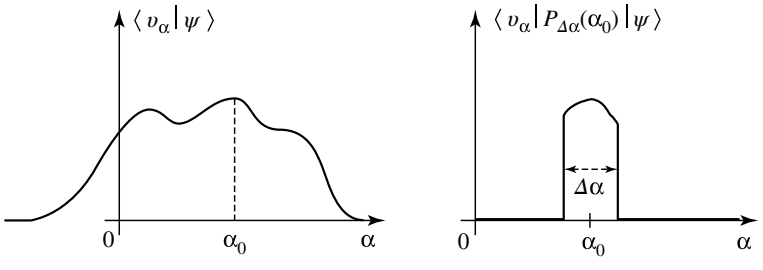


Figure 9: Illustration of the postulate of reduction of wave packets in the case of a continuous spectrum: one measures the observable A, with eigenvectors $|v_a\rangle$ and eigenvalues α. The measurement device has a selectivity $\Delta\alpha$. If the value found is α_0 to within $\Delta\alpha$, the effect of the measurement on the wave function $\langle v_a|\psi\rangle$ is to "truncate" it about the value α_0 (to normalize the new wave function, it is obviously necessary to multiply it by a factor larger than 1).

We see that, even if $\Delta\alpha$ is very small, one can never actually prepare the system in the state $|v_{\alpha_0}\rangle$, which would be represented, in the $\{|v_\alpha\rangle\}$ basis, by $\langle v_\alpha|v_{\alpha_0}\rangle = \delta(\alpha-\alpha_0)$. We can only obtain a narrow function centered at α_0, since $\Delta\alpha$ is never exactly zero.

References and suggestions for further reading:

Development of quantum mechanical concepts: references of section 4 of the bibliography, particularly Jammer (4.8).

Discussion and interpretation of the postulates: references of section 5 of the bibliography; Von Neumann (10.10), Chaps. V and VI; Feynman III (1.2), § 2.6, Chap. 3 and § 8.3.

Quantization rules using Poisson brackets: Dirac (1.13), § 21; Schiff (1.18), § 24.

Probability and statistics: see the corresponding subsection of section 10 of the bibliography.

COMPLEMENTS OF CHAPTER III, READER'S GUIDE

A_{III} : PARTICLE IN AN INFINITE ONE-DIMENSIONAL POTENTIAL WELL **B_{III} : STUDY OF THE PROBABILITY CURRENT IN SOME SPECIAL CASES**	Direct applications of Chapter III to simple cases. The accent is placed on the physical discussion of the results (elementary level).
C_{III} : ROOT MEAN SQUARE DEVIATIONS OF TWO CONJUGATE OBSERVABLES	A little more formal; general proof of the Heisenberg relations; may be skipped in a first reading.
D_{III} : MEASUREMENTS BEARING ON ONLY ONE PART OF A PHYSICAL SYSTEM	Discussion of measurements bearing on only one part of a system; a rather simple but somewhat formal application of Chapter III; may be skipped on a first reading.

(more complements on the next page)

E_{III} : THE DENSITY OPERATOR

F_{III} : THE EVOLUTION OPERATOR

G_{III} : THE SCHRÖDINGER AND HEISENBERG PICTURES

H_{III} : GAUGE INVARIANCE

J_{III} : PROPAGATOR FOR THE SCHRÖDINGER EQUATION

Complements that serve as an introduction to a more advanced quantum mechanics course. Aside from F_{III}, which is simple, they are on a higher level than the rest of this book, but they are comprehensible if Chapter III has been read. May be reserved for subsequent study.

E_{III} : definition and properties of the density operator, which is used in the quantum mechanical description of systems whose state is imperfectly known (statistical mixture of states). Fundamental tool of quantum statistical mechanics.

F_{III} : introduction of the evolution operator, which gives the quantum state of a system at an arbitrary instant t in terms of its state at the instant t_0.

G_{III} : describes the evolution of a quantum system in a way that is different from, but equivalent to, that of Chapter III. The time dependence now appears in the observables and not in the state of the system.

H_{III} : discussion of the quantum formalism in the case where the system is subject to an electromagnetic field. Although the description of the system involves the electromagnetic potentials, the physical properties depend only on the values of the electric and magnetic fields; they remain invariant when the potentials describing the same electromagnetic field are changed.

J_{III} : an introduction to a different way of approaching quantum mechanics, based on a principle analogous to Huygens' principle in classical wave optics.

K_{III} : UNSTABLE LEVELS, LIFETIMES

Simple introduction to the important physical concepts of instability and lifetimes; easy, but can be skipped in a first reading.

L_{III} : EXERCISES

(more complements on the next page)

M_{III} : BOUND STATES OF A PARTICLE IN A "POTENTIAL WELL" OF ARBITRARY SHAPE

N_{III} : UNBOUND STATES OF A PARTICLE IN THE PRESENCE OF A POTENTIAL WELL OR BARRIER OF ARBITRARY SHAPE

O_{III} :QUANTUM PROPERTIES OF A PARTICLE IN A ONE-DIMENSIONAL PERIODIC STRUCTURE

Return to one-dimensional problems, considered from a more general point of view than in Chapter I and its complements.

M_{III} : generalisation to an arbitray potential well of the main results obtained in § 2-c of Complement H_I; recommended, since easy and physically important.

N_{III} : study of unbound stationary states in an arbitrary potential; a little more formal; the definitions and results of this complement are necessary for complement O_{III}.

O_{III} : introduction of the concept (which is fundamental to solid state physics) of energy bands in a potential having a periodic structure (this concept will be treated differently in Complement F_{XI}); rather difficult, can be reserved for later reading.

Complement A_{III}

Particle in an infinite potential well

1 Distribution of the momentum values in a stationary state 271
1-a Calculation of the function $\overline{\varphi}_n(p)$ of $\langle P \rangle$, and, of ΔP 271
1-b Discussion 273
2 Evolution of the particle's wave function 275
2-a Wave function at the instant t 276
2-b Evolution of the shape of the wave packet 276
2-c Motion of the center of the wave packet 277
3 Perturbation created by a position measurement 279

In Complement H_I (§ 2-c-β), we studied the stationary states of a particle in a one-dimensional infinite potential well. Here we intend to re-examine this subject from a physical point of view. This will allow us to apply some of the postulates of Chapter III to a concrete case. We shall focus in particular on the results obtained when the position or momentum of the particle is measured.

1. Distribution of the momentum values in a stationary state

1-a. Calculation of the function $\overline{\varphi}_n(p)$ of $\langle P \rangle$, and, of ΔP

We have seen that the stationary states of the particle correspond to the energies[1]:

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2} \tag{1}$$

and to the wave functions:

$$\varphi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) \tag{2}$$

(where a is the width of the well and n is any positive integer).

Consider a particle in the state $|\varphi_n\rangle$, with energy E_n. The probability of a measurement of the momentum P of the particle yielding a result between p and $p + \mathrm{d}p$ is:

$$\overline{\mathcal{P}}_n(p)\mathrm{d}p = |\overline{\varphi}_n(p)|^2 \, \mathrm{d}p \tag{3}$$

with:

$$\overline{\varphi}_n(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_0^a \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) \mathrm{e}^{-ipx/\hbar} \mathrm{d}x \tag{4}$$

[1] We shall use the notation of Complement H_I.

This integral is easy to calculate; it is equal to:

$$\begin{aligned}\overline{\varphi}_n(p) &= \frac{1}{2i\sqrt{\pi\hbar a}} \int_0^a \left[\mathrm{e}^{i\left(\frac{n\pi}{a}-\frac{p}{\hbar}\right)x} - \mathrm{e}^{-i\left(\frac{n\pi}{a}+\frac{p}{\hbar}\right)x} \right] \mathrm{d}x \\ &= \frac{1}{2i\sqrt{\pi\hbar a}} \left[\frac{\mathrm{e}^{i\left(\frac{n\pi}{a}-\frac{p}{\hbar}\right)a}-1}{i\left(\frac{n\pi}{a}-\frac{p}{\hbar}\right)} - \frac{\mathrm{e}^{-i\left(\frac{n\pi}{a}+\frac{p}{\hbar}\right)a}-1}{-i\left(\frac{n\pi}{a}+\frac{p}{\hbar}\right)} \right] \end{aligned} \tag{5}$$

that is:

$$\overline{\varphi}_n(p) = \frac{1}{2i}\sqrt{\frac{a}{\pi\hbar}} \mathrm{e}^{i\left(\frac{n\pi}{2}-\frac{pa}{2\hbar}\right)} \left[F\left(p - \frac{n\pi\hbar}{a}\right) + (-1)^{n+1} F\left(p + \frac{n\pi\hbar}{a}\right) \right] \tag{6}$$

with:

$$F(p) = \frac{\sin(pa/2\hbar)}{pa/2\hbar} \tag{7}$$

To within a proportionality factor, the function $\overline{\varphi}_n(p)$ is the sum (or the difference) of two "diffraction functions" $F\left(p \pm \dfrac{n\pi\hbar}{a}\right)$, centered at $p = \mp\dfrac{n\pi\hbar}{a}$. The "width" of these functions (the distance between the first two zeros, symmetric with respect to the central value) does not depend on n and is equal to $\dfrac{4\pi\hbar}{a}$. Their "amplitude" does not depend on n either.

The function inside brackets in expression (6) is even if n is odd, and odd if n is even. The probability density $\overline{\mathcal{P}}_n(p)$ given in (3) is therefore an even function of p in all cases, so that:

$$\langle P \rangle_n = \int_{-\infty}^{+\infty} \overline{\mathcal{P}}_n(p)\, p\, \mathrm{d}p = 0 \tag{8}$$

The mean value of the momentum of the particle in the energy state E_n is therefore zero.

Let us calculate, in the same way, the mean value $\langle P^2 \rangle_n$ of the square of the momentum. Using the fact that in the $\{|x\rangle\}$ representation P acts like $\dfrac{\hbar}{i}\dfrac{\mathrm{d}}{\mathrm{d}x}$, and performing an integration by parts, we obtain[2]:

$$\begin{aligned}\langle P^2 \rangle_n &= \hbar^2 \int_0^a \left| \frac{\mathrm{d}\varphi_n}{\mathrm{d}x} \right|^2 \mathrm{d}x \\ &= \hbar^2 \int_0^a \frac{2}{a} \left(\frac{n\pi}{a}\right)^2 \cos^2\left(\frac{n\pi x}{a}\right) \mathrm{d}x \\ &= \left(\frac{n\pi\hbar}{a}\right)^2 \end{aligned} \tag{9}$$

[2]Result (9) could also be derived from (6) by performing the integral $\langle P^2 \rangle_n = \int_{-\infty}^{+\infty} |\overline{\varphi}_n(p)|^2 p^2\, \mathrm{d}p$. This calculation, which presents no theoretical difficulties, is nevertheless not as direct as the one given here.

From (8) and (9), we get:

$$\Delta P_n = \sqrt{\langle P^2 \rangle_n - \langle P \rangle_n^2} = \frac{n\pi\hbar}{a} \tag{10}$$

The root mean square deviation therefore increases linearly with n.

1-b. Discussion

Let us plot, for different values of n, the curves giving the probability density $\overline{\mathcal{P}}_n(p)$. To do this, let us begin by studying the function inside brackets in expression (6). For the ground state ($n = 1$), it is the sum of two functions F, the centers of these two diffraction curves being separated by half their width (Fig. 1-a). For the first excited level ($n = 2$), the distance between these centers is twice as large, and in this case, moreover, the difference of two functions F must be taken (Fig. 2-a). Finally, for an excited level corresponding to a large value of n, the centers of the two diffraction curves are separated by a distance much greater than their width.

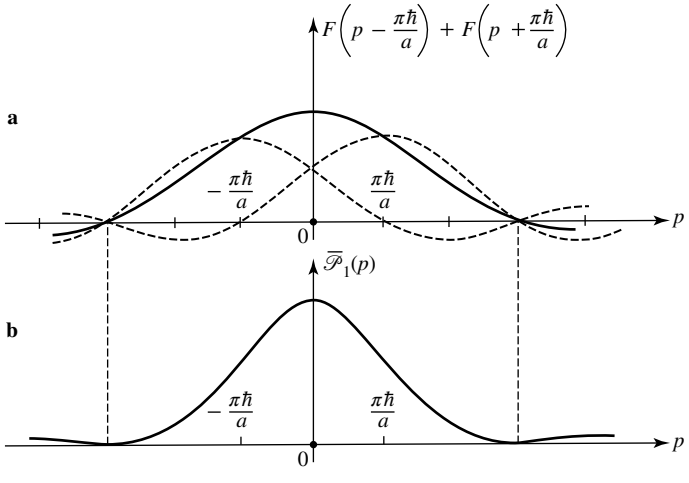


Figure 1: The wave function $\overline{\varphi}_1(p)$, associated in the $\{|p\rangle\}$ representation with the ground state of a particle in an infinite well, is obtained by adding two diffraction functions F (dashed lines curves in figure a). Since the centers of these two functions F are separated by half their width, their sum has the shape represented by the solid-line curve in figure a. Squaring this sum, one obtains the probability density $\overline{\mathcal{P}}_1(p)$ associated with a measurement of the momentum of the particle (fig. b).

Squaring these functions, one obtains the probability density $\overline{\mathcal{P}}_n(p)$ (*cf.* Fig. 1-b and 2-b). Note that for large n the interference term between $F\left(p - \frac{n\pi\hbar}{a}\right)$ and $F\left(p + \frac{n\pi\hbar}{a}\right)$ is negligible (because of the separation of the centers of the two curves):

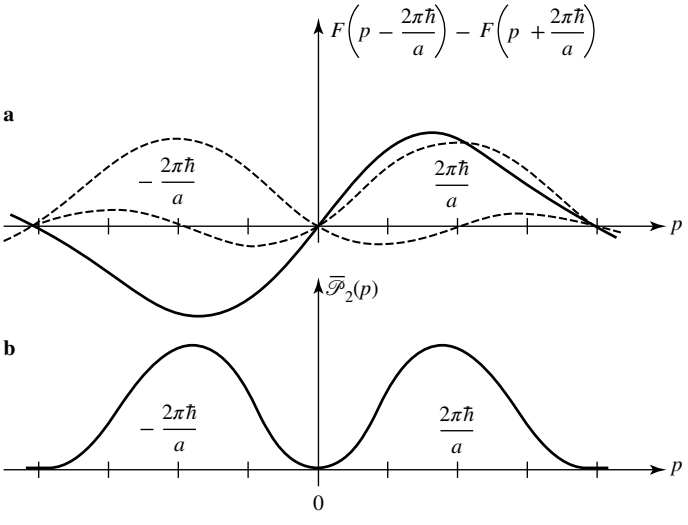


Figure 2: For the first excited level, the function $\overline{\varphi}_2(p)$ is obtained by taking the difference between two functions F, which have the same width as in Figure 1-a but are now more widely separated (dash-line curves in figure a). The curve obtained in this way is the solid line in figure a. The probability density $\overline{\mathcal{P}}_2(p)$ then has two maxima located in the neighborhood of $p = \pm 2\pi\hbar/a$ (fig. b).

$$\overline{\mathcal{P}}_n(p) = \frac{a}{4\pi\hbar}\left[F\left(p - \frac{n\pi\hbar}{a}\right) + (-1)^{n+1}F\left(p + \frac{n\pi\hbar}{a}\right)\right]^2$$
$$\simeq \frac{a}{4\pi\hbar}\left[F^2\left(p - \frac{n\pi\hbar}{a}\right) + F^2\left(p + \frac{n\pi\hbar}{a}\right)\right] \tag{11}$$

The function $\overline{\mathcal{P}}_n(p)$ then has the shape shown in Figure 3.

It can be seen that when n is large, the probability density has two symmetrical peaks, of width $\frac{4\pi\hbar}{a}$, centered at $p = \pm\frac{n\pi\hbar}{a}$. It is then possible to predict with almost complete certainty the results of a measurement of the momentum of the particle in the state $|\varphi_n\rangle$: the value found will be nearly equal to $+\frac{n\pi\hbar}{a}$ or $-\frac{n\pi\hbar}{a}$, the relative accuracy[3] improving as n increases (the two opposite values $\pm\frac{n\pi\hbar}{a}$ being equally probable). This is simple to understand: for large n, the function $\varphi_n(x)$, which varies sinusoidally, performs numerous oscillations inside the well; it can then be considered to be practically the sum of two progressive waves corresponding to opposite momenta $p = \pm\frac{n\pi\hbar}{a}$.

When n decreases, the relative accuracy with which one can predict the possible values of the momentum diminishes. We see, for example, in Figure 2-b, that when $n = 2$,

[3]The absolute accuracy is independent of n, since the width of the curves is always $\frac{4\pi\hbar}{a}$.

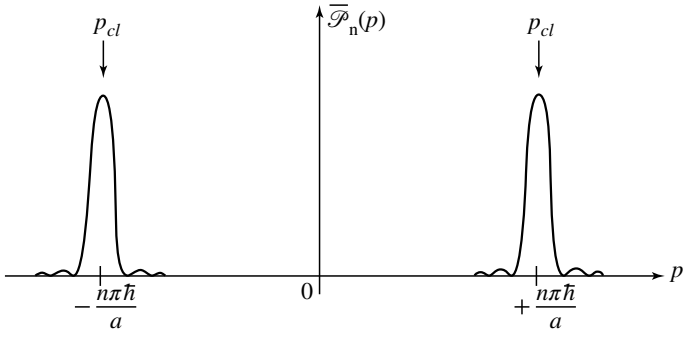


Figure 3: When n is large (a very excited level), the probability density has two pronounced peaks, centered at the values $p = \pm n\pi\hbar/a$, which are the momenta associated with the classical motion at the same energy.

the function $\overline{\mathcal{P}}_n(p)$ has two peaks whose widths are comparable to their distance from the origin. In this case, the wave function undergoes only one oscillation inside the well. It is not surprising that, for this sinusoid "truncated" at $x = 0$ and $x = a$, the wavelength (and therefore, the momentum of the particle) is poorly defined. Finally, for the ground state, the wave function is represented by half a sinusoidal arc: the relative values of the wavelength and momentum of the particle are then very poorly known (Fig. 1-b).

Comments:

(*i*) Let us calculate the momentum of a classical particle of energy E_n given in (1); we have:

$$\frac{p_{cl}^2}{2m} = \frac{n^2\pi^2\hbar^2}{2ma^2} \tag{12}$$

that is:

$$p_{cl} = \pm\frac{n\pi\hbar}{a} \tag{13}$$

When n is large, the two peaks of $\overline{\mathcal{P}}_n(p)$ therefore correspond to the classical values of the momentum.

(*ii*) We see that, for large n, although the absolute value of the momentum is well-defined, its sign is not. This is why ΔP_n is large: for probability distributions with two maxima like that of Figure 3, the root mean square deviation reflects the distance between the two peaks; it is no longer related to their widths.

2. Evolution of the particle's wave function

Each of the states $|\varphi_n\rangle$, with its wave function $\varphi_n(x)$, describes a stationary state, which leads to time-independent physical predictions. Time evolution appears only when the

state vector is a linear combination of several kets $|\varphi_n\rangle$. We shall consider here a very simple case, for which at time $t = 0$ the state vector $|\psi(0)\rangle$ is:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}\left[|\varphi_1\rangle + |\varphi_2\rangle\right] \tag{14}$$

2-a. Wave function at the instant t

Apply formula (D-54) of Chapter III; we immediately obtain:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\left[\mathrm{e}^{-i\frac{\pi^2\hbar}{2ma^2}t}|\varphi_1\rangle + \mathrm{e}^{-2i\frac{\pi^2\hbar}{ma^2}t}|\varphi_2\rangle\right] \tag{15}$$

or, omitting a *global* phase factor of $|\psi(t)\rangle$:

$$|\psi(t)\rangle \propto \frac{1}{\sqrt{2}}\left[|\varphi_1\rangle + \mathrm{e}^{-i\omega_{21}t}|\varphi_2\rangle\right] \tag{16}$$

with:

$$\omega_{21} = \frac{E_2 - E_1}{\hbar} = \frac{3\pi^2\hbar}{2ma^2} \tag{17}$$

2-b. Evolution of the shape of the wave packet

The shape of the wave packet is given by the probability density:

$$|\psi(x,t)|^2 = \frac{1}{2}\varphi_1^2(x) + \frac{1}{2}\varphi_2^2(x) + \varphi_1(x)\,\varphi_2(x)\cos\,\omega_{21}t \tag{18}$$

We see that the time variation of the probability density is due to the interference term in $\varphi_1\varphi_2$. Only one Bohr frequency appears, $\nu_{21} = (E_2 - E_1)/h$, since the initial state

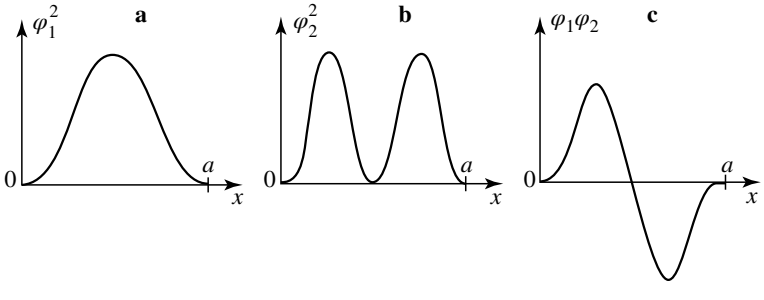


Figure 4: Graphical representation of the functions φ_1^2 (the probability density of the particle in the ground state), φ_2^2 (the probability density of the particle in the first excited state) and $\varphi_1\varphi_2$ (the cross term responsible for the evolution of the shape of the wave packet).

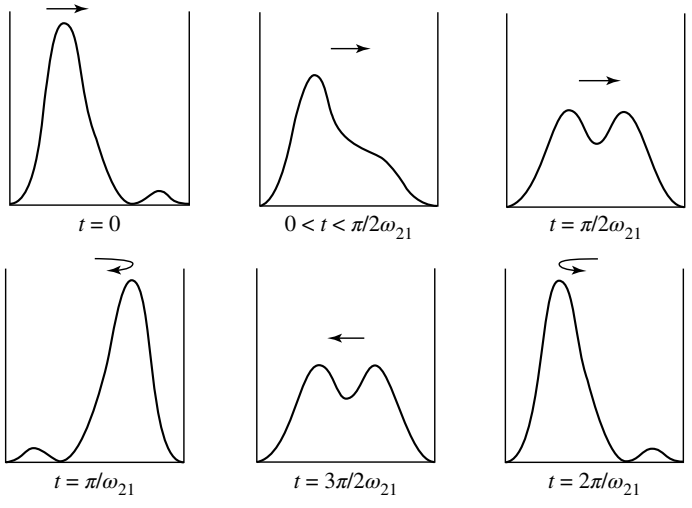


Figure 5: Periodic motion of a wave packet obtained by superposing the ground state and the first excited state of a particle in an infinite well. The frequency of the motion is the Bohr frequency $\omega_{21}/2\pi$.

(14) is composed only of the two states $|\varphi_1\rangle$ and $|\varphi_2\rangle$. The curves corresponding to the variation of the functions φ_1^2, φ_2^2 and $\varphi_1\varphi_2$ are plotted in Figures 4-a, b and c.

Using these figures and relation (18), it is not difficult to represent graphically the variation in time of the shape of the wave packet (*cf.* Fig. 5): we see that the wave packet oscillates between the two walls of the well.

2-c. Motion of the center of the wave packet

Let us calculate the mean value $\langle X\rangle(t)$ of the position of the particle at time t. It is convenient to take:

$$X' = X - a/2 \tag{19}$$

since, by symmetry, the diagonal matrix elements of X' are zero:

$$\langle\varphi_1|X'|\varphi_1\rangle \propto \int_0^a \left(x - \frac{a}{2}\right) \sin^2\left(\frac{\pi x}{a}\right) \mathrm{d}x = 0$$

$$\langle\varphi_2|X'|\varphi_2\rangle \propto \int_0^a \left(x - \frac{a}{2}\right) \sin^2\left(\frac{2\pi x}{a}\right) \mathrm{d}x = 0 \tag{20}$$

We then have:

$$\langle X'\rangle(t) = \mathrm{Re}\left\{\mathrm{e}^{-i\omega_{21}t}\langle\varphi_1|X'|\varphi_2\rangle\right\} \tag{21}$$

with:

$$\langle\varphi_1|X'|\varphi_2\rangle = \langle\varphi_1|X|\varphi_2\rangle - \frac{a}{2}\langle\varphi_1|\varphi_2\rangle$$
$$= \frac{2}{a}\int_0^a x \sin\frac{\pi x}{a}\sin\frac{2\pi x}{a}\mathrm{d}x$$
$$= -\frac{16a}{9\pi^2} \tag{22}$$

Therefore:

$$\langle X\rangle(t) = \frac{a}{2} - \frac{16a}{9\pi^2}\cos\omega_{21}t \tag{23}$$

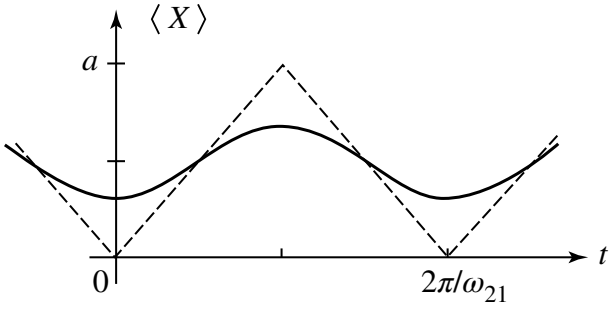


Figure 6: Time variation of the mean value $\langle X\rangle$ corresponding to the wave packet's motion plotted in Figure 5. The dashed line represents the position of a classical particle moving with the same period. Quantum mechanics predicts that the center of the wave packet will turn back before reaching the wall, as explained by the action of the potential on the "edges" of the wave packet.

The variation of $\langle X\rangle(t)$ is represented in Figure 6. In dashed lines, the variation of the position of a classical particle has been traced, for a particle moving to and fro in the well with an angular frequency of ω_{21} (since it is not subjected to any force except at the walls, its position varies linearly with t between 0 and a during each half-period).

We immediately notice a very clear difference between these two types of motion, classical and quantum mechanical. The center of the quantum wave packet, instead of turning back at the walls of the well, executes a movement of smaller amplitude and retraces its steps before reaching the regions where the potential is not zero. We see again here a result of § D-2 of Chapter I: since the potential varies infinitely quickly at $x = 0$ and $x = a$, its variation within a domain of the order of the dimension of the wave packet is not negligible, and the motion of the center of the wave packet does not obey the laws of classical mechanics (see also Chapter III, § D-1-d-γ). The physical explanation of this phenomenon is the following: before the center of the wave packet has touched the wall, the action of the potential on the "edges" of this packet is sufficient to make it turn back.

Comment:

The mean value of the energy of the particle in the state $|\psi(t)\rangle$ calculated in (15) is easy to obtain:

$$\langle H\rangle = \frac{1}{2}E_1 + \frac{1}{2}E_2 = \frac{5}{2}E_1 \tag{24}$$

as is:

$$\langle H^2\rangle = \frac{1}{2}E_1^2 + \frac{1}{2}E_2^2 = \frac{17}{2}E_1^2 \tag{25}$$

which gives:

$$\Delta H = \frac{3}{2}E_1 \tag{26}$$

Note in particular that $\langle H\rangle$, $\langle H^2\rangle$ and ΔH are not time-dependent; this could have been foreseen, since H is a constant of the motion. In addition, we see from the preceding discussion that the wave packet evolves appreciably over a time of the order of:

$$\Delta t \simeq \frac{1}{\omega_{21}} \tag{27}$$

Using (26) and (27), we find:

$$\Delta H.\Delta t \simeq \frac{3}{2}E_1 \times \frac{\hbar}{3E_1} = \frac{\hbar}{2} \tag{28}$$

We again find the time-energy uncertainty relation (§ D-2-e of Chapter III).

3. Perturbation created by a position measurement

Consider a particle in the state $|\varphi_1\rangle$. Assume that the position of the particle is measured at time $t = 0$, with the result $x = a/2$. What are the probabilities of the different results that can be obtained in a measurement of the energy, performed immediately after this first measurement?

One must beware of the following false argument: after the measurement, the particle is in the eigenstate of X corresponding to the result found, and its wave function is therefore proportional to $\delta(x-a/2)$; if a measurement of the energy is then performed, the various values E_n can be found, with probabilities proportional to:

$$\left|\int_0^a \mathrm{d}x\ \delta\left(x-\frac{a}{2}\right)\varphi_n^*(x)\right|^2 = \left|\varphi_n\left(\frac{a}{2}\right)\right|^2 = \begin{cases} 2/a \text{ if } n \text{ is odd} \\ \\ 0 \text{ if } n \text{ is even} \end{cases} \tag{29}$$

Using this *incorrect* argument, one would find the probabilities of all values of E_n corresponding to odd n to be equal. This is absurd, since the sum of these probabilities would then be infinite.

This error results from the fact that we have not taken the norm of the wave function into account. To apply the fourth postulate of Chapter III correctly, it is necessary to write the wave function as normalized just after the first measurement. However it is not possible[4] to normalize the function $\delta(x - a/2)$. The problem posed above must be stated more precisely.

As we saw in § E-2-b of Chapter III, an experiment in which the measurement of an observable with a continuous spectrum is performed never yields any result with complete accuracy. For the case with which we are concerned, we can only say that:

$$\frac{a}{2} - \frac{\varepsilon}{2} \leq x \leq \frac{a}{2} + \frac{\varepsilon}{2} \tag{30}$$

where ε depends on the measurement device used but is never zero.

If we assume ε to be much smaller than the extension of the wave function before the measurement (here a), the wave function after the measurement will be practically $\sqrt{\varepsilon}\ \delta^{(\varepsilon)}\left(x - \frac{a}{2}\right)$ [$\delta^{(\varepsilon)}(x)$ is the null function everywhere except in the interval defined in (30), where it takes on the value $1/\varepsilon$; *cf.* Appendix II, § 1-a]. This wave function is indeed normalized since:

$$\int \mathrm{d}x \left|\sqrt{\varepsilon}\ \delta^{(\varepsilon)}\left(x - \frac{a}{2}\right)\right|^2 = 1 \tag{31}$$

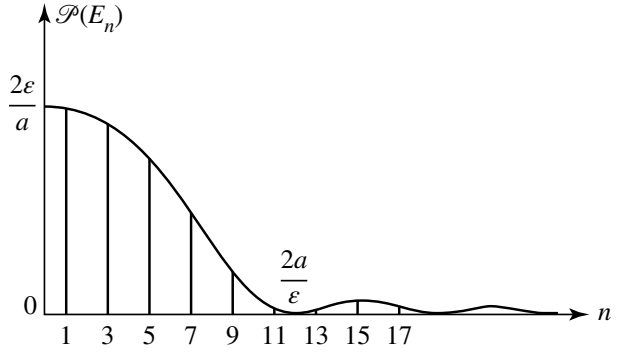


Figure 7: Variation with n of the probability $\mathcal{P}(E_n)$ of finding the energy E_n after a measurement of the particle's position has yielded the result $a/2$ with an accuracy of $\varepsilon(\varepsilon \ll a)$. The smaller ε, the greater the probability of finding high energy values.

What happens now if the energy is measured? Each value E_n can be found with

[4] We see concretely in this example that a δ-function cannot represent a physically realizable state.

the probability:

$$\mathcal{P}(E_n) = \left| \int \varphi_n^*(x)\sqrt{\varepsilon}\ \delta^{(\varepsilon)}\left(x - \frac{a}{2}\right) \mathrm{d}x \right|^2$$

$$= \begin{cases} \dfrac{8a}{\varepsilon}\left(\dfrac{1}{n\pi}\right)^2 \sin^2\left(\dfrac{n\pi\varepsilon}{2a}\right) & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases} \tag{32}$$

The variation with respect to n of $\mathcal{P}(E_n)$, for fixed ε and odd n, is shown in Figure 7. This figure shows that the probability $\mathcal{P}(E_n)$ becomes negligible when n is much larger than a/ε. Therefore, however small ε may be, the distribution of probabilities $\mathcal{P}(E_n)$ depends strongly on ε. This is why, in the first argument, where we set $\varepsilon = 0$ at the beginning, we could not obtain the correct result. We also see from the figure that the smaller ε is, the more the curve extends towards large values of n. The interpretation of this result is the following: according to Heisenberg's uncertainty relations (*cf.* Chap. I, § C-3), if one measures the position of the particle with great accuracy, one drastically changes its momentum. Thus kinetic energy is transferred to the particle, the amount increasing as ε decreases.

Complement B_{III}

Study of the probability current in some special cases

1 Expression for the current in constant potential regions . . 283

2 Application to potential step problems 284

2-a Case where $E > V_0$ 284

2-b Case where $E < V_0$ 285

3 Probability current of incident and evanescent waves, in the case of reflection from a two-dimensional potential step 285

The probability current associated with a particle having a wave function $\psi(\mathbf{r}, t)$ was defined in Chapter III by the relation:

$$\mathbf{J}(\mathbf{r}, t) = \frac{\hbar}{2mi} \left[\psi^*(\mathbf{r}, t)\boldsymbol{\nabla}\psi(\mathbf{r}, t) - \text{c.c.}\right] \tag{1}$$

(where c.c. is an abbreviation for complex conjugate). In this complement, we shall study this probability current in greater detail in some special cases: one- and two-dimensional "square" potentials.

1. Expression for the current in constant potential regions

Consider a one-dimensional problem, with a particle of energy E placed in a constant potential V_0. In Complement H_I, we distinguished between several cases.

(*i*) When $E > V_0$, the wave function is written:

$$\psi(x) = A\,\mathrm{e}^{ikx} + A'\,\mathrm{e}^{-ikx} \tag{2}$$

with:

$$E - V_0 = \frac{\hbar^2 k^2}{2m} \tag{3}$$

Substituting (2) into (1), we obtain:

$$J_x = \frac{\hbar k}{m}\left[|A|^2 - |A'|^2\right] \tag{4}$$

The interpretation of this result is simple: the wave function given in (2) corresponds to two plane waves of opposite momenta $p = \pm\hbar k$ with probability densities $|A|^2$ and $|A'|^2$.

(*ii*) When $E < V_0$, we have:

$$\psi(x) = B\,\mathrm{e}^{\rho x} + B'\,\mathrm{e}^{-\rho x} \tag{5}$$

with:

$$V_0 - E = \frac{\hbar^2 \rho^2}{2m} \tag{6}$$

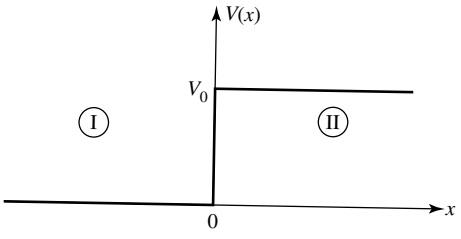


Figure 1: Potential step of height V_0.

Substituting (5) into (1), we obtain:

$$J_x = \frac{\hbar\rho}{m}\left[iB^*B' + \text{c.c.}\right] \tag{7}$$

In this case, we see that the two exponential waves must both have non-zero coefficients for the probability current to be non-zero.

2. Application to potential step problems

Let us apply these results to the potential step problems studied in Complements H_I, and J_I. We shall therefore consider a particle of mass m and energy E propagating in the Ox direction and arriving at $x = 0$ at a potential step of height V_0 (Fig. 1).

2-a. Case where $E > V_0$

Apply formula (4) to wave functions (11) and (12) of Complement H_I, setting, as in that complement:

$$A_2' = 0 \tag{8}$$

In region I, the probability current is:

$$J_I = \frac{\hbar k_1}{m}\left[|A_1|^2 - |A_1'|^2\right] \tag{9}$$

and in region II:

$$J_{II} = \frac{\hbar k_2}{m}|A_2|^2 \tag{10}$$

J_I is the difference of two terms, the first one corresponding to the incident current and the second, to the reflected current. The ratio of these two currents gives the reflection coefficient R of the barrier:

$$R = \left|\frac{A_1'}{A_1}\right|^2 \tag{11}$$

which is precisely formula (15) of Complement H_I.

Similarly, the transmission coefficient T of the barrier is the ratio of the transmitted current J_{II} to the incident current; we therefore have:

$$T = \frac{k_2}{k_1}\left|\frac{A_2}{A_1}\right|^2 \tag{12}$$

and we again find relation (16) of Complement $\mathrm{H_I}$.

2-b. Case where $E < V_0$

Since the expression for the wave function $\varphi_1(x)$ is the same as in § 2-a, relation (9) is still valid. However, in region II, the wave function is:

$$\varphi_{\mathrm{II}}(x) = B_2' \, \mathrm{e}^{-\rho_2 x} \tag{13}$$

[since, in equation (20) of Complement $\mathrm{H_I}$, $B_2 = 0$]. Using (7), we thus obtain:

$$J_{\mathrm{II}} = 0 \tag{14}$$

The transmitted flux is zero, as is consistent with relation (24) of $\mathrm{H_I}$.

How should we interpret the fact that, in region II, the probability current is zero while the probability of finding the particle in this region is not? Let us refer to the results obtained in § 1 of Complement $\mathrm{J_I}$. We saw that part of the incident wave packet enters the classically forbidden region II, and then turns back, before setting out in the negative x direction (this incursion into region II being responsible for the delay upon reflection). In the steady state, we shall therefore have two probability currents in region II: a positive current corresponding to the entrance into this region of part of the incident wave packet; a negative current corresponding to the return towards region I of this part of the wave packet. These two currents are exactly equal, so the overall result is zero.

In the case of a one-dimensional problem, the structure of the probability current of the evanescent wave is therefore masked by the fact that the two opposite currents balance. This is why we are going to consider a two-dimensional problem, for the case of oblique reflection, so as to obtain a non-zero current and interpret its structure.

3. Probability current of incident and evanescent waves, in the case of reflection from a two-dimensional potential step

We shall consider the following two-dimensional problem: a particle of mass m, moving in the xOy plane, has a potential energy $V(x, y)$ which is independent of y and given by:

$$\begin{aligned} &V(x, y) = 0 \quad \text{if} \quad x < 0 \\ &V(x, y) = V_0 \quad \text{if} \quad x > 0 \end{aligned} \tag{15}$$

The present case corresponds to the one studied in § 2 of Complement $\mathrm{F_I}$: the potential energy $V(x, y)$ is the sum of a function $V_1(x)$ (potential energy of a one-dimensional step) and a function $V_2(y)$, which is zero here. We can therefore look for a solution of the eigenvalue equation of the Hamiltonian in the form of a product:

$$\varphi(x, y) = \varphi_1(x)\varphi_2(y) \tag{16}$$

The functions $\varphi_1(x)$ and $\varphi_2(y)$ satisfy one-dimensional eigenvalue equations which correspond respectively to $V_1(x)$ and $V_2(y)$ and to energies E_1 and E_2 such that:

$$E_1 + E_2 = E \quad \text{(total energy of the particle)} \tag{17}$$

We shall assume $E_1 < V_0$: the equation giving $\varphi_1(x)$ therefore corresponds to total reflection in a one-dimensional problem, and we can use formulas (11) and (20) of Complement H_I. As for the function $\varphi_2(y)$, it can be obtained immediately since it corresponds to the case of a free particle ($V_2 = 0$): it is a plane wave. We therefore have, in region I ($x < 0$):

$$\varphi_1(x, y) = A\, e^{i(k_x x + k_y y)} + A'\, e^{i(-k_x x + k_y y)} \tag{18}$$

with:

$$k_x = \sqrt{\frac{2mE_1}{\hbar^2}} \qquad k_y = \sqrt{\frac{2mE_2}{\hbar^2}} \tag{19}$$

and, in region II ($x > 0$):

$$\varphi_{II}(x, y) = B\, e^{-\rho_x x}\, e^{ik_y y} \tag{20}$$

with:

$$\rho_x = \sqrt{\frac{2m(V_0 - E_1)}{\hbar^2}} \tag{21}$$

Equations (22) and (23) of H_I give us the ratios A'/A and B/A. Introducing the parameter θ defined by:

$$\tan\theta = \frac{\rho_x}{k_x} = \sqrt{\frac{V_0 - E_1}{E_1}}; \qquad 0 \leq \theta \leq \frac{\pi}{2} \tag{22}$$

we obtain:

$$\frac{A'}{A} = \frac{k_x - i\rho_x}{k_x + i\rho_x} = e^{-2i\theta} \tag{23}$$

and:

$$\frac{B}{A} = \frac{2k_x}{k_x + i\rho_x} = 2\cos\theta\, e^{-i\theta} \tag{24}$$

Let us apply relation (1), which defines the probability current. We obtain, in region I:

$$\mathbf{J}_I \begin{cases} (J_I)_x = \dfrac{\hbar k_x}{m}\left[|A|^2 - |A'|^2\right] = 0 \\ (J_I)_y = \dfrac{\hbar k_y}{m}\left|A\, e^{ik_x x} + A'\, e^{-ik_x x}\right|^2 = \dfrac{\hbar k_y}{m}\ |A|^2\ [2 + 2\cos(2k_x x + 2\theta)] \end{cases} \tag{25}$$

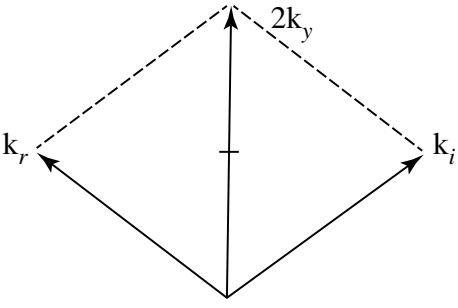


Figure 2: The sum of the probability currents associated with the incident and reflected waves yields a probability current parallel to Oy.

and in region II:

$$\mathbf{J}_{\mathrm{II}} \begin{cases} (J_{\mathrm{II}})_x = 0 \\ (J_{\mathrm{II}})_y = \dfrac{\hbar k_y}{m} \, |B|^2 \, \mathrm{e}^{-2\rho_x x} = \dfrac{\hbar k_y}{m} \, 4|A|^2 \, \cos^2\theta \, \mathrm{e}^{-2\rho_x x} \end{cases} \tag{26}$$

In region I, only the $(J_{\mathrm{I}})_y$ component of the probability current is non-zero; this component is the sum of two terms:

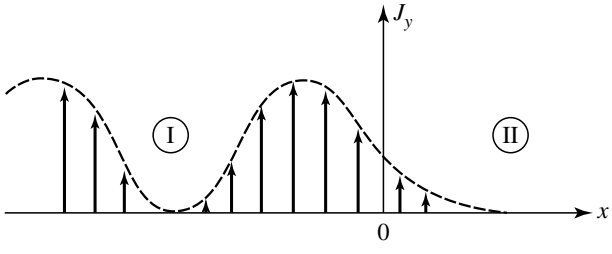


Figure 3: Because of the interference between the incident and reflected waves, the probability current in region I is an oscillatory function of x; in region II, it decreases exponentially (evanescent wave).

– the term proportional to $2|A|^2$ which results from the sum of the currents of the incident and reflected waves (*cf.* Fig. 2);

– the term containing $\cos(2k_x x + 2\theta)$, which represents an interference effect between the two waves; it is responsible for the oscillation of the probability current with respect to x (*cf.* Fig. 3).

In region II, the probability current is again parallel to Oy. Its exponential decay corresponds to the decay of the evanescent wave. This probability current arises from the fact that the wave packets do enter the second region (*cf.* Fig. 4) and, before turning back, propagate in the Oy direction for a time of the order of the reflection delay τ [*cf.* Complement $\mathrm{J_I}$, equation (10)]. This penetration is also related to the lateral shift of the wave packet upon reflection (*cf.* Fig. 4).

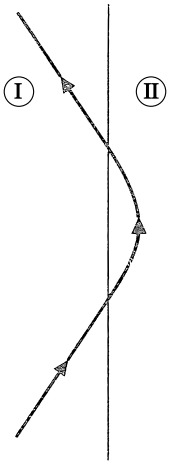


Figure 4: The penetration of the particle into region II leads to a lateral shift upon reflection.

Complement C_{III}

Root mean square deviations of two conjugate observables

1 The Heisenberg relation for P and Q 289

2 The "minimum" wave packet 290

Two conjugate observables P and Q are two observables whose commutator $[Q, P]$ is equal to $i\hbar$. We shall show in this complement that the root mean square deviations (*cf.* § C-5 of Chapter III) ΔP and ΔQ, for any state vector of the system under study, satisfy the relation:

$$\Delta P.\Delta Q \geq \frac{\hbar}{2} \tag{1}$$

We shall then show that if the system is in a state where the product $\Delta P.\Delta Q$ is exactly equal to $\hbar/2$, the wave function associated with this state in the $\{|q\rangle\}$ representation is a Gaussian wave packet (as is the wave function in the $\{|p\rangle\}$ representation).

1. The Heisenberg relation for P and Q

Consider the ket:

$$|\varphi\rangle = (Q + i\lambda P)|\psi\rangle \tag{2}$$

where λ is an arbitrary real parameter. For all λ, the square of the norm $\langle\varphi|\varphi\rangle$ is positive. This is written:

$$\begin{aligned}\langle\varphi|\varphi\rangle &= \langle\psi|(Q - i\lambda P)(Q + i\lambda P)|\psi\rangle \\ &= \langle\psi|Q^2|\psi\rangle + \langle\psi|(i\lambda QP - i\lambda PQ)|\psi\rangle + \langle\psi|\lambda^2 P^2|\psi\rangle \\ &= \langle Q^2\rangle + i\lambda\langle[Q, P]\rangle + \lambda^2\langle P^2\rangle \\ &= \langle Q^2\rangle - \lambda\hbar + \lambda^2\langle P^2\rangle \geq 0\end{aligned} \tag{3}$$

The discriminant of this expression, of second order in λ, is therefore negative or zero:

$$\hbar^2 - 4\langle P^2\rangle\,\langle Q^2\rangle \leq 0 \tag{4}$$

and we have:

$$\langle P^2\rangle\,\langle Q^2\rangle \geq \frac{\hbar^2}{4} \tag{5}$$

Assuming $|\psi\rangle$ to be given, let us now introduce the two observables Q' and P' defined by:

$$P' = P - \langle P \rangle = P - \langle \psi | P | \psi \rangle$$
$$Q' = Q - \langle Q \rangle = Q - \langle \psi | Q | \psi \rangle \tag{6}$$

P' and Q' are also conjugate observables, since we have:

$$[Q', P'] = [Q, P] = i\hbar \tag{7}$$

Result (5), obtained above for P and Q, is therefore also valid for P' and Q':

$$\langle P'^2 \rangle \langle Q'^2 \rangle \geq \frac{\hbar^2}{4} \tag{8}$$

In addition, referring to definition (C-23) (Chap. III) of the root mean square deviation and using (6), we see that:

$$\Delta P = \sqrt{\langle P'^2 \rangle}$$
$$\Delta Q = \sqrt{\langle Q'^2 \rangle} \tag{9}$$

Relation (8) can therefore also be written:

$$\boxed{\Delta P . \Delta Q \geq \frac{\hbar}{2}} \tag{10}$$

Thus, if two observables are conjugate (as is the case when they correspond to a classical position x_i and its conjugate momentum p_i), there exists an exact lower bound for the product $\Delta P . \Delta Q$. We thus generalize the Heisenberg uncertainty relation.

Comment:

This argument can easily be generalized to two arbitrary observables A and B. One obtains:

$$\Delta A . \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle| \tag{11}$$

2. The "minimum" wave packet

When the minimum value of the product $\Delta P . \Delta Q$ is attained:

$$\Delta P . \Delta Q = \frac{\hbar}{2} \tag{12}$$

the state vector $|\psi\rangle$ is said to correspond to a minimum wave packet for the observables P and Q.

According to the preceding argument, relation (12) requires that the square of the norm of the ket:

$$|\varphi'\rangle = (Q' + i\lambda P')|\psi\rangle \tag{13}$$

be a second-order polynomial in λ with a double root λ_0. When $\lambda = \lambda_0$, the ket $|\varphi'\rangle$ is therefore zero:

$$(Q' + i\lambda_0 P')|\psi\rangle = [Q - \langle Q\rangle + i\lambda_0(P - \langle P\rangle)]|\psi\rangle = 0 \tag{14}$$

On the other hand, if $\Delta P.\Delta Q > \hbar/2$, the polynomial which gives $\langle\varphi'|\varphi'\rangle$ can never be equal to zero (it is positive for all λ).

Therefore, the necessary and sufficient condition for the product $\Delta P.\Delta Q$ to take on its minimum value $\hbar/2$ is that the kets $(Q-\langle Q\rangle)|\psi\rangle$ and $(P-\langle P\rangle)|\psi\rangle$ be proportional. The proportionality coefficient $-i\lambda_0$ can easily be calculated. When $\Delta Q.\Delta P = \hbar/2$, the equation:

$$\langle\varphi'|\varphi'\rangle = \lambda^2(\Delta P)^2 - \lambda\hbar + (\Delta Q)^2 = 0 \tag{15}$$

has for its double root:

$$\lambda_0 = \frac{\hbar}{2(\Delta P)^2} = \frac{2(\Delta Q)^2}{\hbar} \tag{16}$$

Let us write relation (14) in the $\{|q\rangle\}$ representation (for simplicity, we assume that the eigenvalues q of Q to be non-degenerate). Using the fact (*cf.* Complement $\mathrm{E_{II}}$) that in this representation, P acts like $\dfrac{\hbar}{i}\dfrac{\mathrm{d}}{\mathrm{d}q}$, we obtain:

$$\left[q + \hbar\lambda_0\frac{\mathrm{d}}{\mathrm{d}q} - \langle Q\rangle - i\lambda_0\langle P\rangle\right]\psi(q) = 0 \tag{17}$$

with:

$$\psi(q) = \langle q|\psi\rangle \tag{18}$$

To integrate equation (17), it is convenient to introduce the function $\theta(q)$ defined by:

$$\psi(q) = \mathrm{e}^{i\langle P\rangle q/\hbar}\theta(q - \langle Q\rangle) \tag{19}$$

Substituting (19) into (17), we thus obtain a more simple equation:

$$\left[q + \lambda_0\hbar\frac{\mathrm{d}}{\mathrm{d}q}\right]\theta(q) = 0 \tag{20}$$

whose solution is:

$$\theta(q) = C\,\mathrm{e}^{-q^2/2\lambda_0\hbar} \tag{21}$$

(where C is an arbitrary complex constant). Substituting (16) and (21) into (19), we obtain:

$$\psi(q) = C\,\mathrm{e}^{i\langle P\rangle q/\hbar}\,\mathrm{e}^{-\left[\frac{q-\langle Q\rangle}{2\Delta Q}\right]^2} \tag{22}$$

This function can be normalized by setting:

$$C = \left[2\pi(\Delta Q)^2\right]^{-1/4} \tag{23}$$

We thus arrive at the following conclusion: when the product $\Delta P.\Delta Q$ takes on its minimum value $\hbar/2$, the wave function in the $\{|q\rangle\}$ representation is a Gaussian wave packet, obtained from the Gaussian function $\theta(q)$ by transformation (19) (which is equivalent to two changes of the origin, one on the q-axis and one on the p-axis).

Comment:

This argument in the $\{|q\rangle\}$ representation can be repeated in the $\{|p\rangle\}$ representation. One then finds that the wave function $\overline{\psi}(p)$ defined by:

$$\overline{\psi}(p) = \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}q\, \mathrm{e}^{-ipq/\hbar} \psi(q) \tag{24}$$

is also a Gaussian function, given by:

$$\overline{\psi}(p) = \left[2\pi(\Delta P)^2\right]^{-1/4} \mathrm{e}^{-i\langle Q\rangle p/\hbar}\, \mathrm{e}^{-\left[\frac{p-\langle P\rangle}{2\Delta P}\right]^2} \tag{25}$$

to within a phase factor $\exp(i\langle Q\rangle\langle P\rangle/\hbar)$.

Complement D_{III}

Measurements bearing on only one part of a physical system

1 Calculation of the physical predictions 293
2 Physical meaning of a tensor product state 295
3 Physical meaning of a state that is not a tensor product . . 296

The concept of a tensor product, introduced in § F of Chapter II, enabled us to see how to construct, starting with the state spaces of two subsystems, that of the global system obtained by considering them together. We intend to pursue this study here, using the postulates of Chapter III to see what results can be obtained, when the state of the global system is known, from measurements bearing on only one subsystem.

1. Calculation of the physical predictions

Consider a system composed of two parts (1) and (2) (for example, a system of two electrons). If $\mathcal{E}(1)$ and $\mathcal{E}(2)$ are the state spaces of parts (1) and (2), the state space of the global system (1)+(2) is the tensor product $\mathcal{E}(1) \otimes \mathcal{E}(2)$. For example, the state of a two-electron system is described by a wave function of six variables, $\psi(x_1, y_1, z_1 ; x_2, y_2, z_2)$, associated with a ket of $\mathcal{E}_{\mathbf{r}}(1) \otimes \mathcal{E}_{\mathbf{r}}(2)$ (*cf.* Chap. II, § F-4-b).

It is possible to imagine measurements that bear on only one of the two parts [part (1), for example] of the global system. The observables $\tilde{A}(1)$ corresponding to these measurements are defined in $\mathcal{E}(1) \otimes \mathcal{E}(2)$ by extending[1] the observables $A(1)$ acting only in $\mathcal{E}(1)$ (*cf.* Chap. II, § F-2-b):

$$A(1) \Longrightarrow \tilde{A}(1) = A(1) \otimes \mathbb{1}(2) \tag{1}$$

where $\mathbb{1}(2)$ is the identity operator in $\mathcal{E}(2)$.

The spectrum of $\tilde{A}(1)$ in $\mathcal{E}(1) \otimes \mathcal{E}(2)$ is the same as that of $A(1)$ in $\mathcal{E}(1)$. On the other hand, we have seen that all the eigenvalues of $\tilde{A}(1)$ are degenerate in $\mathcal{E}(1) \otimes \mathcal{E}(2)$, even if none of the eigenvalues of $A(1)$ is degenerate in $\mathcal{E}(1)$ [on the condition, of course, that the dimension of $\mathcal{E}(2)$ be greater than 1]. When a measurement is made on system (1) alone, the global system may therefore be in several different states after the measurement, whatever the result (the state after the measurement depends not only on the result but also on the state before the measurement). From a physical point of view, this multiplicity of states corresponds to the degrees of freedom of system (2), about which no information is sought in the measurement.

Let $P_n(1)$ be the projector, in $\mathcal{E}(1)$, onto the eigensubspace related to the eigenvalue a_n of $A(1)$:

$$P_n(1) = \sum_{i=1}^{g_n} |u_n^i(1)\rangle \langle u_n^i(1)| \tag{2}$$

[1] For the sake of clarity, we shall adopt throughout this complement different notations for $A(1)$ and its extension $\tilde{A}(1)$.

where the kets $|u_n^i(1)\rangle$ are g_n orthonormal eigenvectors associated with a_n. Let $\tilde{P}_n(1)$ be the projector, in $\mathcal{E}(1) \otimes \mathcal{E}(2)$, onto the eigensubspace related to the same eigenvalue a_n of $\tilde{A}(1)$. $\tilde{P}_n(1)$ is obtained by extending $P_n(1)$ into $\mathcal{E}(1) \otimes \mathcal{E}(2)$:

$$\tilde{P}_n(1) = P_n(1) \otimes \mathbb{1}(2) \tag{3}$$

To write the identity operator $\mathbb{1}(2)$ of $\mathcal{E}(2)$, let us use the closure relation for an arbitrary orthonormal basis $\{|v_k(2)\rangle\}$ of $\mathcal{E}(2)$:

$$\mathbb{1}(2) = \sum_k |v_k(2)\rangle \langle v_k(2)| \tag{4}$$

Substituting (4) into (3) and using (2), we obtain:

$$\tilde{P}_n(1) = \sum_{i=1}^{g_n} \sum_k |u_n^i(1) v_k(2)\rangle \langle u_n^i(1) v_k(2)| \tag{5}$$

Thus, knowing the state $|\psi\rangle$ of the global system (assumed to be normalized to 1), we can calculate the probability $\mathcal{P}^{(1)}(a_n)$ of finding the result a_n in a measurement of $A(1)$ on part (1) of this system. Using general formula (B-14) of Chapter III, which here gives:

$$\mathcal{P}^{(1)}(a_n) = \langle \psi | \tilde{P}_n(1) | \psi \rangle \tag{6}$$

we find:

$$\mathcal{P}^{(1)}(a_n) = \sum_{i=1}^{g_n} \sum_k |\langle u_n^i(1) v_k(2) | \psi \rangle|^2 \tag{7}$$

Similarly, the state $|\psi'\rangle$ of the system after the measurement can be calculated; according to formula (B-31) of Chapter III, it is given by:

$$|\psi'\rangle = \frac{\tilde{P}_n(1)|\psi\rangle}{\sqrt{\langle \psi | \tilde{P}_n(1) | \psi \rangle}} \tag{8}$$

that is, using (5):

$$|\psi'\rangle = \frac{\sum_{i=1}^{g_n} \sum_k |u_n^i(1) v_k(2)\rangle \langle u_n^i(1) v_k(2) | \psi \rangle}{\sqrt{\sum_{i=1}^{g_n} \sum_k |\langle u_n^i(1) v_k(2) | \psi \rangle|^2}} \tag{9}$$

Comments:

(*i*) The choice of an orthonormal basis $\{|v_k(2)\rangle\}$ in $\mathcal{E}(2)$ is arbitrary. We see from (3), (6) and (8) that the predictions concerning subsystem (1) do not depend on this choice. Physically, it is clear that if no measurement is performed on system (2), no state or set of states of this system can play a preferential role.

(*ii*) If the state $|\psi\rangle$ before the measurement is a tensor product:

$$|\psi\rangle = |\varphi(1)\rangle \otimes |\chi(2)\rangle \tag{10}$$

[where $|\varphi(1)\rangle$ and $|\chi(2)\rangle$ are two normalized states of $\mathcal{E}(1)$ and $\mathcal{E}(2)$ respectively], it is easy to see, using (3) and (8), that the state $|\psi'\rangle$ is also a tensor product:

$$|\psi'\rangle = |\varphi'(1)\rangle \otimes |\chi(2)\rangle \tag{11}$$

with:

$$|\varphi'(1)\rangle = \frac{P_n(1)|\varphi(1)\rangle}{\sqrt{\langle\varphi(1)|P_n(1)|\varphi(1)\rangle}} \tag{12}$$

The state of system (1) has therefore changed, but not that of system (2).

(*iii*) If the eigenvalue a_n of $A(1)$ is non-degenerate in $\mathcal{E}(1)$ – or, more generally, if $A(1)$ actually represents a complete set of commuting observables of $\mathcal{E}(1)$ – the index i is no longer necessary in formula (2) and those which follow. The state of the system after a measurement yielding the result a_n can always be put in the form of a product of two vectors. This can be seen by writing relation (9) in the form:

$$|\psi'\rangle = |u_n(1)\rangle \otimes |\chi'(2)\rangle \tag{13}$$

where the normalized vector $|\chi'(2)\rangle$ of $\mathcal{E}(2)$ is given by:

$$|\chi'(2)\rangle = \frac{\sum\limits_k |v_k(2)\rangle \langle u_n(1)v_k(2)|\psi\rangle}{\sqrt{\sum\limits_k |\langle u_n(1)v_k(2)|\psi\rangle|^2}} \tag{14}$$

Therefore, whatever the state $|\psi\rangle$ of the global system before the measurement, the state of the system after a measurement bearing on part (1) alone is always a tensor product when this measurement is complete with respect to part (1) [although partial as regards the global system (1) + (2)].

2. Physical meaning of a tensor product state

To see what a product state represents physically, let us apply the results of the preceding paragraph to the particular case where the initial state of the global system is of the form (10). We immediately obtain, using (6) and (3):

$$\mathcal{P}^{(1)}(a_n) = \langle\varphi(1)\chi(2)|P_n(1) \otimes \mathbb{1}(2)|\varphi(1)\chi(2)\rangle \tag{15}$$

The very definition of the tensor product $P_n(1) \otimes \mathbb{1}(2)$ and the fact that $|\chi(2)\rangle$ is normalized then allow us to write:

$$\begin{aligned}\mathcal{P}^{(1)}(a_n) &= \langle\varphi(1)|P_n(1)|\varphi(1)\rangle \langle\chi(2)|\mathbb{1}(2)|\chi(2)\rangle \\ &= \langle\varphi(1)|P_n(1)|\varphi(1)\rangle\end{aligned} \tag{16}$$

$\mathcal{P}^{(1)}(a_n)$ does not depend on $|\chi(2)\rangle$, but only on $|\varphi(1)\rangle$. When the state of the global system has the form (10), all physical predictions relating to only one of the two systems do not, therefore, depend on the state of the other one and are expressed entirely in terms of $|\varphi(1)\rangle$ [or of $|\chi(2)\rangle$], depending on whether it is system (1) alone [or system (2) alone] that is being observed.

A product state $|\varphi(1)\rangle \otimes |\chi(2)\rangle$ can therefore be considered to represent the simple juxtaposition of two systems, one in the state $|\varphi(1)\rangle$ and the other in the state $|\chi(2)\rangle$. In such a state, the two systems are also said to be *uncorrelated* (more precisely, the results of the two types of measurements, bearing either on one system or on the other, correspond to *independent random variables*). Such a situation is realized when the two systems have been separately prepared in the states $|\varphi(1)\rangle$ and $|\chi(2)\rangle$ and then united without interacting.

3. Physical meaning of a state that is not a tensor product

Now consider the case in which the state of the global system is not a product state, that is, where $|\psi\rangle$ cannot be written in the form $|\varphi(1)\rangle \otimes |\chi(2)\rangle$. The predictions of measurement results bearing on only one of the two systems can then no longer be expressed in terms of a ket $|\varphi(1)\rangle$ [or $|\chi(2)\rangle$] in which system (1) [or (2)] would be found. In this case, general formulas (6) and (7) must be used to find the probabilities of the various possible results. We assume here without proof that such a situation generally reflects the existence of *correlations* between systems (1) and (2). The results of measurements bearing on either system (1) or system (2) correspond to random variables which are not independent and can therefore be correlated. It can be shown, for example, that an interaction between the two systems transforms an initial state which is a product into one which is no longer a product: any interaction between two systems therefore introduces, in general, correlations between them.

When the state of the global system is not a product $|\varphi(1)\rangle \otimes |\chi(2)\rangle$, how can each partial system (1) or (2) be characterized, since the ket $|\varphi(1)\rangle$ or $|\chi(2)\rangle$ can no longer be associated with it? This question is very important since, in general, every physical system has interacted with others in the past (even if it is isolated at the instant when it is being studied). The state of the global system: { system (1) + system (2) with which it has interacted in the past } is therefore not in general a product state, and it is not possible to associate a state vector $|\varphi(1)\rangle$ with system (1) alone. To resolve these difficulties, one must describe system (1), not by a state vector, but by an operator, called the *density operator*. The corresponding formalism, fundamental to statistical quantum mechanics, is introduced in Complement E_{III} (*cf.* § 5-b).

However, system (1) can always be described by a state vector after a complete set of measurements has been performed on it. We have seen that whatever the state of the global system (1) + (2) before the measurement, a complete measurement on system (1) places the global system in a product state [*cf.* formulas (13) and (14)]. The vector associated with (1) is the unique eigenvector (to within a multiplicative factor) associated with the results of the complete set of measurements done on it. This set of measurements has therefore erased all correlations resulting from previous interactions between the two systems. If, at the moment of measurement, system (2) is already far away and no longer interacting with system (1), it can then be completely forgotten if one is interested only in system (1).

Comment:

Nervertheless, the quantum state $|\chi'(2)\rangle$ reached by system (2) depends in general[2] on the result of the measurements performed on system (1). This is easy to check with (14), when the state $|\psi\rangle$ before the measurement is not a product state. Such a result may seem extremely surprising: how can the state of system (2) after a series of measurements made on system (1) change as a function of the result of measurements performed on an arbitrarily remote system (1), with which is does not interact any longer? To this "paradox", studied in detail by many physicists, are attached the names of Einstein, Podolsky and Rosen; it is discussed in § F-1 of Chapter XXI.

References and suggestions for further reading:

The Einstein-Podolsky-Rosen paradox/argument: see also *Do we really understand quantum mechanics?*, F. Laloë, Cambridge University Press (2018), as well the subsection "Hidden variables and paradoxes" of section 5 of the bibliography; Bohm (5.1), §§ 22.15 to 22.19; d'Espagnat (5.3), Chap. 7.

Photons produced in the decay of positronium: Feynman III (1.2), § 18.3; Dicke and Wittke (1.14), Chap. 7.

[2]Recall that this is not the case when $|\psi\rangle$ is a product state; *cf.* comment (*ii*) of § 1

Complement E_{III}

The density operator

1 Outline of the problem 299
2 The concept of a statistical mixture of states 299
3 The pure case. Introduction of the density operator 301
3-a Description by a state vector 301
3-b Description by a density operator 302
3-c Properties of the density operator in a pure case 303
4 A statistical mixture of states (non-pure case) 304
4-a Definition of the density operator 304
4-b General properties of the density operator 305
4-c Populations; coherences 307
5 Use of the density operator: some applications 308
5-a System in thermodynamic equilibrium 308
5-b Separate description of part of a physical system. Concept of a partial trace 309

1. Outline of the problem

Until now, we have considered systems whose state is perfectly well known. We have shown how to study their time evolution and how to predict the results of various measurements performed on them. To determine the state of a system at a given instant, it suffices to perform on the system a set of measurements corresponding to a C.S.C.O. For example, in the experiment studied in § A-3 of Chapter I, the polarization state of the photons is perfectly well known when the light beam has traversed the polarizer.

However, in practice, the state of the system is often not perfectly determined. This is true, for example, of the polarization state of photons coming from a source of natural (unpolarized) light, and also for the atoms of a beam emitted by a furnace at temperature T, where the atoms' kinetic energy is known only statistically. The problem posed by the quantum description of such systems is the following: how can we incorporate into the formalism the *incomplete information* we possess about the state of the system, so that our predictions make maximum use of this partial information? To do this, we shall introduce here a very useful mathematical tool, the density operator, which facilitates the simultaneous application of the postulates of quantum mechanics and the results of probability calculations.

2. The concept of a statistical mixture of states

When one has incomplete information about a system, one typically appeals to the concept of probability. For example, we know that a photon emitted by a source of

natural light can have any polarization state with equal probability. Similarly, a system in thermodynamic equilibrium at a temperature T has a probability proportional to $e^{-E_n/kT}$ of being in a state of energy E_n.

More generally, the incomplete information one has about the system usually presents itself, in quantum mechanics, in the following way: the state of this system may be either the state $|\psi_1\rangle$ with a probability p_1 or the state $|\psi_2\rangle$ with a probability p_2, etc... Obviously:

$$p_1 + p_2 + ... = \sum_k p_k = 1 \tag{1}$$

We then say that we are dealing with a *statistical mixture* of states $|\psi_1\rangle$, $|\psi_2\rangle$, ... with probabilities p_1, p_2, ...

Now let us see what happens to the predictions concerning the results of measurements performed on this system. If the state of the system were $|\psi_k\rangle$, we could use the postulates stated in Chapter III to determine the probability of obtaining one or another measurement result. Since such a possibility (the state $|\psi_k\rangle$) has a probability of p_k, it is clear that the results obtained must be weighted by the p_k and then summed over the various values of k, that is, over all the states of the statistical mixture.

Comments:

(*i*) The various states $|\psi_1\rangle$, $|\psi_2\rangle$, ... are not necessarily orthogonal. However, they can always be chosen normalized; in this complement, we shall assume that this is the case.

(*ii*) Note that, in the present case, probabilities intervene at two different levels:

- first, in the initial information about the system (until now, we have not introduced probabilities at this stage: we considered the state vector to be perfectly well known, in which case all the probabilities p_k are zero, except one, which is equal to 1);
- again, when the postulates concerning the measurement are applied (leading to probabilistic predictions, even if the initial state of the system is perfectly well known).

There are thus two totally different reasons necessitating the introduction of probabilities: the incomplete nature of the initial information about the state of the system (such situations are also envisaged in classical statistical mechanics), and the (specifically quantum mechanical) uncertainty related to the measurement process.

(*iii*) A system described by a statistical mixture of states (with the probability p_k of the state vector being $|\psi_k\rangle$) must not be confused with a system whose state $|\psi\rangle$ is a linear superposition of states[1]:

$$|\psi\rangle = \sum_k c_k |\psi_k\rangle \tag{2}$$

[1] We assume, in this comment (*iii*), that the states $|\psi_k\rangle$ are orthonormal. This hypothesis is not essential but it simplifies the discussion.

It is often said in quantum mechanics, when the state vector is the ket $|\psi\rangle$ given in (2), that "the system has a probability $|c_k|^2$ of being in the state $|\psi_k\rangle$".

If we want to be precise, this must be understood to mean that if we perform a set of measurements corresponding to a C.S.C.O. which has $|\psi_k\rangle$ as an eigenvector, the probability of finding the set of eigenvalues associated with $|\psi_k\rangle$ is $|c_k|^2$. But we have stressed, in § E-1 of Chapter III, the fact that a system in the state $|\psi\rangle$ given in (2) *is not simply equivalent* to a system having the probability $|c_1|^2$ of being in the state $|\psi_1\rangle$, $|c_2|^2$ of being in the state $|\psi_2\rangle$, etc... In fact, for a linear combination of $|\psi_k\rangle$, there exist, in general, interference effects between these states (due to cross terms of the type $c_k c_{k'}^*$, obtained when the modulus of the probability amplitudes is squared) which are very important in quantum mechanics.

We therefore see that it is impossible, in general, to describe a statistical mixture by an "average state vector" which would be a superposition of the states $|\psi_k\rangle$. As we indicated earlier, when we take *a weighted sum of probabilities*, we can never obtain interference terms between the various states $|\psi_k\rangle$ of a statistical mixture.

3. The pure case. Introduction of the density operator

To study the behavior of a statistical mixture of states, we have envisaged one method: calculation of the physical predictions corresponding to a possible state $|\psi_k\rangle$; weighting the results so obtained by the probability p_k associated with this state and summation over k. Although correct in principle, this method often leads to clumsy calculations. We have indicated [comment (*iii*)] that it is impossible to associate an "average state vector" with the system. Actually, it is an "average operator" and not an "average vector" which permits a simple description of the statistical mixture of states: the density operator.

Before studying this general case, we shall return to the simple case where the state of the system is perfectly known (all the probabilities p_k are zero, except one). The system is then said to be in a *pure state*. We shall show that characterizing the system by its state vector is completely equivalent to characterizing it by a certain operator acting in the state space, the density operator. The usefulness of this operator will become apparent in § 4, where we shall show that nearly all the formulas involving this operator, and derived for the pure case, remain valid for the description of a statistical mixture of states.

3-a. Description by a state vector

Consider a system whose state vector at the instant t is:

$$|\psi(t)\rangle = \sum_n c_n(t)|u_n\rangle \tag{3}$$

where the $\{|u_n\rangle\}$ form an orthonormal basis of the state space, assumed to be discrete (extension to the case of a continuous basis presents no difficulties). The coefficients

$c_n(t)$ satisfy the relation:

$$\sum_n |c_n(t)|^2 = 1 \tag{4}$$

which expresses the fact that $|\psi(t)\rangle$ is normalized.

If A is an observable, with matrix elements:

$$\langle u_n|A|u_p\rangle = A_{np} \tag{5}$$

the mean value of A at the instant t is:

$$\langle A\rangle(t) = \langle\psi(t)|A|\psi(t)\rangle = \sum_{n,p} c_n^*(t)\, c_p(t)\, A_{np} \tag{6}$$

Finally, the evolution of $|\psi(t)\rangle$ is described by the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle \tag{7}$$

where $H(t)$ is the Hamiltonian of the system.

3-b. Description by a density operator

Relation (6) shows that the coefficients $c_n(t)$ enter into the mean values through quadratic expressions of the type $c_n^*(t)\, c_p(t)$. These are simply the matrix elements of the operator $|\psi(t)\rangle\langle\psi(t)|$, the projector onto the ket $|\psi(t)\rangle$ (*cf.* Chap. II, § B-3-b), as can be seen from (3):

$$\langle u_p|\psi(t)\rangle\langle\psi(t)|u_n\rangle = c_n^*(t)\, c_p(t) \tag{8}$$

It is therefore natural to introduce the density operator $\rho(t)$, defined by:

$$\rho(t) = |\psi(t)\rangle\langle\psi(t)| \tag{9}$$

The density operator is represented in the $\{|u_n\rangle\}$ basis by a matrix called the *density matrix* whose elements are:

$$\rho_{pn}(t) = \langle u_p|\rho(t)|u_n\rangle = c_n^*(t)\, c_p(t) \tag{10}$$

We are going to show that the specification of $\rho(t)$ suffices to characterize the quantum state of the system; that is, it enables us to obtain all the physical predictions that can be calculated from $|\psi(t)\rangle$. Let us write formulas (4), (6) and (7) in terms of the operator $\rho(t)$. According to (10), relation (4) indicates that the sum of the diagonal elements of the density matrix is equal to 1:

$$\sum_n |c_n(t)|^2 = \sum_n \rho_{nn}(t) = \mathrm{Tr}\,\rho(t) = 1 \tag{11}$$

In addition, using (5) and (10), formula (6) becomes:

$$\begin{aligned}\langle A\rangle(t) &= \sum_{n,p}\langle u_p|\rho(t)|u_n\rangle\langle u_n|A|u_p\rangle \\ &= \sum_p \langle u_p|\rho(t)A|u_p\rangle \\ &= \mathrm{Tr}\,\{\rho(t)A\}\end{aligned} \tag{12}$$

Finally, the time evolution of the operator $\rho(t)$ can be deduced from the Schrödinger equation (7):

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle\right)\langle\psi(t)| + |\psi(t)\rangle\left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\right) \\ &= \frac{1}{i\hbar}H(t)|\psi(t)\rangle\langle\psi(t)| + \frac{1}{(-i\hbar)}|\psi(t)\rangle\langle\psi(t)|H(t) \\ &= \frac{1}{i\hbar}\left[H(t),\,\rho(t)\right] \end{aligned} \tag{13}$$

Therefore, in terms of the density operator, conservation of probability is expressed by:

$$\mathrm{Tr}\,\rho(t) = 1 \tag{14}$$

The mean value of an observable A is calculated using the formula:

$$\langle A\rangle(t) = \mathrm{Tr}\left\{\rho(t)A\right\} = \mathrm{Tr}\left\{A\rho(t)\right\} \tag{15}$$

and the time evolution obeys the equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = \left[H(t),\,\rho(t)\right] \tag{16}$$

For completeness, we must also indicate how to calculate from $\rho(t)$ the probabilities $\mathcal{P}(a_n)$ of the various results a_n which can be obtained in the measurement of an observable A at time t. Actually, formula (15) can be used to do this. We know [see equation (B-14) of Chapter III] that $\mathcal{P}(a_n)$ can be written as a mean value, that of the projector P_n onto the eigensubspace associated with a_n:

$$\mathcal{P}(a_n) = \langle\psi(t)|P_n|\psi(t)\rangle \tag{17}$$

Using (15), we therefore obtain:

$$\mathcal{P}(a_n) = \mathrm{Tr}\left\{P_n\,\rho(t)\right\} \tag{18}$$

3-c. Properties of the density operator in a pure case

In a pure case, a system can be described just as well by a density operator as by a state vector. However, the density operator presents a certain number of advantages.

First of all, we see from (9) that two state vectors $|\psi(t)\rangle$ and $\mathrm{e}^{i\theta}|\psi(t)\rangle$ (where θ is a real number), which describe the same physical state, correspond to the same density operator. Using this operator therefore eliminates the drawbacks related to the existence of an arbitrary global phase factor for the state vector. Moreover, we see from (14), (15) and (18) that the formulas using the density operator are linear with respect to it, while expressions (6) and (17) are quadratic with respect to $|\psi(t)\rangle$. This is an important property which will be useful subsequently.

Finally, let us mention some properties of $\rho(t)$, which can be deduced directly from its definition (9):

$$\rho^{\dagger}(t) = \rho(t) \tag{19}$$

(the density operator is Hermitian)

$$\rho^2(t) = \rho(t) \tag{20}$$

$$\mathrm{Tr}\,\rho^2(t) = 1 \tag{21}$$

These last two relations, which follow from the fact that $\rho(t)$ is a projector, are true only in a pure case. We shall see later that they are not valid for a statistical mixture of states.

4. A statistical mixture of states (non-pure case)

4-a. Definition of the density operator

We now return to the general case described in § 1, and consider a system for which (at a given instant) the various probabilities p_1, p_2, ... p_k, ... are arbitrary, on the condition that they satisfy the relations:

$$\begin{cases} 0 \leq p_1,\, p_2,\, ...,\, p_k,\, ... \leq 1 \\ \\ \sum\limits_k p_k = 1 \end{cases} \tag{22}$$

Under these conditions, how does one calculate the probability $\mathcal{P}(a_n)$ that a measurement of the observable A will yield the result a_n?

Let:

$$\mathcal{P}_k(a_n) = \langle\psi_k|P_n|\psi_k\rangle \tag{23}$$

be the probability of finding a_n if the state vector were $|\psi_k\rangle$. To obtain the desired probability $\mathcal{P}(a_n)$, one must, as we have already indicated, weight $\mathcal{P}_k(a_n)$ by p_k and then sum over k:

$$\mathcal{P}(a_n) = \sum_k p_k\, \mathcal{P}_k(a_n) \tag{24}$$

Now, from (18), we have:

$$\mathcal{P}_k(a_n) = \mathrm{Tr}\,\{\rho_k\, P_n\} \tag{25}$$

where:

$$\rho_k = |\psi_k\rangle\,\langle\psi_k| \tag{26}$$

is the density operator corresponding to the state $|\psi_k\rangle$. Substituting (25) into (24), we have:

$$\begin{aligned} \mathcal{P}(a_n) &= \sum_k p_k\, \mathrm{Tr}\,\{\rho_k\, P_n\} \\ &= \mathrm{Tr}\left\{\sum_k p_k\, \rho_k\, P_n\right\} \\ &= \mathrm{Tr}\,\{\rho P_n\} \end{aligned} \tag{27}$$

where we have set:

$$\rho = \sum_k p_k \, \rho_k \tag{28}$$

We therefore see that the linearity of the formulas which use the density operator enables us to express all physical predictions in terms of ρ, the average of the density operators ρ_k with weights p_k; ρ is, by definition, the density operator of the system.

Comment:

An ensemble of pure states $|\psi_k\rangle$ with probabilities p_k leads to a single density operator ρ, but the reverse is not true in general: the same density operator can be interpreted as several different statistical mixtures of pure states. For instance, in a space of states with dimension N, a statistical mixture of the N states of a given basis $|u_n\rangle$ with the same probability $1/N$ leads to a density operator that is the unit operator, divided by N; but the same operator may be obtained with a statistical mixture of the kets of any other basis $|v_n\rangle$. These various ensembles lead to the same probabilities $\mathcal{P}(a_n)$, and cannot be distinguished by measuring these probabilities. This situation is sometimes described as the "multiple preparations" of the same density operator.

4-b. General properties of the density operator

Since the coefficients p_k are real, ρ is obviously a Hermitian operator like each of the ρ_k.

Let us calculate the trace of ρ; it is equal to:

$$\mathrm{Tr}\,\rho = \sum_k p_k \mathrm{Tr}\,\rho_k \tag{29}$$

Now, as we saw in § 3-b, the trace of ρ_k is always equal to 1; it follows that:

$$\mathrm{Tr}\,\rho = \sum_k p_k = 1 \tag{30}$$

Relation (14) is therefore valid in the general case.

We have already given, in (27), the expression that enables us to calculate the probability $\mathcal{P}(a_n)$ in terms of ρ. Using this expression, we can easily generalize formula (15) to statistical mixtures:

$$\begin{aligned}\langle A\rangle &= \sum_n a_n \mathcal{P}(a_n) = \mathrm{Tr}\left\{\rho \sum_n a_n P_n\right\} \\ &= \mathrm{Tr}\,\{\rho A\}\end{aligned} \tag{31}$$

[we have used formula (D-36-b) of Chapter II].

Now let us calculate the time evolution of the density operator. To do this, we shall assume that, unlike the state of the system, its Hamiltonian $H(t)$ is perfectly well known. One can then easily show that if the system at the initial time t_0 has the probability p_k of being in the state $|\psi_k\rangle$, then, at a subsequent time t, it has the same probability p_k of being in the state $|\psi_k(t)\rangle$ given by:

$$\begin{cases} i\hbar \dfrac{\mathrm{d}}{\mathrm{d}t}|\psi_k(t)\rangle = H(t)|\psi_k(t)\rangle \\ |\psi_k(t_0)\rangle = |\psi_k\rangle \end{cases} \tag{32}$$

The density operator at the instant t will then be:

$$\rho(t) = \sum_k p_k \rho_k(t) \tag{33}$$

with:

$$\rho_k(t) = |\psi_k(t)\rangle \langle\psi_k(t)| \tag{34}$$

According to (16), $\rho_k(t)$ obeys the evolution equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\rho_k(t) = [H(t), \rho_k(t)] \tag{35}$$

The linearity of formulas (33) and (35) with respect to $\rho_k(t)$ implies that:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = [H(t), \rho(t)] \tag{36}$$

We can therefore generalize to a statistical mixture of states all the equations of § 3, with the exception of (20) and (21). We see that, since ρ is no longer a projector, we have, in general[2]:

$$\rho^2 \neq \rho \tag{37}$$

and, consequently:

$$\mathrm{Tr}\,\rho^2 \leq 1 \tag{38}$$

Moreover, it is sufficient that one of the equations, (20) or (21), be satisfied for us to be sure that we are dealing with a pure state.

Finally, we see from definition (28) that, for any ket $|u\rangle$, we have:

$$\begin{aligned} \langle u|\rho|u\rangle &= \sum_k p_k \langle u|\rho_k|u\rangle \\ &= \sum_k p_k |\langle u|\psi_k\rangle|^2 \end{aligned} \tag{39}$$

and consequently:

$$\langle u|\rho|u\rangle \geq 0 \tag{40}$$

ρ is therefore a positive operator.

[2]Assume, for example, that the states $|\psi_k\rangle$ are mutually orthogonal. In an orthonormal basis including the $|\psi_k\rangle$, ρ is diagonal and its elements are the p_k. To obtain ρ^2, we simply replace p_k by p_k^2. Relations (37) and (38) then follow from the fact that the p_k are always less than 1 (except in the particular case where only one of them is non-zero: the pure case).

4-c. Populations; coherences

What is the physical meaning of the matrix elements ρ_{np} of ρ in the $\{|u_n\rangle\}$ basis? First, let us consider the diagonal element ρ_{nn}. According to (28), we have:

$$\rho_{nn} = \sum_k p_k [\rho_k]_{nn} \tag{41}$$

that is, using (26) and introducing the components:

$$c_n^{(k)} = \langle u_n | \psi_k \rangle \tag{42}$$

of $|\psi_k\rangle$ in the $\{|u_n\rangle\}$ basis:

$$\rho_{nn} = \sum_k p_k \left| c_n^{(k)} \right|^2 \tag{43}$$

$\left| c_n^{(k)} \right|^2$ is a positive real number, whose physical interpretation is the following: if the state of the system is $|\psi_k\rangle$, this number is the probability of finding, in a measurement, this system in the state $|u_n\rangle$. According to (41), if we take into account the indeterminacy of the state before the measurement, ρ_{nn} represents the average probability of finding the system in the state $|u_n\rangle$. For this reason, ρ_{nn} is called the *population* of the state $|u_n\rangle$: if the same measurement is carried out N times under the same initial conditions, where N is a large number, $N\rho_{nn}$ systems will be found in the state $|u_n\rangle$. It is evident from (43) that ρ_{nn} is a positive real number, equal to zero only if all the $|c_n^{(k)}|^2$ are zero.

A calculation analogous to the preceding one gives the following expression for the non-diagonal element ρ_{np}:

$$\rho_{np} = \sum_k p_k \; c_n^{(k)} c_p^{(k)*} \tag{44}$$

$c_n^{(k)} c_p^{(k)*}$ is a cross term, of the same type as those studied in § E-1 of Chapter III. It reflects the interference effects between the states $|u_n\rangle$ and $|u_p\rangle$ which can appear when the state $|\psi_k\rangle$ is a coherent linear superposition of these states. According to (44), ρ_{np} is the average of these cross terms, taken over all the possible states of the statistical mixture. In contrast to the populations, ρ_{np} can be zero even if none of the products $c_n^{(k)} c_p^{(k)*}$ is: while ρ_{nn} is a sum of real positive (or zero) numbers, ρ_{np} is a sum of complex numbers. If ρ_{np} is zero, this means that the average (44) has cancelled out any interference effects between $|u_n\rangle$ and $|u_p\rangle$. On the other hand, if ρ_{np} is different from zero, a certain coherence subsists between these states. This is why the non-diagonal elements of ρ are often called *coherences*.

Comments:

(*i*) The distinction between populations and coherences obviously depends on the basis $\{|u_n\rangle\}$ chosen in the state space. Since ρ is Hermitian, it is always

possible to find an orthonormal basis $\{|\chi_l\rangle\}$ where ρ is diagonal. ρ can then be written:

$$\rho = \sum_l \pi_l \ |\chi_l\rangle \langle\chi_l| \tag{45}$$

Since ρ is positive [relation (40)] and Tr $\rho = 1$, we have:

$$\begin{cases} 0 \leq \pi_l \leq 1 \\ \sum_l \pi_l = 1 \end{cases} \tag{46}$$

ρ can thus be considered to describe a statistical mixture of the states $|\chi_l\rangle$ with the probabilities π_l (there are no coherences between the states $|\chi_l\rangle$).

(*ii*) If the kets $|u_n\rangle$ are eigenvectors of the Hamiltonian H, which is assumed to be time-independent:

$$H|u_n\rangle = E_n|u_n\rangle \tag{47}$$

we obtain directly from (36):

$$\begin{cases} i\hbar \dfrac{\mathrm{d}}{\mathrm{d}t}\rho_{nn}(t) = 0 \\ i\hbar \dfrac{\mathrm{d}}{\mathrm{d}t}\rho_{np}(t) = (E_n - E_p)\rho_{np} \end{cases} \tag{48}$$

that is:

$$\begin{cases} \rho_{nn}(t) = \text{constant} \\ \rho_{np}(t) = \mathrm{e}^{\frac{i}{\hbar}(E_p - E_n)t}\rho_{np}(0) \end{cases} \tag{49}$$

The populations are constant, and the coherences oscillate at the Bohr frequencies of the system.

(*iii*) Using (40), one can prove the inequality:

$$\rho_{nn}\rho_{pp} \geq |\rho_{np}|^2 \tag{50}$$

It follows, for example, that ρ can have coherences only between states whose populations are not zero.

5. Use of the density operator: some applications

5-a. System in thermodynamic equilibrium

The first example we shall consider is borrowed from quantum statistical mechanics. Consider a system in thermodynamic equilibrium with a reservoir at the absolute temperature T. It can be shown that its density operator is then:

$$\rho = Z^{-1}\,\mathrm{e}^{-H/kT} \tag{51}$$

where H is the Hamiltonian operator of the system, k is the Boltzmann constant, and Z is a normalization coefficient chosen so as to make the trace of ρ equal to 1:

$$Z = \mathrm{Tr}\left(\mathrm{e}^{-H/kT}\right) \tag{52}$$

(Z is called the "partition function").

In the $\{|u_n\rangle\}$ basis of eigenvectors of H, we have (*cf.* Complement $\mathrm{B_{II}}$, § 4-a):

$$\begin{aligned}\rho_{nn} &= Z^{-1}\,\langle u_n|\mathrm{e}^{-H/kT}|u_n\rangle \\ &= Z^{-1}\,\mathrm{e}^{-E_n/kT}\end{aligned} \tag{53}$$

and:

$$\begin{aligned}\rho_{np} &= Z^{-1}\,\langle u_n|\mathrm{e}^{-H/kT}|u_p\rangle \\ &= Z^{-1}\,\mathrm{e}^{-E_p/kT}\langle u_n|u_p\rangle \\ &= 0\end{aligned} \tag{54}$$

At thermodynamic equilibrium, the populations of the stationary states are exponentially decreasing functions of the energy (the lower the temperature T, the more rapid the decrease), and the coherences between stationary states are zero.

More details on the use of the density operator in statistical mechanics and at thermal equilibrium are given in Appendix VI of Volume III.

5-b. Separate description of part of a physical system. Concept of a partial trace

We now return to the problem mentioned in § 3 of Complement $\mathrm{D_{III}}$. Consider two different systems (1) and (2) and the global system (1) + (2), whose state space is the tensor product:

$$\mathcal{E} = \mathcal{E}(1) \otimes \mathcal{E}(2) \tag{55}$$

Let $\{|u_n(1)\rangle\}$ be a basis of $\mathcal{E}(1)$ and $\{|v_p(2)\rangle\}$ a basis of $\mathcal{E}(2)$; the kets $|u_n(1)\rangle\ |v_p(2)\rangle$ form a basis of $\mathcal{E}$.

The density operator ρ of the global system is an operator which acts in $\mathcal{E}$. We saw in Chapter II (*cf.* § F-2-b) how to extend into $\mathcal{E}$ an operator which acts only in $\mathcal{E}(1)$ [or $\mathcal{E}(2)$]. We are going to show here how to perform the inverse operation: we shall construct from ρ an operator $\rho(1)$ [or $\rho(2)$] acting only in $\mathcal{E}(1)$ [or $\mathcal{E}(2)$]. This will enable us to make all the physical predictions about measurements bearing only on system (1), or on system (2). This operation will be called a *partial trace* with respect to (2) [or (1)].

Let us introduce the operator $\rho(1)$ whose matrix elements are:

$$\langle u_n(1)|\rho(1)|u_{n'}(1)\rangle = \sum_p \left(\langle u_n(1)|\langle v_p(2)|\right)\rho\left(|u_{n'}(1)\rangle|v_p(2)\rangle\right) \tag{56}$$

By definition, $\rho(1)$ is obtained from ρ by performing a partial trace on (2):

$$\rho(1) = \mathrm{Tr}_2\rho \tag{57}$$

Similarly, the operator:

$$\rho(2) = \mathrm{Tr}_1\rho \tag{58}$$

has matrix elements:

$$\langle v_p(2)|\rho(2)|v_{p'}(2)\rangle = \sum_n \Big(\langle u_n(1)|\langle v_p(2)|\Big)\rho\Big(|u_n(1)\rangle|v_{p'}(2)\rangle\Big) \tag{59}$$

It is clear why these operations are called partial traces. We know that the (total) trace of ρ is:

$$\mathrm{Tr}\rho = \sum_n \sum_p \Big(\langle u_n(1)|\langle v_p(2)|\Big)\rho\Big(|u_n(1)\rangle|v_p(2)\rangle\Big) \tag{60}$$

The difference between (60) and (56) [or (59)] is the following: for the partial traces, the indices n and n' (or p and p') are not required to be equal and the summation is performed only over p (or n). We have, moreover:

$$\mathrm{Tr}\rho = \mathrm{Tr}_1(\mathrm{Tr}_2\rho) = \mathrm{Tr}_2(\mathrm{Tr}_1\rho) \tag{61}$$

$\rho(1)$ and $\rho(2)$ are therefore, like ρ, operators whose trace is equal to 1. It can be verified from their definitions that they are Hermitian, and that they satisfy all the general properties of a density operator (*cf.* § 4-b).

Now let $A(1)$ be an observable acting in $\mathcal{E}(1)$, and $\tilde{A}(1) = A(1)\otimes \mathbb{1}(2)$ its extension in $\mathcal{E}$. We obtain, using (31), the definition of the trace, and the closure relation on the $\{|u_n(1)\rangle|v_p(2)\rangle\}$ basis:

$$\begin{aligned}
\langle \tilde{A}(1)\rangle &= \mathrm{Tr}\,\Big\{\rho\tilde{A}(1)\Big\} \\
&= \sum_{n,p}\sum_{n',p'} \Big(\langle u_n(1)|\langle v_p(2)|\Big)\rho\Big(|u_{n'}(1)\rangle|v_{p'}(2)\rangle\Big) \\
&\qquad\qquad \times \Big(\langle u_{n'}(1)|\langle v_{p'}(2)|\Big)A(1)\otimes\mathbb{1}(2)\Big(|u_n(1)\rangle|v_p(2)\rangle\Big) \\
&= \sum_{n,p,n',p'} \Big(\langle u_n(1)|\langle v_p(2)|\Big)\rho\Big(|u_{n'}(1)\rangle|v_{p'}(2)\rangle\Big) \\
&\qquad\qquad \times \langle u_{n'}(1)|A(1)|u_n(1)\rangle\langle v_{p'}(2)|v_p(2)\rangle
\end{aligned} \tag{62}$$

Now:

$$\langle v_{p'}(2)|v_p(2)\rangle = \delta_{pp'} \tag{63}$$

We can therefore write (62) in the form:

$$\langle \tilde{A}(1)\rangle = \sum_{n,n'}\left[\sum_p \langle u_n(1)v_p(2)|\rho|u_{n'}(1)v_p(2)\rangle\right]\langle u_{n'}(1)|A(1)|u_n(1)\rangle \tag{64}$$

Inside the brackets on the right-hand side of (64), we recognize the matrix element of $\rho(1)$ defined in (56). We therefore have:

$$\begin{aligned}
\langle \tilde{A}(1)\rangle &= \sum_{n,n'}\langle u_n(1)|\rho(1)|u_{n'}(1)\rangle\langle u_{n'}(1)|A(1)|u_n(1)\rangle \\
&= \sum_n \langle u_n(1)|\rho(1)A(1)|u_n(1)\rangle \\
&= \mathrm{Tr}\,\{\rho(1)A(1)\}
\end{aligned} \tag{65}$$

Let us compare this result with (31). We see that the partial trace $\rho(1)$ enables us to calculate all the mean values $\langle\tilde{A}(1)\rangle$ as if the system (1) were isolated and had $\rho(1)$ for a density operator. Making the same comment as for formula (17), we see that $\rho(1)$ also enables us to obtain the probabilities of all the results of measurements bearing on system (1) alone.

Comments:

(*i*) We saw in Complement D_{III} that it is impossible to assign a state vector to system (1) [or (2)] when the state of the global system (1) + (2) is not a product state. We now see that the density operator is a much more simple tool than the state vector. In all cases (whether the global system is in a product state or not, whether it corresponds to a pure case or to a statistical mixture), one can always, thanks to the partial trace operation, assign a density operator to subsystem (1) [or (2)]. This permits us to calculate all the physical predictions about this subsystem.

(*ii*) Even if ρ describes a pure state ($\mathrm{Tr}\,\rho^2 = 1$), this is not in general true of the density operators $\rho(1)$ and $\rho(2)$ obtained from ρ by a partial trace. It can be verified from (56) [or (59)] that $\mathrm{Tr}\,\{\rho^2(1)\}$ [or $\mathrm{Tr}\,\{\rho^2(2)\}$] is not generally equal to 1. This is another way of saying that it is not in general possible to assign a state vector to (1) [or (2)], except, of course, if the global system is in a product state.

(*iii*) If the global system is in a product state:

$$|\psi\rangle = |\varphi(1)\rangle|\chi(2)\rangle \tag{66}$$

we can verify directly that the corresponding density operator is written:

$$\rho = \sigma(1) \otimes \tau(2) \tag{67}$$

with:

$$\begin{aligned} \sigma(1) &= |\varphi(1)\rangle\langle\varphi(1)| \\ \tau(2) &= |\chi(2)\rangle\langle\chi(2)| \end{aligned} \tag{68}$$

More generally, we can envisage states of the global system for which the density operator ρ can be factored as in (67) [$\sigma(1)$ and $\tau(2)$ can correspond to statistical mixtures as well as to pure cases]. The partial trace operation then yields:

$$\begin{aligned} \mathrm{Tr}_2\,\{\sigma(1) \otimes \tau(2)\} &= \sigma(1) \\ \mathrm{Tr}_1\,\{\sigma(1) \otimes \tau(2)\} &= \tau(2) \end{aligned} \tag{69}$$

An expression such as (67) therefore represents the simple juxtaposition of a system (1), described by the density operator $\sigma(1)$, and a system (2), described by the density operator $\tau(2)$.

(*iv*) Starting with an arbitrary density operator ρ [that cannot be factored as in (67)], let us calculate $\rho(1) = \mathrm{Tr}_2\,\rho$ and $\rho(2) = \mathrm{Tr}_1\,\rho$. Then let us form the product:

$$\rho' = \rho(1) \otimes \rho(2) \tag{70}$$

Unlike the case envisaged in comment (*iii*), ρ' is in general different from ρ. When the density operator cannot be factored as in (67), there is therefore a certain "correlation" between systems (1) and (2), which is no longer contained in the operator ρ' of formula (70).

(*v*) If the evolution of the global system is described by equation (36), it is in general impossible to find a Hamiltonian operator relating to system (1) alone that would enable us to write an analogous equation for $\rho(1)$. While the definition, at any time, of $\rho(1)$ in terms of ρ is simple, the evolution of $\rho(1)$ is much more difficult to describe.

References and suggestions for further reading:

Articles by Fano (2.31) and Ter Haar (2.32). Using the density operator to study relaxation phenomena: Abragam (14.1), Chap. VIII; Slichter (14.2), Chap. 5; Sargent, Scully and Lamb (15.5), Chap. VII.

Complement F_{III}

The evolution operator

1 General properties 313
2 Case of conservative systems 315

In § D-1-b of Chapter III, we saw that the transformation of $|\psi(t_0)\rangle$ (the state vector at the initial instant t_0) into $|\psi(t)\rangle$ (the state vector at an arbitrary instant) is linear. There therefore exists a linear operator $U(t, t_0)$ such that:

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle \tag{1}$$

We intend to study here the principal properties of $U(t, t_0)$, which is, by definition, the *evolution operator* of the system.

1. General properties

Since the ket $|\psi(t_0)\rangle$ is arbitrary, it is clear from (1) that:

$$U(t_0, t_0) = \mathbb{1} \tag{2}$$

Also, substituting (1) into the Schrödinger equation, we obtain:

$$i\hbar \frac{\partial}{\partial t} U(t, t_0)|\psi(t_0)\rangle = H(t)U(t, t_0)|\psi(t_0)\rangle \tag{3}$$

from which, for the same reason as above:

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = H(t)U(t, t_0) \tag{4}$$

The first-order differential equation (4) completely defines $U(t, t_0)$, taking the initial condition (2) into account. Note, moreover, that (2) and (4) can be condensed into a single integral equation:

$$U(t, t_0) = \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^{t} H(t')U(t', t_0)\mathrm{d}t' \tag{5}$$

Now let us consider the parameter t_0, which appears in $U(t, t_0)$ as a variable t', just like t. We then write (1) in the form:

$$|\psi(t)\rangle = U(t, t')|\psi(t')\rangle \tag{6}$$

But $|\psi(t')\rangle$ can itself be obtained from a formula of the same type:

$$|\psi(t')\rangle = U(t', t'')|\psi(t'')\rangle \tag{7}$$

Substitute (7) into (6):

$$|\psi(t)\rangle = U(t,t')\,U(t',t'')|\psi(t'')\rangle \tag{8}$$

Since, moreover, $|\psi(t)\rangle = U(t,t'')|\psi(t'')\rangle$, we deduce ($|\psi(t'')\rangle$ being arbitrary):

$$U(t,t'') = U(t,t')\,U(t',t'') \tag{9}$$

It is easy to generalize this procedure and to obtain:

$$U(t_n,t_1) = U(t_n,t_{n-1})\,...\,U(t_3,t_2)\,U(t_2,t_1) \tag{10}$$

where t_1, t_2, ..., t_n are arbitrary. If we assume that $t_1 < t_2 < t_3 < ... < t_n$, formula (10) is simple to interpret: to go from t_1 to t_n, the system progresses from t_1 to t_2, then from t_2 to t_3, ..., then, finally from t_{n-1} to t_n.

Set $t'' = t$ in (9); taking (2) into consideration, we obtain:

$$\mathbb{1} = U(t,t')\,U(t',t) \tag{11}$$

or, interchanging the roles of t and t':

$$\mathbb{1} = U(t',t)\,U(t,t') \tag{12}$$

We therefore have:

$$U(t',t) = U^{-1}(t,t') \tag{13}$$

Now let us calculate the evolution operator between two instants separated by $\mathrm{d}t$. To do this, write the Schrödinger equation in the form:

$$\begin{aligned} \mathrm{d}|\psi(t)\rangle &= |\psi(t+\mathrm{d}t)\rangle - |\psi(t)\rangle \\ &= -\frac{i}{\hbar}H(t)|\psi(t)\rangle\mathrm{d}t \end{aligned} \tag{14}$$

that is:

$$|\psi(t+\mathrm{d}t)\rangle = \left[\mathbb{1} - \frac{i}{\hbar}H(t)\mathrm{d}t\right]|\psi(t)\rangle \tag{15}$$

We then obtain, using the very definition of $U(t+\mathrm{d}t,t)$:

$$U(t+\mathrm{d}t,t) = \mathbb{1} - \frac{i}{\hbar}H(t)\mathrm{d}t \tag{16}$$

$U(t+\mathrm{d}t,t)$ is called the infinitesimal evolution operator. Since $H(t)$ is Hermitian, $U(t+\mathrm{d}t,t)$ is unitary (*cf.* Complement C_{II}, § 3). It follows that $U(t,t')$ is also unitary since the interval $[t,t']$ can be divided into a very large number of infinitesimal intervals. Formula (10) then shows that $U(t,t')$ is a product of unitary operators; it is therefore a unitary operator. One can consequently write (13) in the form:

$$U^{\dagger}(t,t') = U^{-1}(t,t') = U(t',t) \tag{17}$$

It is not surprising that the transformation $U(t,t')$ is unitary, that is, that it conserves the norm of vectors on which it acts. We saw in Chapter III (*cf.* § D-1-c) that the norm of the state vector does not change over time.

2. Case of conservative systems

When the operator H does not depend on time, equation (4) can easily be integrated; taking the initial condition (2) into account, we obtain:

$$U(t,t_0) = \mathrm{e}^{-iH(t-t_0)/\hbar} \tag{18}$$

One can verify directly from this formula all the properties of the evolution operator cited in § 1.

It is very simple to go from formula (D-52) to (D-54) of Chapter III using (18). It suffices to apply the operator $U(t,t_0)$ to both sides of (D-52), noting that, since $|\varphi_{n,\tau}\rangle$ is an eigenvector of H with the eigenvalue E_n:

$$U(t,t_0)|\varphi_{n,\tau}\rangle = \mathrm{e}^{-iH(t-t_0)/\hbar}|\varphi_{n,\tau}\rangle = \mathrm{e}^{-iE_n(t-t_0)/\hbar}|\varphi_{n,\tau}\rangle \tag{19}$$

Comments:

(i) When H is time-dependent, one might be tempted to believe, by analogy with formula (18), that the evolution operator is equal to the operator $V(t,t_0)$ defined by:

$$V(t,t_0) = \mathrm{e}^{-\frac{i}{\hbar}\int_{t_0}^{t} H(t')\mathrm{d}t'} \tag{20}$$

Actually, this is not true, since the derivative of an operator of the form $\mathrm{e}^{F(t)}$ is not in general equal to $F'(t)\mathrm{e}^{F(t)}$ (*cf.* Complement $\mathrm{B_{II}}$, § 5-c):

$$i\hbar\frac{\partial}{\partial t}V(t,t_0) \neq H(t)V(t,t_0) \tag{21}$$

(ii) Let us again consider the experiments described in § E-1-b of Chapter III. As we have already indicated [comment (ii) of § E-1-b-β], it is not necessary to assume that the measurements of the various observables A, B and C are made very close together in time. When the system has had the time to evolve between two successive measurements, the variations of the state vector can easily be taken into account by using the evolution operator. If t_0, t_1 and t_2 designate respectively the instants at which the measurements of A, B and C are performed, we then replace (E-15) by:

$$\mathcal{P}_a(c) = |\langle v_c|U(t_2,t_0)|u_a\rangle|^2 \tag{22}$$

and (E-17) by:

$$\mathcal{P}_a(b,c) = |\langle v_c|U(t_2,t_1)|w_b\rangle|^2|\langle w_b|U(t_1,t_0)|u_a\rangle|^2 \tag{23}$$

We then have, using (9):

$$\begin{aligned}\langle v_c|U(t_2,t_0)|u_a\rangle &= \langle v_c|U(t_2,t_1)\,U(t_1,t_0)|u_a\rangle \\ &= \sum_b \langle v_c|U(t_2,t_1)|w_b\rangle\,\langle w_b|U(t_1,t_0)|u_a\rangle\end{aligned} \tag{24}$$

Substituting (24) into (22), we see, as in (E-21), that $\mathcal{P}_a(c)$ is not equal to $\sum_b \mathcal{P}_a(b,c)$.

References and suggestions for further reading:

The evolution operator is of fundamental importance in collision theory (see Chapter VIII) and time-dependent perturbation theory (see Chapter XIII), as well as in the study of the interactions between atoms and photons (see Chapter XX).

Complement G_{III}

The Schrödinger and Heisenberg pictures

In the formalism developed in Chapter III, it is the time-independent operators which correspond to the observables of the system (*cf.* Chap. III, § D-1-d). For example, the position, momentum and kinetic energy operators of a particle do not depend on time. The evolution of the system is entirely contained in that of the state vector $|\psi(t)\rangle$ [here written $|\psi_S(t)\rangle$, for reasons which will be evident later] and is obtained from the Schrödinger equation. This is why this approach is called the *Schrödinger picture.*

Nevertheless, we know that all the predictions of quantum mechanics (probabilities, mean values) are expressed in terms of scalar products of a bra and a ket or matrix elements of operators. Now, as we saw in Complement C_{II}, these quantities are invariant when the same unitary transformation is performed on the kets and on the operators. This transformation can be chosen so as to make the transform of the ket $|\psi_S(t)\rangle$ a time-independent ket. Of course, the transforms of the observables cited above then depend on time. We thus obtain the *Heisenberg picture.*

To avoid confusion, in this complement, we shall systematically assign an index S to the kets and operators of the Schrödinger picture and an index H to those of the Heisenberg picture. The index S can be considered to be implicit in all the other complements and chapters where only the Schrödinger picture is used.

The state vector $|\psi_S(t)\rangle$ at the instant t is expressed in terms of $|\psi_S(t_0)\rangle$ by the relation:

$$|\psi_S(t)\rangle = U(t,t_0)|\psi_S(t_0)\rangle \tag{1}$$

where $U(t,t_0)$ is the evolution operator (*cf.* Complement F_{III}). Since this operator is unitary, it is sufficient to perform the unitary transformation associated with the operator $U^\dagger(t,t_0)$ to obtain a constant transformed vector $|\psi_H\rangle$:

$$\begin{aligned}|\psi_H\rangle &= U^\dagger(t,t_0)|\psi_S(t)\rangle = U^\dagger(t,t_0)\,U(t,t_0)|\psi_S(t_0)\rangle \\ &= |\psi_S(t_0)\rangle\end{aligned} \tag{2}$$

In the Heisenberg picture, the state vector, which is constant, is therefore equal to $|\psi_S(t)\rangle$ at time t_0.

The transform $A_H(t)$ of an operator $A_S(t)$ is given by (Complement C_{II}, § 2):

$$A_H(t) = U^\dagger(t,t_0)A_S(t)\,U(t,t_0) \tag{3}$$

As we have already seen, $A_H(t)$ generally depends on time, even if A_S does not.

Nevertheless, there exists an interesting special case in which, if A_S is time-independent, the same is true of A_H: the case in which the system is conservative (H_S does not depend on time) and A_S commutes with H_S (A_S is then a constant of the motion; *cf.* Chap. III, § D-2-c). In this case, we have:

$$U(t,t_0) = \mathrm{e}^{-iH_S(t-t_0)/\hbar} \tag{4}$$

If the operator A_S commutes with H_S, it also commutes with $U(t,t_0)$ (*cf.* Complement B_{II}, § 4-c), so that:

$$A_H(t) = U^\dagger(t,t_0)\,U(t,t_0)A_S = A_S \tag{5}$$

The operators A_S and A_H are therefore simply equal in this case (in particular, $H_S = H_H$, and the indices S and H are, in reality, unnecessary for the Hamiltonian). Since they are time-independent, we see that they indeed correspond to a constant of the motion.

When $A_S(t)$ is arbitrary, let us calculate the evolution of the operator $A_H(t)$. Using relation (4) of Complement F_{III}, as well as its adjoint, we obtain:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}A_H(t) = &-\frac{1}{i\hbar}U^\dagger(t,t_0)H_S(t)A_S(t)\,U(t,t_0) + U^\dagger(t,t_0)\frac{\mathrm{d}A_S(t)}{\mathrm{d}t}U(t,t_0)\\ &+\frac{1}{i\hbar}U^\dagger(t,t_0)A_S(t)H_S(t)U(t,t_0)\end{aligned} \tag{6}$$

In the first and last terms of this expression, let us insert between A_S and H_S the product $U(t,t_0)U^\dagger(t,t_0)$, which is equal to the identity operator [formula (17) of Complement F_{III}]:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}A_H(t) = &-\frac{1}{i\hbar}U^\dagger(t,t_0)H_S(t)U(t,t_0)U^\dagger(t,t_0)A_S(t)U(t,t_0)\\ &+U^\dagger(t,t_0)\frac{\mathrm{d}A_S(t)}{\mathrm{d}t}U(t,t_0)\\ &+\frac{1}{i\hbar}U^\dagger(t,t_0)A_S(t)U(t,t_0)U^\dagger(t,t_0)H_S(t)U(t,t_0)\end{aligned} \tag{7}$$

According to definition (3), we finally obtain:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}A_H(t) = [A_H(t),\, H_H(t)] + i\hbar\left(\frac{\mathrm{d}}{\mathrm{d}t}A_S(t)\right)_H \tag{8}$$

Comments:

(*i*) Historically, the first picture was developed by Schrödinger, leading him to the equation which bears his name, and the second one, by Heisenberg, who calculated the evolution of matrices representing the various operators $A_H(t)$ (hence the name "matrix mechanics"). It was not until later that the equivalence of the two approaches was proved.

(*ii*) Using (8), one immediately obtains equation (D-27) of Chapter III as we shall now show. In the Heisenberg picture, the evolution of the mean value

$$\langle A\rangle(t) = \langle\psi_S(t)|A_S(t)|\psi_S(t)\rangle$$

can be calculated, since:

$$\langle A\rangle(t) = \langle\psi_H|A_H(t)|\psi_H\rangle \tag{9}$$

On the right-hand side of (9), only $A_H(t)$ depends on time, so (D-27) can be obtained directly by differentiation. Note, nevertheless, that equation (8) is more general than (D-27) since, instead of expressing the equality of two mean values (that is, two matrix elements of operators), it expresses the equality of two operators.

(*iii*) When the system under consideration is composed of a particle of mass m under the influence of a potential, equation (8) becomes very simple. We then have (confining ourselves to one dimension):

$$H_S(t) = \frac{{P_S}^2}{2m} + V(X_S, t) \tag{10}$$

and therefore [*cf.* formula (35) of Complement $\mathrm{C_{II}}$]:

$$H_H(t) = \frac{{P_H}^2}{2m} + V(X_H, t) \tag{11}$$

Substituting (11) into (8) and using the fact that $[X_H, P_H] = [X_S, P_S] = i\hbar$, we obtain, by an argument analogous to that of § D-1-d of Chapter III:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} X_H(t) &= \frac{1}{m} P_H(t) \\ \frac{\mathrm{d}}{\mathrm{d}t} P_H(t) &= -\frac{\partial V}{\partial X}(X_H, t) \end{aligned} \tag{12}$$

These equations generalize the Ehrenfest theorem [*cf.* Chap. III, relations (D-34) and (D-35)]. They are similar to those giving the evolution of the classical quantities x and p [*cf.* Chap. III, relations (D-36a) and (D-36b)]. An advantage of the Heisenberg picture is that it leads to equations formally similar to those of classical mechanics.

References and suggestions for further reading:

The *interaction picture*: see exercise 15 of Complement $\mathrm{L_{III}}$ as well as § A-2 of Chapter XX; Messiah (1.17), Chap. VIII, § 14; Schiff (1.18), § 24; Merzbacher (1.16), Chap. 18, § 7.

Complement H_{III}

Gauge invariance

1 Outline of the problem: scalar and vector potentials associated with an electromagnetic field; concept of a gauge . . 321

2 Gauge invariance in classical mechanics 322

2-a Newton's equations 322

2-b The Hamiltonian formalism 322

3 Gauge invariance in quantum mechanics 327

3-a Quantization rules 327

3-b Unitary transformation of the state vector; form invariance of the Schrödinger equation 328

3-c Invariance of physical predictions under a gauge transformation 331

1. Outline of the problem: scalar and vector potentials associated with an electromagnetic field; concept of a gauge

Consider an electromagnetic field, characterized by the values $\mathbf{E}(\mathbf{r};t)$ of the electric field and $\mathbf{B}(\mathbf{r};t)$ of the magnetic field at every instant and at all points in space: $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$ are not independent since they satisfy Maxwell's equations. Instead of specifying these two vector fields, it is possible to introduce a scalar potential $U(\mathbf{r};t)$ and a vector potential $\mathbf{A}(\mathbf{r};t)$ such that:

$$\begin{cases} \mathbf{E}(\mathbf{r};t) = -\boldsymbol{\nabla} U(\mathbf{r};t) - \dfrac{\partial}{\partial t}\mathbf{A}(\mathbf{r};t) \\ \mathbf{B}(\mathbf{r};t) = \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r};t) \end{cases} \tag{1}$$

It can be shown from Maxwell's equations (*cf.* Appendix III, § 4-b-α) that there always exist functions $U(\mathbf{r};t)$ and $\mathbf{A}(\mathbf{r};t)$ that allow $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$ to be expressed in the form (1). All electromagnetic fields can therefore be described by scalar and vector potentials. However, when $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$ are given, $U(\mathbf{r};t)$ and $\mathbf{A}(\mathbf{r};t)$ are not uniquely determined. It can easily be verified that if we have a set of possible values for $U(\mathbf{r};t)$ and $\mathbf{A}(\mathbf{r};t)$, we obtain other potentials $U'(\mathbf{r};t)$ and $\mathbf{A}'(\mathbf{r};t)$ describing the *same* electromagnetic field by the transformation:

$$\begin{aligned} U'(\mathbf{r};t) &= U(\mathbf{r};t) - \frac{\partial}{\partial t}\chi(\mathbf{r};t) \\ \mathbf{A}'(\mathbf{r};t) &= \mathbf{A}(\mathbf{r};t) + \boldsymbol{\nabla}\chi(\mathbf{r};t) \end{aligned} \tag{2}$$

where $\chi(\mathbf{r};t)$ is an arbitrary function of $\mathbf{r}$ and t. This can be seen by replacing $U(\mathbf{r};t)$ by $U'(\mathbf{r};t)$ and $\mathbf{A}(\mathbf{r};t)$ by $\mathbf{A}'(\mathbf{r};t)$ in (1) and verifying that $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$ remain

unchanged. Moreover, it can be shown that relations (2) give all the possible scalar and vector potentials associated with a given electromagnetic field.

When a particular set of potentials has been chosen to describe an electromagnetic field, a *choice of gauge* is said to have been made. As we just mentioned, an infinite number of different gauges can be used for the same field, characterized by $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$. When one changes from one to another, one is said to perform a *gauge transformation*.

It often happens in physics that the equations of motion of a system involve, not the fields $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$, but the potentials $U(\mathbf{r};t)$ and $\mathbf{A}(\mathbf{r};t)$. We saw an example of this in § B-5-b of Chapter III, when we wrote the Schrödinger equation for a particle of charge q in an electromagnetic field [*cf.* relation (B-48) of that chapter]. The following question can then be posed: do the physical results predicted by the theory depend only on the values of the fields $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$ at all points in space, or do they also depend on the gauge used to write the equations? In the latter case, it would obviously be necessary, in order for the theory to make sense, to specify in which gauge the equations are valid.

The aim of this complement is to answer this question. We shall see that in classical mechanics (§ 2), as in quantum mechanics (§ 3), physical results are not modified when a gauge transformation is performed. The scalar and vector potentials can then be seen to be calculation tools; actually, all that counts are the values of the electric and magnetic fields at all points in space. We shall express this result by saying that classical and quantum mechanics possess the property of *gauge invariance*.

2. Gauge invariance in classical mechanics

2-a. Newton's equations

In classical mechanics, the motion of a particle[1] of charge q and mass m placed in an electromagnetic field can be calculated from the force $\mathbf{f}$ exerted on it. This force is given by Lorentz' law:

$$\mathbf{f} = q\,[\mathbf{E}(\mathbf{r};t) + \mathbf{v} \times \mathbf{B}(\mathbf{r};t)] \tag{3}$$

where $\mathbf{v}$ is the velocity of the particle. To obtain the equations of motion which allow one to calculate the position $\mathbf{r}(t)$ of the particle at any instant t, one substitutes relation (3) into the fundamental dynamical equation (Newton's law):

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathbf{r}(t) = \mathbf{f} \tag{4}$$

In this approach, only the values of the electric and magnetic fields enter into the calculation; therefore, the problem of gauge invariance does not arise.

2-b. The Hamiltonian formalism

Instead of adopting the point of view of the preceding section, one can use other equations of motion, the Hamilton-Jacobi equations. It is not difficult to show (*cf.*

[1] For simplicity, we shall assume in this complement that the system under study is composed of a single particle. Generalization to a more complex system formed by several particles placed in an electromagnetic field presents no difficulties.

Appendix III) that the latter equations are completely equivalent to Newton's equations. However, since we used the Hamiltonian formalism in Chapter III to quantize a physical system, it is useful to study how a gauge transformation appears in this formalism. Although the scalar and vector potentials do not enter into Newton's equations, they are indispensable for writing those of Hamilton. The property of gauge invariance is therefore less obvious for this second point of view.

α. The dynamical variables of the system and their evolution

To determine the motion of a particle subjected to the Lorentz force written in (3), one can use the Lagrangian[2]:

$$\mathcal{L}(\mathbf{r}, \mathbf{v}; t) = \frac{1}{2} m\mathbf{v}^2 - q\left[U(\mathbf{r}; t) - \mathbf{v} \cdot \mathbf{A}(\mathbf{r}; t)\right] \tag{5}$$

This expression permits the calculation of the momentum $\mathbf{p}$, which is written:

$$\mathbf{p} = \boldsymbol{\nabla}_{\mathbf{v}} \mathcal{L}(\mathbf{r}, \mathbf{v}; t) = m\mathbf{v} + q\mathbf{A}(\mathbf{r}; t) \tag{6}$$

It is then possible to introduce the classical Hamiltonian:

$$\mathcal{H}(\mathbf{r}, \mathbf{p}; t) = \frac{1}{2m}\left[\mathbf{p} - q\mathbf{A}(\mathbf{r}; t)\right]^2 + qU(\mathbf{r}; t) \tag{7}$$

In the Hamiltonian formalism, the state of the particle at a given time is defined by its position $\mathbf{r}$ and its momentum $\mathbf{p}$, which we shall call the fundamental dynamical variables, and no longer by its position and its velocity, as in § 2-a above (and as in the Lagrange point of view). The momentum $\mathbf{p}$ (conjugate momentum of the position $\mathbf{r}$) must not be confused with the mechanical momentum $\boldsymbol{\pi}$:

$$\boldsymbol{\pi} = m\mathbf{v} \tag{8}$$

They are indeed different since, according to (6):

$$\boldsymbol{\pi} = \mathbf{p} - q\mathbf{A}(\mathbf{r}; t) \tag{9}$$

This relation allows us to calculate the mechanical momentum (and therefore the velocity) whenever the values of $\mathbf{r}$ and $\mathbf{p}$ are known. Similarly, all the other quantities associated with the particle (kinetic energy, angular momentum, etc...) are expressed in the Hamiltonian formalism as functions of the fundamental dynamical variables $\mathbf{r}$ and $\mathbf{p}$ (and, if necessary, of time).

The evolution of the system is governed by Hamilton's equations:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{r}(t) = \boldsymbol{\nabla}_{\mathbf{p}} \mathcal{H}[\mathbf{r}(t), \mathbf{p}(t); t] \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{p}(t) = -\boldsymbol{\nabla}_{\mathbf{v}} \mathcal{H}[\mathbf{r}(t), \mathbf{p}(t); t] \end{cases} \tag{10}$$

where $\mathcal{H}$ is the function of $\mathbf{r}$ and $\mathbf{p}$ written in (7). These equations give the values, for all times, of the fundamental dynamical variables if they are known at the initial instant.

[2] We state without proof a certain number of results of analytical mechanics which are established in Appendix III.

To write equations (10), it is necessary to choose a gauge $\mathcal{J}$, that is, a pair of potentials $\{U(\mathbf{r};t),\ \mathbf{A}(\mathbf{r};t)\}$ describing the electromagnetic field. What happens if, instead of this gauge $\mathcal{J}$, we choose another one $\mathcal{J}'$, characterized by different potentials $U'(\mathbf{r};t)$ and $\mathbf{A}'(\mathbf{r};t)$, but describing the same fields $\mathbf{E}(\mathbf{r};t)$ and $\mathbf{B}(\mathbf{r};t)$? We shall label with a prime the values of the dynamical variables associated with the motion of the particle when the gauge chosen is $\mathcal{J}'$. As we pointed out in § a, Newton's equations indicate that the position $\mathbf{r}$ and the velocity $\mathbf{v}$ take on, at every instant, values independent of the gauge. Consequently, we have:

$$\begin{cases} \mathbf{r}'(t) = \mathbf{r}(t) & \text{(11a)} \\ \boldsymbol{\pi}'(t) = \boldsymbol{\pi}(t) & \text{(11b)} \end{cases}$$

Now, from (9):

$$\begin{aligned} \boldsymbol{\pi}(t) &= \mathbf{p}(t) - q\mathbf{A}[\mathbf{r}(t);t] \\ \boldsymbol{\pi}'(t) &= \mathbf{p}'(t) - q\mathbf{A}'[\mathbf{r}(t);t] \end{aligned} \qquad (12)$$

Therefore, the values $\mathbf{p}(t)$ and $\mathbf{p}'(t)$ of the momentum in the gauges $\mathcal{J}$ and $\mathcal{J}'$ are different; they must satisfy:

$$\mathbf{p}'(t) - q\mathbf{A}'[\mathbf{r}'(t);t] = \mathbf{p}(t) - q\mathbf{A}[\mathbf{r}(t);t] \qquad (13)$$

If $\chi(\mathbf{r};t)$ is the function appearing in formulas (2) which govern the gauge transformation from $\mathcal{J}$ to $\mathcal{J}'$, the values of the fundamental dynamical variables are transformed according to the formulas:

$$\begin{cases} \mathbf{r}'(t) = \mathbf{r}(t) & \text{(14a)} \\ \mathbf{p}'(t) = \mathbf{p}(t) + q\boldsymbol{\nabla}\chi\,[\mathbf{r}(t);t] & \text{(14b)} \end{cases}$$

In the Hamiltonian formalism, the value at each instant of the dynamical variables describing a given motion depends on the gauge chosen. Moreover, such a result is not surprising since, in (7) and (10), the scalar and vector potentials appear explicitly in the equations of motion for the position and momentum.

β. "True physical quantities" and "non-physical quantities"

(*i*) Definitions

We have just seen, in relations (14) for example, that it is possible to distinguish between two types of quantities associated with the particle: those which, like $\mathbf{r}$ or $\boldsymbol{\pi}$, have identical values at all times in two different gauges, and those which, like $\mathbf{p}$, have values that depend on the arbitrarily chosen gauge. We are thus led to the following general definition:

– A *true physical quantity* associated with the system under consideration is a quantity whose value at any time does not depend (for a given motion of the system) on the gauge used to describe the electromagnetic field.

– A *non-physical quantity*, on the other hand, is a quantity whose value is modified by a gauge transformation; thus, like the scalar and vector potentials, it is seen to be a calculation tool, rather than an actually observable quantity.

The problem then posed is the following: in the Hamiltonian formalism, all the quantities associated with the system appear in the form of functions of the fundamental dynamical variables $\mathbf{r}$ and $\mathbf{p}$; how can we know whether such a function corresponds to a true physical quantity or not?

(*ii*) Characteristic relation of true physical quantities

Let us first assume that a quantity associated with the particle is described, in the gauge $\mathcal{J}$, by a function of $\mathbf{r}$ and $\mathbf{p}$ (which may depend on time) which we shall denote by $\mathcal{F}(\mathbf{r}, \mathbf{p}; t)$. If to this quantity corresponds, in another gauge $\mathcal{J}'$, the same function $\mathcal{F}(\mathbf{r}', \mathbf{p}'; t)$, the quantity is clearly not truly physical [except in the special case where the function $\mathcal{F}$ depends only on $\mathbf{r}$ and not on $\mathbf{p}$; see equations (14)]. Since the values of the momentum are different in the two gauges $\mathcal{J}$ and $\mathcal{J}'$, the same is obviously true for the values of the function $\mathcal{F}$.

To obtain the true physical quantities associated with the system, *we must therefore consider functions* $\mathcal{G}_{\mathcal{J}}(\mathbf{r}, \mathbf{p}; t)$ *whose form depends on the gauge chosen* (this is why we label these functions with an index $\mathcal{J}$). We have already seen an example of such a function: the mechanical momentum $\boldsymbol{\pi}$ is a function of $\mathbf{r}$ and $\mathbf{p}$ via the vector potential $\mathbf{A}$ [*cf.* (9)]. In this case, the function is different in the two gauges $\mathcal{J}$ and $\mathcal{J}'$; that is, it is of the form $\boldsymbol{\pi}_{\mathcal{J}}(\mathbf{r}, \mathbf{p}; t)$. The definition given in (*i*) thus implies that the function $\mathcal{G}_{\mathcal{J}}(\mathbf{r}, \mathbf{p}; t)$ describes a true physical quantity on the condition that:

$$\mathcal{G}_{\mathcal{J}}\left[\mathbf{r}(t), \mathbf{p}(t); t\right] = \mathcal{G}_{\mathcal{J}'}\left[\mathbf{r}'(t), \mathbf{p}'(t); t\right] \tag{15}$$

where $\mathbf{r}(t)$ and $\mathbf{p}(t)$ are the values taken on by the position and momentum in the gauge $\mathcal{J}$, and $\mathbf{r}'(t)$ and $\mathbf{p}'(t)$ are their values in the gauge $\mathcal{J}'$. If we substitute relations (14) into (15), we obtain:

$$\mathcal{G}_{\mathcal{J}}\left[\mathbf{r}(t), \mathbf{p}(t); t\right] = \mathcal{G}_{\mathcal{J}'}\left[\mathbf{r}(t), \mathbf{p}(t) + q\boldsymbol{\nabla}\chi(\mathbf{r}(t); t); t\right] \tag{16}$$

This relation must be satisfied at every instant t and for all possible motions of the system. Since, when t is fixed, the values of the position and the momentum can be chosen independently, both sides of (16) must in fact be the same function of $\mathbf{r}$ and $\mathbf{p}$, which is written:

$$\mathcal{G}_{\mathcal{J}}\left[\mathbf{r}, \mathbf{p}; t\right] = \mathcal{G}_{\mathcal{J}'}\left[\mathbf{r}, \mathbf{p} + q\boldsymbol{\nabla}\chi(\mathbf{r}; t); t\right] \tag{17}$$

This relation is characteristic of the functions $\mathcal{G}_{\mathcal{J}}[\mathbf{r}, \mathbf{p}; t]$ associated with true physical quantities. Therefore, *if one considers the function* $\mathcal{G}_{\mathcal{J}'}[\mathbf{r}, \mathbf{p}; t]$ *for the gauge* $\mathcal{J}'$*, and if one replaces* $\mathbf{p}$ *by* $\mathbf{p} + q\boldsymbol{\nabla}\chi(\mathbf{r}; t)$ [where $\chi(\mathbf{r}; t)$ defines, according to (2), the gauge transformation from $\mathcal{J}$ to $\mathcal{J}'$], *one obtains a new function of* $\mathbf{r}$ *and* $\mathbf{p}$ *which must be identical to* $\mathcal{G}_{\mathcal{J}}[\mathbf{r}, \mathbf{p}; t]$. If this is not the case, the function considered corresponds to a quantity which is not truly physical.

(*iii*) Examples

Let us give some examples of functions $\mathcal{G}_{\mathcal{J}}[\mathbf{r}, \mathbf{p}; t]$ which describe true physical quantities. We have already encountered two: those corresponding to the position and to the mechanical momentum; the first is simply equal to $\mathbf{r}$ and the second to:

$$\boldsymbol{\pi}_{\mathcal{J}}(\mathbf{r}, \mathbf{p}; t) = \mathbf{p} - q\mathbf{A}(\mathbf{r}; t) \tag{18}$$

Since relations (11) express the fact that $\mathbf{r}$ and $\boldsymbol{\pi}$ are true physical quantities, we know *a priori* that relation (17) is satisfied by the corresponding functions.

However, let us verify this directly in order to familiarize ourselves with the use of this relation. As regards $\mathbf{r}$, we are dealing with a function that does not depend on $\mathbf{p}$ and whose form does not depend on the gauge[3]; this immediately implies (17). As regards $\boldsymbol{\pi}$, relation (18) yields:

$$\boldsymbol{\pi}_{\mathcal{J}}(\mathbf{r},\mathbf{p};t) = \mathbf{p} - q\mathbf{A}'(\mathbf{r};t) \tag{19}$$

Replace in this function $\mathbf{p}$ by $\mathbf{p} + q\boldsymbol{\nabla}\chi(\mathbf{r};t)$; we obtain the function:

$$\mathbf{p} + q\boldsymbol{\nabla}\chi(\mathbf{r};t) - q\mathbf{A}'(\mathbf{r};t) = \mathbf{p} - q\mathbf{A}(\mathbf{r};t) \tag{20}$$

which is none other than $\boldsymbol{\pi}_{\mathcal{J}}(\mathbf{r},\mathbf{p};t)$; relation (17) is therefore satisfied.

Other true physical quantities are the kinetic energy:

$$\gamma_{\mathcal{J}}(\mathbf{r},\mathbf{p};t) = \frac{1}{2m}[\mathbf{p} - q\mathbf{A}(\mathbf{r};t)]^2 \tag{21}$$

and the moment, with respect to the origin, of the mechanical momentum:

$$\boldsymbol{\lambda}_{\mathcal{J}}(\mathbf{r},\mathbf{p};t) = \mathbf{r} \times [\mathbf{p} - q\mathbf{A}(\mathbf{r};t)]^2 \tag{22}$$

In general, we see that whenever a function of $\mathbf{r}$ and $\mathbf{p}$ has the form:

$$\mathcal{G}_{\mathcal{J}}(\mathbf{r},\mathbf{p};t) = F[\mathbf{r},\mathbf{p} - q\mathbf{A}(\mathbf{r};t)] \tag{23}$$

(where F is a function whose form is independent of the gauge $\mathcal{J}$ chosen), we obtain a true *physical quantity*[4]. This result makes sense since (23) really expresses the fact that the values taken on by the quantity considered are obtained from those of $\mathbf{r}$ and $\boldsymbol{\pi}$, which we know to be gauge-invariant.

Let us also give some examples of functions describing quantities that are *not* true physical quantities. In addition to the momentum $\mathbf{p}$, we can cite the function:

$$\mathcal{C}(\mathbf{p}) = \frac{\mathbf{p}^2}{2m} \tag{24}$$

which must not be confused with the kinetic energy written in (21), and, in general, any function of $\mathbf{p}$ alone (and, possibly, of the time). Similarly, the angular momentum:

$$\mathcal{L}(\mathbf{r},\mathbf{p}) = \mathbf{r} \times \mathbf{p} \tag{25}$$

cannot be considered to be a true physical quantity. Finally, let us cite the classical Hamiltonian, which, according to (7), is the sum of the kinetic energy $\gamma_{\mathcal{J}}(\mathbf{r},\mathbf{p};t)$, which is a true physical quantity, and the potential energy qU. Now, the latter [which should rigorously be written in the form of a gauge-dependent function $U_{\mathcal{J}}(\mathbf{r};t)$] is not a true physical quantity since, at every point in space, its value changes when the gauge is changed.

[3] It is not difficult to verify that, in general, any function $\mathcal{G}(\mathbf{r},t)$ that depends only on $\mathbf{r}$ (and, possibly, on the time), and whose form is the same in any gauge $\mathcal{J}$ chosen, describes a true physical quantity.

[4] One could also construct functions associated with true physical quantities in which the potentials are involved in a more complex way than in (23) (for example, the scalar product of the particle velocity and the electric field at the position of the particle).

3. Gauge invariance in quantum mechanics

In Chapter III, we introduced the postulates of quantum mechanics by starting from the Hamiltonian formulation of classical mechanics. We are thus led to ask if the problem of gauge invariance, easily resolved in classical mechanics because of the existence of Newton's equations, is more complex in the framework of quantum mechanics. The following question then arises: are the postulates stated in Chapter III valid for any arbitrarily chosen gauge $\mathcal{J}$ or only for a particular gauge?

In answering this question, we shall be guided by the results obtained in the preceding paragraph. Following the same type of reasoning, we shall see that there exists a close analogy between the consequences of a gauge transformation in the classical Hamiltonian formalism and in the quantum mechanical formalism. We shall thus establish the gauge invariance of quantum mechanics.

To do this, we shall begin (§ 3-a) by examining the results obtained when the quantization rules are applied in the same way in two different gauges. We shall then see (§ 3-b) that, like in classical mechanics, where the values of the dynamical variables generally change when the gauge is changed, *a given physical system must be characterized by a mathematical state vector* $|\psi\rangle$ *that depends on the gauge.* To pass from a state vector corresponding to one gauge $\mathcal{J}$ to that of another gauge $\mathcal{J}'$, we use a *unitary transformation.* The form of the Schrödinger equation, however, always remains the same (as do Hamilton's equations in classical mechanics). Finally, we shall examine the behavior, under a gauge transformation, of the observables associated with the system (§ 3-c). We shall see that the simultaneous modification of the state vector and the observables is such that *the physical content of quantum mechanics does not depend on the gauge chosen.* Moreover, we shall demonstrate this by showing that the density and the probability current values are gauge invariant.

3-a. Quantization rules

The state space of a (spinless) particle is always $\mathcal{E}_{\mathbf{r}}$. However, we are clearly led by the results of § 2 above to expect that the operator associated with a given quantity may be different in two different gauges. We shall therefore label these operators with an index $\mathcal{J}$.

The quantization rules associate, with the position $\mathbf{r}$ and the momentum $\mathbf{p}$ of the particle, operators $\mathbf{R}$ and $\mathbf{P}$ acting in $\mathcal{E}_{\boldsymbol{r}}$ such that:

$$[X, P_x] = [Y, P_y] = [Z, P_z] = i\hbar \tag{26}$$

(where all the other commutators between components of $\mathbf{R}$ and $\mathbf{P}$ are zero). In the $\{|\mathbf{r}\rangle\}$ representation, the operator $\mathbf{R}$ acts like a multiplication by $\mathbf{r}$, and $\mathbf{P}$ like the differential operator $\frac{\hbar}{i}\boldsymbol{\nabla}$. These rules are the same in all gauges. We can therefore write:

$$\begin{cases} \mathbf{R}_{\mathcal{J}'} = \mathbf{R}_{\mathcal{J}} & \text{(27a)} \\ \mathbf{P}_{\mathcal{J}'} = \mathbf{P}_{\mathcal{J}} & \text{(27b)} \end{cases}$$

In fact, these equations enable us to omit the index $\mathcal{J}$ for the observables $\mathbf{R}$ and $\mathbf{P}$, and we shall henceforth do so.

The quantization of all other quantities associated with the particle is obtained as follows: in a given gauge, take the function of $\mathbf{r}$ and $\boldsymbol{p}$ giving the classical quantity

considered and (after having symmetrized, if necessary) replace $\mathbf{r}$ by the operator $\mathbf{R}$ and $\mathbf{p}$ by $\mathbf{P}$. We thus obtain the operator which, in the gauge chosen, describes this quantity. Consider some examples:

– The angular momentum operator, obtained from $\mathbf{r} \times \mathbf{p}$, is the same in all gauges:

$$\mathbf{L}_{\mathcal{J}'} = \mathbf{L}_{\mathcal{J}} \tag{28}$$

– The operator associated with the mechanical momentum, on the other hand, depends on the gauge chosen. In the gauge $\mathcal{J}$, it is given by:

$$\mathbf{\Pi}_{\mathcal{J}} = \mathbf{P} - q\mathbf{A}(\mathbf{R};t) \tag{29}$$

If the gauge is changed, it becomes:

$$\mathbf{\Pi}_{\mathcal{J}'} = \mathbf{P} - q\mathbf{A}'(\mathbf{R};t) \tag{30}$$

whose action in $\mathcal{E}_{\mathbf{r}}$ is different from that of $\mathbf{\Pi}_{\mathcal{J}}$:

$$\mathbf{\Pi}_{\mathcal{J}'} = \mathbf{\Pi}_{\mathcal{J}} - q\boldsymbol{\nabla}\chi(\mathbf{R};t) \tag{31}$$

– Similarly, the operator[5]:

$$\mathbf{\Lambda}_{\mathcal{J}} = \mathbf{R} \times \mathbf{\Pi}_{\mathcal{J}} = \mathbf{R} \times [\mathbf{P} - q\mathbf{A}(\mathbf{R};t)] \tag{32}$$

which describes the moment of the mechanical momentum, explicitly involves the vector potential chosen.

– Finally, the Hamiltonian operator is obtained from formula (7):

$$H_{\mathcal{J}} = \frac{1}{2m}\left[\mathbf{P} - q\mathbf{A}(\mathbf{R};t)\right]^2 + qU(\mathbf{R};t) \tag{33}$$

It is obvious that in another gauge, it becomes a different operator, since:

$$H_{\mathcal{J}'} = \frac{1}{2m}\left[\mathbf{P} - q\mathbf{A}'(\mathbf{R};t)\right]^2 + qU'(\mathbf{R};t) \neq H_{\mathcal{J}} \tag{34}$$

3-b. Unitary transformation of the state vector; form invariance of the Schrödinger equation

α. The unitary operator $T_\chi(t)$

In classical mechanics, we denoted by $\{\mathbf{r}(t), \mathbf{p}(t)\}$ and $\{\mathbf{r}'(t), \mathbf{p}'(t)\}$ the values of the fundamental dynamical variables characterizing the state of the particle in two different gauges $\mathcal{J}$ and $\mathcal{J}'$. In quantum mechanics, we shall therefore denote by $|\psi(t)\rangle$ and $|\psi'(t)\rangle$ the state vectors relative to these two gauges, and the analogue of relations (14) is thus given by relations between mean values:

$$\langle\psi'(t)|\mathbf{R}_{\mathcal{J}'}|\psi'(t)\rangle = \langle\psi(t)|\mathbf{R}_{\mathcal{J}}|\psi(t)\rangle \tag{35a}$$

$$\langle\psi'(t)|\mathbf{P}_{\mathcal{J}'}|\psi'(t)\rangle = \langle\psi(t)|\mathbf{P}_{\mathcal{J}} + q\boldsymbol{\nabla}\chi(\mathbf{R};t)|\psi(t)\rangle \tag{35b}$$

[5]It can be verified, by using the commutation relations of $\mathbf{R}$ and $\mathbf{\Pi}_{\mathcal{J}}$, that it is not necessary to symmetrize expression (32).

Using (27), we immediately see that this is possible only if $|\psi(t)\rangle$ and $|\psi'(t)\rangle$ are two different kets. We shall therefore seek a unitary transformation $T_\chi(t)$ that enables us to go from $|\psi(t)\rangle$ to $|\psi'(t)\rangle$:

$$|\psi'(t)\rangle = T_\chi(t)|\psi(t)\rangle \tag{36a}$$

$$T_\chi^\dagger(t)T_\chi(t) = T_\chi(t)T_\chi^\dagger(t) = \mathbb{1} \tag{36b}$$

Taking (27) into account, we see that equations (35) are satisfied for any $|\psi(t)\rangle$ on condition that:

$$\begin{cases} T_\chi^\dagger(t)\mathbf{R}\,T_\chi(t) = \mathbf{R} & \text{(37a)} \\ T_\chi^\dagger(t)\mathbf{P}\,T_\chi(t) = \mathbf{P} + q\boldsymbol{\nabla}\chi(\mathbf{R};t) & \text{(37b)} \end{cases}$$

Multiplying (37a) on the left by $T_\chi(t)$, we obtain:

$$\mathbf{R}\,T_\chi(t) = T_\chi(t)\mathbf{R} \tag{38}$$

The desired unitary operator commutes with the three components of $\mathbf{R}$; it can therefore be written in the form:

$$T_\chi(t) = \mathrm{e}^{iF(\mathbf{R};t)} \tag{39}$$

where $F(\mathbf{R};t)$ is a Hermitian operator. Relation (48) of Complement $\mathrm{B_{II}}$ then allows us to write:

$$[\mathbf{P}, T_\chi(t)] = \hbar\boldsymbol{\nabla}\{F(\mathbf{R};t)\}T_\chi(t) \tag{40}$$

If we multiply this equation on the left by $T_\chi^\dagger(t)$ and substitute it into (37b), we easily obtain the relation:

$$\hbar\boldsymbol{\nabla}\{F(\mathbf{R};t)\} = q\boldsymbol{\nabla}\chi(\mathbf{R};t) \tag{41}$$

which is satisfied when:

$$F(\mathbf{R};t) = F_0(t) + \frac{q}{\hbar}\chi(\mathbf{R};t) \tag{42}$$

Omitting the coefficient $F_0(t)$, which corresponds, for the state vector $|\psi(t)\rangle$, to a global phase factor of no physical consequence, we obtain the operator $T_\chi(t)$:

$$T_\chi(t) = \mathrm{e}^{i\frac{q}{\hbar}\chi(\mathbf{R};t)} \tag{43}$$

If, in (36a), $T_\chi(t)$ is this operator, relations (35) are automatically satisfied.

Comments:

(*i*) In the $\{|\mathbf{r}\rangle\}$ representation, relations (36a) and (43) imply that the wave functions $\psi(\mathbf{r},t) = \langle\mathbf{r}|\psi(t)\rangle$ and $\psi'(\mathbf{r},t) = \langle\mathbf{r}|\psi'(t)\rangle$ are related by:

$$\psi'(\mathbf{r},t) = \mathrm{e}^{i\frac{q}{\hbar}\chi(\mathbf{r},t)}\psi(\mathbf{r},t) \tag{44}$$

For the wave function, the gauge transformation corresponds to a phase change which varies from one point to another, and is not, therefore, a global phase factor. The gauge invariance of physical predictions obtained by using the wave functions ψ or ψ', is therefore not obvious *a priori*.

(*ii*) If the system under study is composed of several particles having positions $\mathbf{r}_1$, $\mathbf{r}_2$, ... and charges q_1, q_2, ..., (43) must be replaced by:

$$\begin{aligned} T_\chi(t) &= T_\chi^{(1)}(t)\, T_\chi^2(t)... \\ &= \mathrm{e}^{\frac{i}{\hbar}[q_1\chi(\mathbf{R}_1,t)+q_2\chi(\mathbf{R}_2,t)+...]} \end{aligned} \tag{45}$$

β. Time evolution of the state vector

Now let us show that if the evolution of the ket $|\psi(t)\rangle$ obeys, in the gauge $\mathcal{J}$, to the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H_{\mathcal{J}}(t)|\psi(t)\rangle \tag{46}$$

the state vector $|\psi'(t)\rangle$ given by (36) satisfies an equation of the same form in the gauge $\mathcal{J}'$:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi'(t)\rangle = H_{\mathcal{J}'}(t)|\psi'(t)\rangle \tag{47}$$

where $H_{\mathcal{J}'}(t)$ is given by (34).

To do this, let us calculate the left-hand side of (47); it is written:

$$\begin{aligned} i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi'(t)\rangle &= i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\{T_\chi(t)|\psi(t)\rangle\} \\ &= i\hbar\left\{\frac{\mathrm{d}}{\mathrm{d}t}T_\chi(t)\right\}|\psi(t)\rangle + i\hbar T_\chi(t)\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle \end{aligned} \tag{48}$$

that is, according to (43) and (46) [6]:

$$\begin{aligned} i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi'(t)\rangle &= -q\left\{\frac{\partial}{\partial t}\chi(\mathbf{R};t)\right\}T_\chi(t)|\psi(t)\rangle + T_\chi(t)H_{\mathcal{J}}(t)|\psi(t)\rangle \\ &= \left\{-q\frac{\partial}{\partial t}\chi(\mathbf{R};t) + \tilde{H}_{\mathcal{J}}(t)\right\}|\psi'(t)\rangle \end{aligned} \tag{49}$$

where $\tilde{H}_{\mathcal{J}}(t)$ designates the transform of $H_{\mathcal{J}}(t)$ by the unitary operator $T_\chi(t)$:

$$\tilde{H}_{\mathcal{J}}(t) = T_\chi(t)H_{\mathcal{J}}(t)T_\chi^\dagger(t) \tag{50}$$

Equation (47) will therefore be satisfied if:

$$H_{\mathcal{J}'}(t) = \tilde{H}_{\mathcal{J}}(t) - q\frac{\partial}{\partial t}\chi(\mathbf{R};t) \tag{51}$$

Now $\tilde{H}_{\mathcal{J}}(t)$ is given by:

$$\tilde{H}_{\mathcal{J}}(t) = \frac{1}{2m}\left[\tilde{\mathbf{P}} - q\mathbf{A}(\tilde{\mathbf{R}};t)\right]^2 + qU(\tilde{\mathbf{R}};t) \tag{52}$$

[6] The function χ depends on $\mathbf{R}$ and not on $\mathbf{P}$; consequently $\chi(\mathbf{R},t)$ commutes with $\dfrac{\partial}{\partial t}\chi(\mathbf{R};t)$. This is why $T_\chi(t)$ can be differentiated as if $\chi(\mathbf{R},t)$ were an ordinary function of the time and not an operator (*cf.* Complement B_{II}, comment of § 5-c).

where $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{P}}$ designate the transforms of $\mathbf{R}$ and $\mathbf{P}$ by the unitary operator $T_\chi(t)$. According to (37):

$$\begin{cases} \tilde{\mathbf{R}} = T_\chi(t)\mathbf{R}\,T_\chi^\dagger(t) = \mathbf{R} & \text{(53a)} \\ \tilde{\mathbf{P}} = T_\chi(t)\mathbf{P}\,T_\chi^\dagger(t) = \mathbf{P} - q\boldsymbol{\nabla}\chi(\mathbf{R};t) & \text{(53b)} \end{cases}$$

These relations, substituted into (52), indicate that:

$$\tilde{H}_{\mathcal{J}}(t) = \frac{1}{2m}\left[\mathbf{P} - q\mathbf{A}(\mathbf{R};t) - q\boldsymbol{\nabla}\chi(\mathbf{R};t)\right]^2 + qU(\mathbf{R};t) \tag{54}$$

Using relations (2) to replace the potentials relative to the gauge $\mathcal{J}$ by those relative to $\mathcal{J}'$, we then obtain, using (34), relation (51). Therefore, the Schrödinger equation can be written in the same way in any gauge chosen.

3-c. Invariance of physical predictions under a gauge transformation

α. Behavior of the observables

Under the effect of the unitary transformation $T_\chi(t)$, any observable K is transformed into $\tilde{K}$, with:

$$\tilde{K} = T_\chi(t)\,K\,T_\chi^\dagger(t) \tag{55}$$

We have already seen, in (53), that while $\tilde{\mathbf{R}}$ is simply equal to $\mathbf{R}$, $\tilde{\mathbf{P}}$ is not equal to $\mathbf{P}$. Similarly, $\tilde{\boldsymbol{\Pi}}_{\mathcal{J}}$ is different from $\boldsymbol{\Pi}_{\boldsymbol{\mathcal{J}}}$ since:

$$\begin{aligned} \tilde{\boldsymbol{\Pi}}_{\mathcal{J}} &= \tilde{\mathbf{P}} - q\mathbf{A}(\tilde{\mathbf{R}};t) \\ &= \mathbf{P} - q\boldsymbol{\nabla}\chi(\mathbf{R};t) - q\mathbf{A}(\mathbf{R};t) \\ &= \boldsymbol{\Pi}_{\mathcal{J}} - q\boldsymbol{\nabla}\chi(\mathbf{R};t) \end{aligned} \tag{56}$$

Taking (27a) and (31) into account, we see that relations (53a) and (56) imply that the observables $\mathbf{R}$ and $\boldsymbol{\Pi}_{\mathcal{J}}$, associated with true physical quantities (position and mechanical momentum) are such that:

$$\begin{cases} \tilde{\mathbf{R}}_{\mathcal{J}} = \mathbf{R}_{\mathcal{J}'} \\ \tilde{\boldsymbol{\Pi}}_{\mathcal{J}} = \boldsymbol{\Pi}_{\mathcal{J}'} \end{cases} \tag{57}$$

On the other hand, the momentum $\mathbf{P}$ (which is not a true physical quantity) does not satisfy an analogous relation, since, from (27b) and (53b):

$$\tilde{\mathbf{P}}_{\mathcal{J}} \neq \mathbf{P}_{\mathcal{J}'} \tag{58}$$

We shall see that this result is a general one: *in quantum mechanics, for every true physical quantity, there is an operator $G_{\mathcal{J}}(t)$ that satisfies*:

$$\tilde{G}_{\mathcal{J}} = G_{\mathcal{J}'}(t) \tag{59}$$

This relation is the quantum mechanical analogue of the classical relation (16). It shows that, except for the special case of $\mathbf{R}$ or a function of $\mathbf{R}$ alone, *the operator corresponding*

to a true physical quantity depends on the gauge $\mathcal{J}$. We have already seen examples of this in (29) and (32).

To prove (59), one need only apply the quantization rules stated in Chapter III to a function $\mathcal{G}_{\mathcal{J}}(\mathbf{r}, \mathbf{p}; t)$ and use relation (17), the characteristic relation for true physical classical quantities. We therefore replace $\mathbf{r}$ and $\mathbf{p}$ by the operators $\mathbf{R}$ and $\mathbf{P}$ and obtain (if necessary, after a symmetrization with respect to these operators) the operator $G_{\mathcal{J}}(t)$. If the form of the function $\mathcal{G}_{\mathcal{J}}$ depends on the gauge chosen, the operator $G_{\mathcal{J}}(t)$ also depends on $\mathcal{J}$. When the quantity associated with $G_{\mathcal{J}}$ is a true physical quantity, we have, according to (17):

$$\mathcal{G}_{\mathcal{J}}[\mathbf{R}, \mathbf{P}; t] = \mathcal{G}_{\mathcal{J}'}[\mathbf{R}, \mathbf{P} + q\boldsymbol{\nabla}\chi(\mathbf{R}; t); t] \tag{60}$$

Applying the unitary transformation $T_\chi(t)$ to this relation, we obtain:

$$\begin{aligned}\tilde{\mathcal{G}}_{\mathcal{J}}[\mathbf{R}, \mathbf{P}; t] &= \tilde{\mathcal{G}}_{\mathcal{J}'}[\mathbf{R}, \mathbf{P} + q\boldsymbol{\nabla}\chi(\mathbf{R}; t); t] \\ &= \mathcal{G}_{\mathcal{J}'}[\tilde{\mathbf{R}}, \tilde{\mathbf{P}} + q\boldsymbol{\nabla}\chi(\tilde{\mathbf{R}}; t); t]\end{aligned} \tag{61}$$

That is, taking (53) into account:

$$\tilde{\mathcal{G}}_{\mathcal{J}}[\mathbf{R}, \mathbf{P}; t] = \mathcal{G}_{\mathcal{J}'}[\mathbf{R}, \mathbf{P}; t] \tag{62}$$

After symmetrizing, if necessary, both sides of this relation, we indeed obtain (59).

Let us give some examples of true physical observables. In addition to $\mathbf{R}$ and $\boldsymbol{\Pi}_{\mathcal{J}}$, we can cite the moment $\boldsymbol{\Lambda}_{\mathcal{J}}$ of the mechanical momentum [*cf.* (32)], or the kinetic energy:

$$\Gamma_{\mathcal{J}} = \frac{\boldsymbol{\Pi}_{\mathcal{J}}^2}{2m} = \frac{1}{2m}[\mathbf{P} - q\mathbf{A}(\mathbf{R}; t)]^2 \tag{63}$$

On the other hand, $\mathbf{P}$ and $\mathbf{L}$ are not true physical quantities; neither is the Hamiltonian, since relation (51) implies in general that:

$$\tilde{H}_{\mathcal{J}}(t) \neq H_{\mathcal{J}'}(t) \tag{64}$$

Comment:

In classical mechanics, it is well known that the total energy of a particle moving in a time-independent electromagnetic field is a constant of the motion. It is indeed possible in this case to limit oneself to potentials which are also time-independent. We see from (51) that one then has:

$$\tilde{H}_{\mathcal{J}} = H_{\mathcal{J}'} \tag{65}$$

In this particular case, $H_{\mathcal{J}}$ is indeed a true physical observable which can therefore be interpreted to be the total energy of the particle.

β. Probability of the various possible results of a measurement bearing on a true physical quantity

Assume that at time t we want to measure a true physical quantity. In the gauge $\mathcal{J}$, the state of the system is described at this instant by the ket[7] $|\psi\rangle$, and the physical

[7] We do not indicate the time dependence because all the quantities must be evaluated at the time t when we want to perform the measurement.

quantity, by the observable $G_{\mathcal{J}}$. Let $|\varphi_n\rangle$ be an eigenvector of $G_{\mathcal{J}}$, with the eigenvalue g_n (assumed, for simplicity, to be non-degenerate):

$$G_{\mathcal{J}}|\varphi_n\rangle = g_n|\varphi_n\rangle \tag{66}$$

As calculated in the gauge $\mathcal{J}$ from the postulates of quantum mechanics, the probability of obtaining g_n in the measurement envisaged is equal to:

$$\mathcal{P}_n = |\langle\varphi_n|\psi\rangle|^2 \tag{67}$$

What happens to this prediction when the gauge is changed? According to (59), the operator $G_{\mathcal{J}'}$ associated with the quantity under consideration in the new gauge $\mathcal{J}'$ can have the ket:

$$|\varphi'_n\rangle = T_\chi|\varphi_n\rangle \tag{68}$$

as an eigenvector, with the same eigenvalue g_n as in (66). That is:

$$\begin{aligned} G_{\mathcal{J}'}|\varphi'_n\rangle &= T_\chi\, G_{\mathcal{J}}\, T_\chi^\dagger\, T_\chi|\varphi_n\rangle \\ &= T_\chi\, g_n|\varphi_n\rangle = g_n|\varphi'_n\rangle \end{aligned} \tag{69}$$

In the gauge $\mathcal{J}'$, g_n therefore still appears as a possible measurement result. Moreover, calculation of the corresponding probability yields the same value as in the gauge $\mathcal{J}$, since, according to (36a) and (68):

$$\langle\varphi'_n|\psi'\rangle = \langle\varphi_n|T_\chi^\dagger\, T_\chi|\psi\rangle = \langle\varphi_n|\psi\rangle \tag{70}$$

We have thus verified that the postulates of quantum mechanics lead to *gauge-invariant physical predictions*: the possible results of any measurement and the associated probabilities are invariant under a gauge transformation.

γ. *Probability density and current*

Let us calculate, from formulas (D-9) and (D-20) of Chapter III, the probability density $\rho(\mathbf{r},t)$ and probability current $\mathbf{J}(\mathbf{r},t)$ in two different gauges $\mathcal{J}$ and $\mathcal{J}'$. For the first gauge, we have:

$$\rho(\mathbf{r},t) = |\psi(\mathbf{r},t)|^2 \tag{71}$$

and:

$$\mathbf{J}(\mathbf{r},t) = \frac{1}{m}\mathrm{Re}\left\{\psi^*(\mathbf{r},t)\left[\frac{\hbar}{i}\boldsymbol{\nabla} - q\mathbf{A}(\mathbf{r};t)\right]\psi(\mathbf{r},t)\right\} \tag{72}$$

Relation (44) immediately shows that:

$$\rho'(\mathbf{r},t) = |\psi'(\mathbf{r},t)|^2 = \rho(\mathbf{r},t) \tag{73}$$

Moreover, it also implies that:

$$\begin{aligned} \mathbf{J}'(\mathbf{r},t) &= \frac{1}{m}\mathrm{Re}\left\{\mathrm{e}^{-i\frac{q}{\hbar}\chi(\mathbf{r};t)}\psi^*(\mathbf{r},t)\left[\frac{\hbar}{i}\boldsymbol{\nabla} - q\mathbf{A}'(\mathbf{r};t)\right]\mathrm{e}^{i\frac{q}{\hbar}\chi(\mathbf{r};t)}\psi(\mathbf{r},t)\right\} \\ &= \frac{1}{m}\mathrm{Re}\left\{\psi^*(\mathbf{r},t)\left[\frac{\hbar}{i}\boldsymbol{\nabla} - q\mathbf{A}'(\mathbf{r};t) + q\boldsymbol{\nabla}\chi(\mathbf{r};t)\right]\psi(\mathbf{r},t)\right\} \end{aligned} \tag{74}$$

that is, taking (2) into account:

$$\mathbf{J}'(\mathbf{r},t) = \mathbf{J}(\mathbf{r},t) \tag{75}$$

The probability density and current are therefore invariant under a gauge transformation. This result could have been foreseen, moreover, from the conclusions of § 3-c-γ above, since [*cf.* relation (D-19) of Chapter III] $\rho(\mathbf{r},t)$ and $\mathbf{J}(\mathbf{r},t)$ can be considered to be mean values of the operators $|\mathbf{r}\rangle\langle\mathbf{r}|$ and:

$$\mathbf{K}_{\mathcal{J}}(\mathbf{r}) = \frac{1}{2m}\left\{|\mathbf{r}\rangle\langle\mathbf{r}|\mathbf{\Pi}_{\mathcal{J}} + \mathbf{\Pi}_{\mathcal{J}}|\mathbf{r}\rangle\langle\mathbf{r}|\right\} \tag{76}$$

It is not difficult to show that these two operators satisfy relation (59). They therefore describe true physical quantities whose mean values are gauge-invariant.

References and suggestions for further reading:

Messiah (1.17), Chap. XXI, §§ 20 to 22; Sakurai (2.7), § 8-1.

Gauge invariance, extended to other domains, plays an important role in particle physics; see, for example, the article by Abers and Lee (16.35).

Complement J_{III}

Propagator for the Schrödinger equation

1 Introduction 335

2 Existence and properties of a propagator $K(2,1)$ 336

2-a Existence of a propagator 336

2-b Physical interpretation of $K(2,1)$ 337

2-c Expression for $K(2,1)$ in terms of the eigenstates of H 338

2-d Equation satisfied by $K(2,1)$ 339

3 Lagrangian formulation of quantum mechanics 339

3-a Concept of a space-time path 339

3-b Decomposition of $K(2,1)$ into a sum of partial amplitudes . . 340

3-c Feynman's postulates 341

3-d The classical limit and Hamilton's principle 341

1. Introduction

Consider a particle described by the wave function $\psi(\mathbf{r},t)$. The Schrödinger equation enables us to calculate $\dfrac{\partial}{\partial t}\psi(\mathbf{r},t)$, that is, the rate of variation of $\psi(\mathbf{r},t)$ with respect to t. It therefore gives the time evolution of the wave function $\psi(\mathbf{r},t)$, using a differential point of view. One might wonder if it is possible to adopt a more global (but equivalent) point of view that would allow us to determine directly the value $\psi(\mathbf{r}_0,t)$ taken on by the wave function at a given point $\mathbf{r}_0$ and a given time t from knowledge of the whole wave function $\psi(\mathbf{r},t')$ at a previous time t' (which is not necessarily infinitesimally close to t).

To consider this possibility, we can take our inspiration from another domain of physics, electromagnetism, where both points of view are possible. *Maxwell's equations* (the differential point of view) give the rates of variation of the various components of the electric and magnetic fields. *Huygens' principle* (the global point of view) permits the direct calculation, when a monochromatic field is known on a surface Σ, of the field at any point M: one sums the fields radiated at the point M by *fictional secondary sources* N_1, N_2, N_3, ... situated on the surface Σ and whose amplitude and phase are determined by the value of the field at N_1, N_2, N_3, ... (Fig. 1).

We intend to show in this complement that there exists an analogue of Huygens' principle in quantum mechanics. More precisely, we can write, for $t_2 > t_1$:

$$\psi(\mathbf{r}_2,t_2) = \int \mathrm{d}^3 r_1 \, K(\mathbf{r}_2,t_2;\mathbf{r}_1,t_1)\ \psi(\mathbf{r}_1,t_1) \tag{1}$$

a formula whose physical interpretation is the following: the probability amplitude of finding the particle at $\mathbf{r}_2$ at the instant t_2 is obtained by summing all the amplitudes

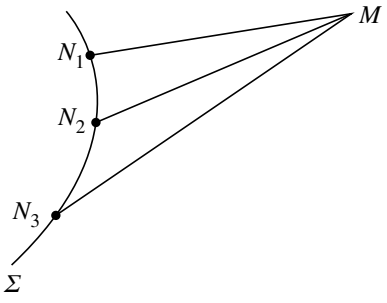


Figure 1: In a diffraction experiment, Huygens' principle permits the calculation of the electric field at the point M, as a sum of fields radiated by secondary sources N_1, N_2, N_3, ... situated on a surface Σ.

"radiated" by the "secondary sources" $(\mathbf{r}_1, t_1)$, $(\mathbf{r}'_1, t_1)$... situated in space-time on the surface $t = t_1$, each of these sources contributing to a degree proportional to $\psi(\mathbf{r}_1, t_1)$, $\psi(\mathbf{r}'_1, t_1)$, ... (Fig. 2). We shall prove the preceding formula, calculate K, called the *propagator* for the Schrödinger equation, and study its properties. We shall then indicate very qualitatively how it is possible to present all of quantum mechanics in terms of K (the Lagrangian formulation of quantum mechanics; Feynman's point of view).

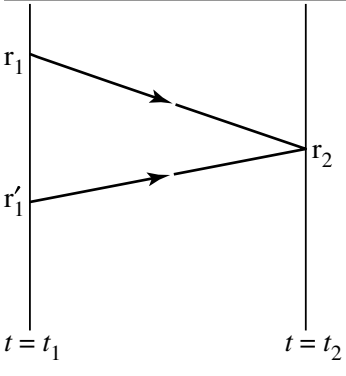


Figure 2: The probability amplitude $\psi(\mathbf{r}_2, t_2)$ can be obtained by summing the contributions of the various amplitudes $\psi(\mathbf{r}_1, t_1)$, $\psi(\mathbf{r}'_1, t_1)$, etc... corresponding to a given previous instant t_1. With each of the arrows of the figure is associated a "propagator" $K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$, $K(\mathbf{r}_2, t_2; \mathbf{r}'_1, t_1)$, etc...

2. Existence and properties of a propagator $K(2, 1)$

2-a. Existence of a propagator

The problem is to link directly the states of the system at two different times. This is possible if we use the evolution operator introduced in Complement F_{III}, since we can write:

$$|\psi(t_2)\rangle = U(t_2, t_1)|\psi(t_1)\rangle \tag{2}$$

Given $|\psi(t_2)\rangle$, it is easy to find the wave function $\psi(\mathbf{r}_2, t_2)$:

$$\psi(\mathbf{r}_2, t_2) = \langle \mathbf{r}_2 | \psi(t_2)\rangle \tag{3}$$

Substituting (2) into (3) and inserting the closure relation:

$$\int \mathrm{d}^3 r_1 |\mathbf{r}_1\rangle \langle \mathbf{r}_1| = \mathbb{1} \tag{4}$$

between $U(t_2, t_1)$ and $|\psi(t_1)\rangle$, we obtain:

$$\begin{aligned} \psi(\mathbf{r}_2, t_2) &= \int \mathrm{d}^3 r_1 \langle \mathbf{r}_2 | U(t_2, t_1) | \mathbf{r}_1 \rangle \langle \mathbf{r}_1 | \psi(t_1) \rangle \\ &= \int \mathrm{d}^3 r_1 \langle \mathbf{r}_2 | U(t_2, t_1) | \mathbf{r}_1 \rangle \psi(\mathbf{r}_1, t_1) \end{aligned} \tag{5}$$

The result is thus a formula identical to (1), on the condition that we set:

$$\langle \mathbf{r}_2 | U(t_2, t_1) | \mathbf{r}_1 \rangle = K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$$

Moreover, since we want to use formulas of the type of (1) only for $t_2 > t_1$, we can set $K = 0$ for $t_2 < t_1$. The exact definition of K then becomes:

$$K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = \langle \mathbf{r}_2 | U(t_2, t_1) | \mathbf{r}_1 \rangle \, \theta(t_2 - t_1) \tag{6}$$

where $\theta(t_2 - t_1)$ is the "step function":

$$\begin{aligned} &\theta(t_2 - t_1) = 1 \quad \text{if} \quad t_2 > t_1 \\ &\theta(t_2 - t_1) = 0 \quad \text{if} \quad t_2 < t_1 \end{aligned} \tag{7}$$

The introduction of $\theta(t_2 - t_1)$ is of both physical and mathematical interest. From the physical point of view, it is a simple way of compelling the secondary sources situated on the surface $t = t_1$ of Figure 2 to "radiate" only towards the future. For this reason, $K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$ as defined by (6) is called the *retarded propagator*. From the mathematical point of view, we shall see later that $K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$, because of the factor $\theta(t_2 - t_1)$, obeys a partial differential equation whose right-hand side is a delta function, which is the equation that defines a *Green's function*.

Comments:

(*i*) Note, however, that equation (5) remains valid even if $t_2 < t_1$. It is possible, moreover, to introduce mathematically an "advanced" propagator which would be different from zero only for $t_2 < t_1$ and which would also obey the equation defining a Green's function. Since the physical meaning of such an advanced propagator is not obvious at this stage, we shall not study it here.

(*ii*) When no ambiguity is possible, we shall simply write $K(2, 1)$ for $K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$.

2-b. Physical interpretation of $K(2, 1)$

This interpretation follows very simply from definition (6): $K(2, 1)$ represents the probability amplitude that the particle, starting from point $\mathbf{r}_1$, at time t_1, will arrive at

point $\mathbf{r}_2$ at a later time t_2. If we take as the initial state at time t_1 a state localized at point $\mathbf{r}_1$:

$$|\psi(t_1)\rangle = |\mathbf{r}_1\rangle \tag{8}$$

at time t_2, the state vector has become:

$$|\psi(t_2)\rangle = U(t_2,t_1)|\psi(t_1)\rangle = U(t_2,t_1)|\mathbf{r}_1\rangle \tag{9}$$

The probability amplitude of finding the particle at point $\mathbf{r}_2$ at this time is then:

$$\langle \mathbf{r}_2|\psi(t_2)\rangle = \langle \mathbf{r}_2|U(t_2,t_1)|\mathbf{r}_1\rangle \tag{10}$$

2-c. Expression for $K(2,1)$ in terms of the eigenstates of H

Assume that the Hamiltonian H does not depend explicitly on time, and call $|\varphi_n\rangle$ and E_n its eigenstates and eigenvalues:

$$H|\varphi_n\rangle = E_n|\varphi_n\rangle \tag{11}$$

According to formula (18) of $\mathrm{F_{III}}$, we have:

$$U(t_2,t_1) = \mathrm{e}^{-iH(t_2-t_1)/\hbar} \tag{12}$$

The closure relation:

$$\sum_n |\varphi_n\rangle \langle\varphi_n| = \mathbb{1} \tag{13}$$

enables us to write (12) in the form:

$$U(t_2,t_1) = \mathrm{e}^{-iH(t_2-t_1)/\hbar} \sum_n |\varphi_n\rangle \langle\varphi_n| \tag{14}$$

that is, taking (11) into account:

$$U(t_2,t_1) = \sum_n \mathrm{e}^{-iE_n(t_2-t_1)/\hbar} |\varphi_n\rangle \langle\varphi_n| \tag{15}$$

To calculate $K(2,1)$, it then suffices to take the matrix element of both sides of (15) between $\langle \mathbf{r}_2|$ and $|\mathbf{r}_1\rangle$ and to multiply it by $\theta(t_2-t_1)$. Since:

$$\langle \mathbf{r}_2|\varphi_n\rangle = \varphi_n(\mathbf{r}_2) \tag{16}$$

$$\langle \varphi_n|\mathbf{r}_1\rangle = \varphi_n^*(\mathbf{r}_1) \tag{17}$$

this leads to:

$$K(\mathbf{r}_2,t_2;\mathbf{r}_1,t_1) = \theta(t_2-t_1) \sum_n \varphi_n^*(\mathbf{r}_1)\varphi_n(\mathbf{r}_2)\, \mathrm{e}^{-iE_n(t_2-t_1)/\hbar} \tag{18}$$

2-d. Equation satisfied by $K(2,1)$

$\varphi_n(\mathbf{r}_2)\,\mathrm{e}^{-iE_nt_2/\hbar}$ is a solution of the Schrödinger equation. We deduce from this that, in the $\{|\mathbf{r}\rangle\}$ representation:

$$\left[i\hbar\frac{\partial}{\partial t_2} - H\left(\mathbf{r}_2, \frac{\hbar}{i}\boldsymbol{\nabla}_2\right)\right]\varphi_n(\mathbf{r}_2)\mathrm{e}^{-iE_nt_2/\hbar} = 0 \tag{19}$$

where $\boldsymbol{\nabla}_2$ is a condensed notation which designates the three operators $\dfrac{\partial}{\partial x_2}$, $\dfrac{\partial}{\partial y_2}$, $\dfrac{\partial}{\partial z_2}$. Let us then apply, to both sides of equation (18), the operator:

$$i\hbar\frac{\partial}{\partial t_2} - H\left(\mathbf{r}_2, \frac{\hbar}{i}\boldsymbol{\nabla}_2\right)$$

which acts only on the variables $\mathbf{r}_2$ and t_2. We know [*cf.* Appendix II, relation (44)] that:

$$\frac{\partial}{\partial t_2}\,\theta(t_2 - t_1) = \delta(t_2 - t_1) \tag{20}$$

Consequently, using (19), we obtain:

$$\left[i\hbar\frac{\partial}{\partial t_2} - H\left(\mathbf{r}_2, \frac{\hbar}{i}\boldsymbol{\nabla}_2\right)\right]K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = i\hbar\;\delta(t_2 - t_1)\sum_n \varphi_n^*(\mathbf{r}_1)\,\varphi_n(\mathbf{r}_2)\;\mathrm{e}^{-iE_n(t_2-t_1)/\hbar} \tag{21}$$

Because of the presence of $\delta(t_2 - t_1)$, we can replace $t_2 - t_1$ by zero in the sum over n appearing on the right-hand side of (21). This makes the exponential equal to 1. We are thus left with the quantity $\sum_n \varphi_n(\mathbf{r}_2)\varphi_n^*(\mathbf{r}_1)$, which, according to (13), (16) and (17), is equal to $\delta(\mathbf{r}_2 - \mathbf{r}_1)$ [taking the matrix element of (13) between $\langle\mathbf{r}_2|$ and $|\mathbf{r}_1\rangle$]. Finally, K satisfies the equation:

$$\left[i\hbar\frac{\partial}{\partial t_2} - H\left(\mathbf{r}_2, \frac{\hbar}{i}\boldsymbol{\nabla}_2\right)\right]K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = i\hbar\;\delta(t_2 - t_1)\,\delta(\mathbf{r}_2 - \mathbf{r}_1) \tag{22}$$

The solutions of equation (22), whose right-hand side is proportional to a four-dimensional "delta function", are called *Green's functions*. It can be shown that, to determine $K(2,1)$ completely, it suffices to associate with (22) the boundary condition:

$$K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = 0 \qquad \text{if} \quad t_2 < t_1 \tag{23}$$

Equations (22) and (23) have interesting implications, in particular with regard to perturbation theory, which we shall study in Chapter XI.

3. Lagrangian formulation of quantum mechanics

3-a. Concept of a space-time path

Let us consider, in space-time, the two points $(\mathbf{r}_1, t_1)$ and $(\mathbf{r}_2, t_2)$ (*cf.* Fig. 3; t is plotted as the abscissa, and the ordinate axis represents the set of the three spatial axes). Choose N intermediate times t_{α_i} $(i = 1, 2, ..., N)$, evenly spaced between t_1 and t_2:

$$t_1 < t_{\alpha_1} < t_{\alpha_2} < ... < t_{\alpha_{N-1}} < t_{\alpha_N} < t_2 \tag{24}$$

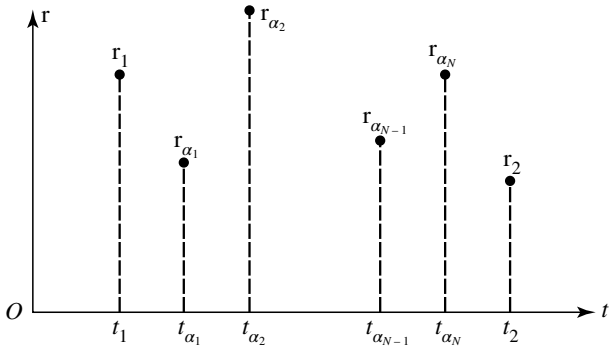


Figure 3: Diagram associated with a "spacetime path": one picks N intermediate times $t_{\alpha_i} (i = 1, 2, ..., N)$ evenly spaced between t_1 and t_2, and chooses for each of them a value of $\mathbf{r}$.

and, for each of them, a position $\mathbf{r}_{\alpha_i}$ in space. We can thus construct, when N approaches infinity, a function $\mathbf{r}(t)$ (which we shall assume to be continuous) such that:

$$\mathbf{r}(t_1) = \mathbf{r}_1 \tag{25a}$$

$$\mathbf{r}(t_2) = \mathbf{r}_2 \tag{25b}$$

$\mathbf{r}(t)$ is said to define a *space-time path* between $(\mathbf{r}_1, t_1)$ and $(\mathbf{r}_2, t_2)$: such a path might be thought of as the trajectory of a physical point leaving point $\mathbf{r}_1$ at time t_1 and arriving at $\mathbf{r}_2$ at time t_2.

3-b. Decomposition of $K(2,1)$ into a sum of partial amplitudes

We first return to the case where the number N of intermediate times is finite. Formula (10) of Complement F_{III} enables us to write:

$$U(t_2, t_1) = U(t_2, t_{\alpha_N})\, U(t_{\alpha_N}, t_{\alpha_{N-1}}) \ldots U(t_{\alpha_2}, t_{\alpha_1})\, U(t_{\alpha_1}, t_1) \tag{26}$$

We now take the matrix elements of both sides of (26) between $\langle \mathbf{r}_2 |$ and $| \mathbf{r}_1 \rangle$ and insert the closure relation relative to the $\{|\mathbf{r}\rangle\}$ representation for each intermediate time t_{α_i}. According to (6) and (24), we thus obtain:

$$K(2,1) = \int \mathrm{d}^3 r_{\alpha_N} \int \mathrm{d}^3 r_{\alpha_{N-1}} \ldots \int \mathrm{d}^3 r_{\alpha_2} \int \mathrm{d}^3 r_{\alpha_1}\; K(2, \alpha_N)\, K(\alpha_N, \alpha_{N-1}) \ldots \times K(\alpha_2, \alpha_1)\, K(\alpha_1, 1) \tag{27}$$

Now consider the product:

$$K(2, \alpha_N)\, K(\alpha_N, \alpha_{N-1}) \ldots K(\alpha_2, \alpha_1)\, K(\alpha_1, 1) \tag{28}$$

Generalizing the argument of § 2-b, we can interpret this term as being the probability amplitude for the particle, having left point 1 $(\mathbf{r}_1, t_1)$, to arrive at point 2 $(\mathbf{r}_2, t_2)$, having

passed successively through all points $\alpha_i(\mathbf{r}_{\alpha_i}, t_{\alpha_i})$ of Figure 3. Note that, in formula (27), one is summing over all possible positions $\mathbf{r}_{\alpha_i}$ at each time t_{α_i}.

We now let N approach infinity[1]. A series of points α_i then defines a space-time path between 1 and 2, and the product (28) associated with it becomes the probability amplitude for the particle to follow this path. Of course, the number of integrations in formula (27) becomes infinite. It is understandable, however, that the summation over the set of possible positions at each time should reduce to a summation over the set of possible paths. $K(2,1)$ is thus seen to be a sum (in fact, an integral) that corresponds to the coherent superposition of the amplitudes associated with all possible space-time paths starting from 1 and ending at 2.

3-c. Feynman's postulates

The concepts of a propagator and a space-time path permit a new formulation of the postulate concerning the time evolution of physical systems. We shall outline here such a formulation for the case of a spinless particle.

We define $K(2,1)$ directly as being the probability amplitude for the particle, starting from $\mathbf{r}_1$ at time t_1, to arrive at $\mathbf{r}_2$ at time t_2. We then postulate that:

(i) $K(2,1)$ is the sum of an infinity of partial amplitudes, one for each of the space-time paths connecting $(\mathbf{r}_1, t_1)$ and $(\mathbf{r}_2, t_2)$.

(ii) The partial amplitude $K_\Gamma(2,1)$ associated with one of these paths Γ is determined in the following manner: let S_Γ be the *classical action* calculated along Γ, that is:

$$S_\Gamma = \int_{(\Gamma)} \mathcal{L}(\mathbf{r}, \mathbf{p}, t)\, \mathrm{d}t \tag{29}$$

where $\mathcal{L}(\mathbf{r}, \mathbf{p}, t)$ is the Lagrangian of the particle (*cf.* ppendix III). $K_\Gamma(2,1)$ is then equal to:

$$K_\Gamma(2,1) = N\, \mathrm{e}^{\frac{i}{\hbar} S_\Gamma} \tag{30}$$

where N is a normalization constant (which can be determined explicitly).

It can be shown that the Schrödinger equation follows as a consequence of these two postulates, which also lead to the canonical commutation relation between the components of the observables $\mathbf{R}$ and $\mathbf{P}$. The two preceding postulates therefore permit a formulation of quantum mechanics which is different from that of Chapter III, but equivalent.

3-d. The classical limit and Hamilton's principle

The formulation we have just evoked is particularly useful for discussing the relation between quantum and classical mechanics.

Consider a situation in which the actions S_Γ are much larger than $\hbar$. In this case, the variation ΔS_Γ of the action between two different paths, even if its relative

[1] In this treatment, we make no attempt to be mathematically rigorous.

value is small $\left(\frac{\Delta S_\Gamma}{S_\Gamma} \ll 1\right)$, is usually much larger than $\hbar$. Consequently, the phase of $K_\Gamma(2,1)$ varies rapidly, and the contributions to the global amplitude $K(2,1)$ of most of the paths Γ cancel out by interference. Let us assume, however, that there exists a path Γ_0 for which the *action is stationary* (meaning that it does not vary, to first order, when one goes from Γ_0 to another infinitesimally close path). The amplitude $K_{\Gamma_0}(2,1)$ then interferes constructively with those of the paths next to Γ_0, since, this time, their phases remain practically equal. Consequently, when the actions S_Γ are much larger than $\hbar$, one is in a "quasi-classical" situation: to obtain $K(2,1)$, one can ignore all the paths except Γ_0 and paths infinitely close to it; it can then be said that, between points 1 and 2, the particle follows the trajectory Γ_0. Now this is indeed the classical trajectory, defined by Hamilton's principle as being the path along which the action is minimal. Feynman's postulates therefore include, at the classical limit, Hamilton's principle of least action. They enable us, moreover, to associate with it the following picture: it is the wave associated with a particle which, "exploring" the various possible paths, picks the one for which the action will be the smallest.

The Lagrangian formulation of quantum mechanics presents numerous other advantages, which we shall not examine in detail. Let us point out, for example, that it lends itself easily to relativistic generalization since one is already reasoning in space-time. Moreover, it can be applied to any classical system (not necessarily mechanical) governed by a variational principle (for example, a field).

However, it has a certain number of drawbacks on the mathematical level (summation over an infinite number of paths, the limit $N \longrightarrow \infty$, ...).

References and suggestions for further reading:

Feynman's original article (2.38); Feynman and Hibbs (2.25); Bjorken and Drell (2.6), Chaps. 6 and 7.

Complement K_{III}

Unstable states. Lifetime

1	**Introduction**	**343**
2	**Definition of the lifetime**	**344**
3	**Phenomenological description of the instability of a state**	**345**

1. Introduction

Consider a conservative system (a system whose Hamiltonian H is time-independent). Assume that at time $t = 0$ the state of the system is one of the eigenstates $|\varphi_n\rangle$ of the Hamiltonian, of energy E_n:

$$|\psi(0)\rangle = |\varphi_n\rangle \tag{1}$$

with:

$$H|\varphi_n\rangle = E_n|\varphi_n\rangle \tag{2}$$

In this case, the system remains indefinitely in the same state (a stationary state, § D-2-b of Chapter III).

We shall study the hydrogen atom in Chapter VII by solving the eigenvalue equation of its Hamiltonian, which is a time-independent operator. The states of the hydrogen atom (that is, the possible values of its energy) which we shall find are in very good agreement with the experimentally measured energies. However, it is known that most of these states are actually *unstable*: if, at the instant $t = 0$, the atom is in an excited state (an eigenstate $|\varphi_n\rangle$ corresponding to an energy E_n greater than that of the ground state, which is the lowest energy state), it generally "falls back" into this ground state by emitting one or several photons. The state $|\varphi_n\rangle$ is not really, therefore, a stationary state in this case.

This problem arises from the fact that, in calculations of the type used in Chapter VII, the system under study (the hydrogen atom) is treated as if it were totally isolated, while it is actually in constant interaction with the electromagnetic field. Although the evolution of the global system "atom + electromagnetic field" can be perfectly well described by a Hamiltonian, it is not rigorously possible to define a Hamiltonian for the hydrogen atom alone [*cf.* comment (v) of § 5-b of Complement E_{III}]. However, since the coupling between the atom and the field happens to be weak (it can be shown that its "force" is characterized by the fine structure constant $\alpha \simeq \dfrac{1}{137}$, which we shall introduce in Chapter VII), the approximation consisting of completely neglecting the existence of the electromagnetic field is very good, except, of course, if we are interested precisely in the instability of the states.

Comments:

(*i*) If, at the initial instant, a strictly conservative and isolated system is in a state formed by a linear combination of several stationary states, it evolves over time, and does not always remain in the same state. But its Hamiltonian is a constant of the motion, and consequently (*cf.* Chapter III, § D-2-c), the probability of finding one energy value or another is independent of time, as is the mean value of the energy. On the other hand, in the case of an unstable state, one state is transformed irreversibly into another, with a loss of energy for the system: this energy is taken away by the photons emitted[1].

(*ii*) The instability of the excited states of an atom is caused by the spontaneous emission of photons; the ground state is stable, since there exists no lower energy state. Recall, nevertheless, that atoms can also absorb light energy and so ascend to higher energy levels.

We intend to indicate here how to take the instability of a state into account phenomenologically. The description will not be rigorous, as we shall continue to consider the system as if it were isolated. We shall try to incorporate this instability as simply as possible into the quantum description of the system.

Complement D_{XIII} presents a more rigorous treatment of this problem, justifying the phenomenological approach used here.

2. Definition of the lifetime

Experiments show that the instability of a state can often be characterized by just one parameter τ, having the dimensions of a time, which is called the *lifetime* of the state. More precisely, if one prepares the system at time $t = 0$ in the unstable state $|\varphi_n\rangle$, one observes that the probability $\mathcal{P}(t)$ of its still being excited at a later time t is equal to:

$$\mathcal{P}(t) = \mathrm{e}^{-t/\tau} \tag{3}$$

This result can also be expressed in the following way. Consider a large number $\mathcal{N}$ of identical independent systems, all prepared at time $t = 0$ in the state $|\varphi_n\rangle$. At time t, there will remain $N(t) = \mathcal{N}\,\mathrm{e}^{-t/\tau}$ in this state. Between times t and $t + \mathrm{d}t$, a certain number $\mathrm{d}n(t)$ of systems leave the unstable state:

$$\mathrm{d}n(t) = N(t) - N(t + \mathrm{d}t) = -\frac{\mathrm{d}N(t)}{\mathrm{d}t}\mathrm{d}t = N(t)\frac{\mathrm{d}t}{\tau} \tag{4}$$

For each of the $N(t)$ systems that are still in the state $|\varphi_n\rangle$ at time t, a probability can therefore be defined:

$$\mathrm{d}\varpi(t) = \frac{\mathrm{d}n(t)}{N(t)} = \frac{\mathrm{d}t}{\tau} \tag{5}$$

of their leaving this state during the time interval $\mathrm{d}t$ following the instant t. We see that $\mathrm{d}\varpi(t)$ is independent of t: the system is said to have a *probability per unit time* $\dfrac{1}{\tau}$ of leaving the unstable state.

[1] which, moreover, may also take away linear and angular momentum.

Comments:

(*i*) Let us calculate the mean value of the time during which the system remains in the unstable state. It is equal to:

$$\int_0^\infty t\ \mathrm{e}^{-t/\tau}\frac{\mathrm{d}t}{\tau} = \tau \tag{6}$$

τ is therefore the mean time the system spends in the state $|\varphi_n\rangle$; this is why it is called the lifetime of this state.

For a stable state, $\mathcal{P}(t)$ is always equal to 1, and the lifetime τ is infinite.

(*ii*) A remarkable property of the lifetime τ is that it does not depend on the procedure used to prepare the system in the unstable state, that is, on its previous "history": the lifetime is a characteristic of the unstable state itself.

(*iii*) According to the time-energy uncertainty relation (§ D-2-e of Chapter III), the time τ characteristic of the evolution of an unstable state is associated with an uncertainty in the energy ΔE given by:

$$\Delta E \simeq \frac{\hbar}{\tau} \tag{7}$$

One indeed finds that the energy of an unstable state cannot be determined with arbitrary accuracy, but only to within an uncertainty of the order of ΔE. ΔE is called the *natural width* of this state. For the case of the hydrogen atom, the width of the various states is negligible compared to their separation. This explains why we can treat them, in a first approximation, as if they were stable.

3. Phenomenological description of the instability of a state

First let us consider a conservative system, prepared, at the initial time, in the eigenstate $|\varphi_n\rangle$ of the Hamiltonian H. According to the rule (D-54) of Chapter III, the state vector, at time t, becomes:

$$|\psi(t)\rangle = \mathrm{e}^{-iE_nt/\hbar}|\varphi_n\rangle \tag{8}$$

The probability $\mathcal{P}_n(t)$ of finding, in a measurement at time t, the system in the state $|\varphi_n\rangle$ is:

$$\mathcal{P}_n(t) = \left|\mathrm{e}^{-iE_nt/\hbar}\right|^2 \tag{9}$$

Since the energy E_n is real (H being an observable), this probability is constant and equal to 1: we again find that $|\varphi_n\rangle$ is a stationary state.

Let us examine what would happen if, in expression (9), we replaced the energy E_n by the complex number:

$$E_n' = E_n - i\hbar\frac{\gamma_n}{2} \tag{10}$$

The probability $\mathcal{P}_n(t)$ then becomes:

$$\mathcal{P}'_n(t) = \left| e^{-i(E_n - i\hbar\frac{\gamma_n}{2})t/\hbar} \right|^2 = e^{-\gamma_n t} \tag{11}$$

In this case, the probability of finding the system in the state $|\varphi_n\rangle$ decreases exponentially with time, as in formula (3). Therefore, to take into account phenomenologically the instability of a state $|\varphi_n\rangle$ whose lifetime is τ_n, it suffices to add, as in (10), an imaginary part to its energy, and set:

$$\gamma_n = \frac{1}{\tau_n} \tag{12}$$

Comment:

When E_n is replaced by E'_n, the norm of the state vector written in (8) becomes $e^{-\gamma_n t/2}$ and therefore varies with time. This result is not surprising. We saw in § D-1-c of Chapter III that the conservation of the norm of the state vector arose from the Hermitian nature of the Hamiltonian operator; now, an operator whose eigenvalues are complex, as are the E'_n, cannot be Hermitian. Of course, as we pointed out in § 1, this is due to the fact that the system under study is part of a larger system (it is interacting with the electromagnetic field) and its evolution cannot be described rigorously by means of a Hamiltonian. It is already rather remarkable that its evolution can be simply explained by introducing a "Hamiltonian" with complex eigenvalues.

Complement L_{III}

Exercises

1. In a one-dimensional problem, consider a particle whose wave function is:

$$\psi(x) = N \frac{e^{ip_0 x/\hbar}}{\sqrt{x^2 + a^2}}$$

where a and p_0 are real constants and N is a normalization coefficient.

a. Determine N so that $\psi(x)$ is normalized.

b. The position of the particle is measured. What is the probability of finding a result between $-\frac{a}{\sqrt{3}}$ and $+\frac{a}{\sqrt{3}}$?

c. Calculate the mean value of the momentum of a particle which has $\psi(x)$ for its wave function.

2. Consider, in a one-dimensional problem, a particle of mass m whose wave function at time t is $\psi(x,t)$.

a. At time t, the distance d of this particle from the origin is measured. Write, as a function of $\psi(x,t)$, the probability $\mathcal{P}(d_0)$ of finding a result greater than a given length d_0. What are the limits of $\mathcal{P}(d_0)$ when $d_0 \longrightarrow 0$ and when $d_0 \longrightarrow \infty$?

b. Instead of performing the measurement of question *a*, one measures the velocity v of the particle at time t. Express, as a function of $\psi(x,t)$, the probability of finding a result greater than a given value v_0.

3. The wave function of a free particle, in a one-dimensional problem, is given at time $t = 0$ by:

$$\psi(x,0) = N \int_{-\infty}^{+\infty} dk\, e^{-|k|/k_0}\, e^{ikx}$$

where k_0 and N are constants.

a. What is the probability $\mathcal{P}(p_1, 0)$ that a measurement of the momentum, performed at time $t = 0$, will yield a result included between $-p_1$ and $+p_1$? Sketch the function $\mathcal{P}(p_1, 0)$.

b. What happens to this probability $\mathcal{P}(p_1, t)$ if the measurement is performed at time t? Interpret.

c. What is the form of the wave packet at time $t = 0$? Calculate for this time the product $\Delta X \cdot \Delta P$; what is your conclusion? Describe qualitatively the subsequent evolution of the wave packet.

4. Spreading of a free wave packet

Consider a free particle.

a. Show, applying Ehrenfest's theorem, that $\langle X \rangle$ is a linear function of time, the mean value $\langle P \rangle$ remaining constant.

b. Write the equations of motion for the mean values $\langle X^2 \rangle$ and $\langle XP+PX \rangle$. Integrate these equations.

c. Show that, with a suitable choice of the time origin, the root mean square deviation ΔX is given by:

$$(\Delta X)^2 = \frac{1}{m^2}(\Delta P)_0^2\, t^2 + (\Delta X)_0^2$$

where $(\Delta X)_0$ and $(\Delta P)_0$ are the root mean square deviations at the initial time.

How does the width of the wave packet vary as a function of time (see § 3-c of Complement G_I)? Give a physical interpretation.

5. Particle subjected to a constant force

In a one-dimensional problem, consider a particle having a potential energy $V(X)$ given by $V(X) = -fX$, where f is a positive constant [$V(X)$ arises, for example, from a gravity field or a uniform electric field].

a. Write Ehrenfest's theorem for the mean values of the position X and the momentum P of the particle. Integrate these equations; compare with the classical motion.

b. Show that the root mean square deviation ΔP does not vary over time.

c. Write the Schrödinger equation in the $\{|p\rangle\}$ representation. Deduce from it a relation between $\dfrac{\partial}{\partial t}|\langle p|\psi(t)\rangle|^2$ and $\dfrac{\partial}{\partial p}|\langle p|\psi(t)\rangle|^2$. Integrate the equation thus obtained; give a physical interpretation.

6. Consider the three-dimensional wave function:

$$\psi(x, y, z) = N\mathrm{e}^{-\left[\frac{|x|}{2a}+\frac{|y|}{2b}+\frac{|z|}{2c}\right]}$$

where a, b and c are three positive lengths.

a. Calculate the constant N which normalizes ψ.

b. Calculate the probability that a measurement of X will yield a result included between 0 and a.

c. Calculate the probability that simultaneous measurements of Y and Z will yield results included respectively between $-b$ and $+b$, and $-c$ and $+c$.

d. Calculate the probability that a measurement of the momentum will yield a result included in the element $\mathrm{d}p_x\ \mathrm{d}p_y\ \mathrm{d}p_z$ centered at the point $p_x = p_y = 0$; $p_z = \hbar/c$.

7. Let $\psi(x, y, z) = \psi(\mathbf{r})$ be the normalized wave function of a particle. Express in terms of $\psi(\mathbf{r})$ the probability for:

a. a measurement of the abscissa X, to yield a result included between x_1 and x_2;

b. a measurement of the component P_x of the momentum, to yield a result included between p_1 and p_2;

c. simultaneous measurements of X and P_z, to yield:

$$x_1 \leq x \leq x_2$$
$$p_z \geq 0$$

d. simultaneous measurements of P_x, P_y, P_z, to yield:

$$p_1 \leq p_x \leq p_2$$
$$p_3 \geq p_y \leq p_4$$
$$p_5 \geq p_z \leq p_6$$

Show that this probability is equal to the result of b when p_3, $p_5 \longrightarrow -\infty$; p_4, $p_6 \longrightarrow +\infty$

e. a measurement of the component $U = \dfrac{1}{\sqrt{3}}(X + Y + Z)$ of the position, to yield a result included between u_1 and u_2.

8. Let $\mathbf{J}(\mathbf{r})$ be the probability current associated with a wave function $\psi(\mathbf{r})$ describing the state of a particle of mass m [Chap. III, relations (D-17) and (D-19)].

a. Show that:

$$m \int \mathrm{d}^3 r\, \mathbf{J}(\mathbf{r}) = \langle \mathbf{P} \rangle$$

where $\langle \mathbf{P} \rangle$ is the mean value of the momentum.

b. Consider the operator $\mathbf{L}$ (orbital angular momentum) defined by $\mathbf{L} = \mathbf{R} \times \mathbf{P}$. Are the three components of $\mathbf{L}$ Hermitian operators? Establish the relation:

$$m \int \mathrm{d}^3 r\, [\mathbf{r} \times \mathbf{J}(\mathbf{r})] = \langle \mathbf{L} \rangle$$

9. One wants to show that the physical state of a (spinless) particle is completely defined by specifying the probability density $\rho(\mathbf{r}) = |\psi(\mathbf{r})|^2$ and the probability current $\mathbf{J}(\mathbf{r})$.

a. Assume the function $\psi(\mathbf{r})$ known and let $\xi(\mathbf{r})$ be its argument:

$$\psi(\mathbf{r}) = \sqrt{\rho(\mathbf{r})}\mathrm{e}^{i\xi(\mathbf{r})}$$

Show that:

$$\mathbf{J}(\mathbf{r}) = \frac{\hbar}{m}\rho(\mathbf{r})\, \boldsymbol{\nabla}\xi(\mathbf{r})$$

Deduce that two wave functions leading to the same density $\rho(\mathbf{r})$ and current $\mathbf{J}(\mathbf{r})$ can differ only by a global phase factor.

b. Given arbitrary functions $\rho(\mathbf{r})$ and $\mathbf{J}(\mathbf{r})$, show that a quantum state $\psi(\mathbf{r})$ can be associated with them only if $\boldsymbol{\nabla} \times \mathbf{v}(\mathbf{r}) = 0$, where $\mathbf{v}(\mathbf{r}) = \mathbf{J}(\mathbf{r})/\rho(\mathbf{r})$ is the velocity associated with the probability fluid.

c. Now assume that the particle is submitted to a magnetic field $\mathbf{B}(\mathbf{r}) = \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r})$ [see Chap. III, definition (D-20) of the probability current in this case]. Show that:

$$\mathbf{J} = \frac{\rho(\mathbf{r})}{m} [\hbar \boldsymbol{\nabla} \xi(\mathbf{r}) - q\mathbf{A}(\mathbf{r})]$$

and:

$$\boldsymbol{\nabla} \times \mathbf{v}(\mathbf{r}) = -\frac{q}{m}\mathbf{B}(\mathbf{r})$$

10. Virial theorem

a. In a one-dimensional problem, consider a particle with the Hamiltonian:

$$H = \frac{P^2}{2m} + V(X)$$

where:

$$V(X) = \lambda X^n$$

Calculate the commutator $[H, XP]$. If there exists one or several stationary states $|\varphi\rangle$ in the potential V, show that the mean values $\langle T \rangle$ and $\langle V \rangle$ of the kinetic and potential energies in these states satisfy the relation: $2\langle T \rangle = n\langle V \rangle$.

b. In a three-dimensional problem, H is written:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R})$$

Calculate the commutator $[H, \mathbf{R}\cdot\mathbf{P}]$. Assume that $V(\mathbf{R})$ is a homogeneous function of nth order in the variables X, Y, Z. What relation necessarily exists between the mean kinetic energy and the mean potential energy of the particle in a stationary state?

Apply this to a particle moving in the potential $V(r) = -e^2/r$ (hydrogen atom).

Recall that a homogeneous function V of nth degree in the variables x, y and z by definition satisfies the relation:

$$V(\alpha x, \alpha y, \alpha z) = \alpha^n V(x, y, z)$$

and satisfies Euler's identity:

$$x\frac{\partial V}{\partial x} + y\frac{\partial V}{\partial y} + z\frac{\partial V}{\partial z} = nV(x, y, z).$$

c. Consider a system of N particles of positions $\mathbf{R}_i$ and momenta $\mathbf{P}_i (i = 1, 2, ...N)$. When their potential energy is a homogeneous (nth degree) function of the set of components X_i, Y_i, Z_i, can the results obtained above be generalized? Apply these results to an arbitrary molecule formed of nuclei of charges $-Z_i q$ and electrons of charge q. All these particles interact by pairs through Coulomb forces. In a stationary state of the molecule, what relation exists between the kinetic energy of the system of particles and their energy of mutual interaction?

11. Two-particle wave function

In a one-dimensional problem, consider a system of two particles (1) and (2) with which is associated the wave function $\psi(x_1, x_2)$.

a. What is the probability of finding, in a measurement of the positions X_1 and X_2 of the two particles, a result such that:

$$x \leq x_1 \leq x + \mathrm{d}x$$
$$\alpha \leq x_2 \leq \beta?$$

b. What is the probability of finding particle (1) between x and $x + \mathrm{d}x$ [when no observations are made on particle (2)]?

c. Give the probability of finding at least one of the particles between α and β.

d. Give the probability of finding one and only one particle between α and β.

e. What is the probability of finding the momentum of particle (1) included between p' and p'' and the position of particle (2) between α and β?

f. The momenta P_1 and P_2 of the two particles are measured; what is the probability of finding $p' \leq p_1 \leq p''$; $p''' \leq p_2 \leq p''''$?

g. The only quantity measured is the momentum P_1 of the first particle. Calculate, first from the results of e and then from those of f, the probability of finding this momentum included between p' and p''. Compare the two results obtained.

h. The algebraic distance $X_1 - X_2$ between the two particles is measured; what is the probability of finding a result included between $-d$ and $+d$? What is the mean value of this distance?

12. Infinite one-dimensional well

Consider a particle of mass m placed in the potential:

$$V(x) = 0 \qquad \text{if} \quad 0 \leq x \leq a.$$
$$V(x) = +\infty \qquad \text{if} \quad x < 0 \quad \text{or} \quad x > a.$$

$|\varphi_n\rangle$ are the eigenstates of the Hamiltonian H of the system, and their eigenvalues are $E_n = \dfrac{n^2\pi^2\hbar^2}{2ma^2}$ (*cf.* Complement $\mathrm{H_I}$). The state of the particle at the instant $t = 0$ is:

$$|\psi(0)\rangle = a_1|\varphi_1\rangle + a_2|\varphi_2\rangle + a_3|\varphi_3\rangle + a_4|\varphi_4\rangle$$

a. What is the probability, when the energy of the particle in the state $|\psi(0)\rangle$ is measured, of finding a value smaller than $\dfrac{3\pi^2\hbar^2}{ma^2}$?

b. What is the mean value and what is the root mean square deviation of the energy of the particle in the state $|\psi(0)\rangle$?

c. Calculate the state vector $|\psi(t)\rangle$ at the instant t. Do the results found in a and b at the instant $t = 0$ remain valid at an arbitrary time t?

d. When the energy is measured, the result of $\dfrac{8\pi^2\hbar^2}{ma^2}$ is found. After the measurement, what is the state of the system? What is the result if the energy is measured again?

13. Infinite two-dimensional well (*cf.* Complement G_{II})

In a two-dimensional problem, consider a particle of mass m; its Hamiltonian H is written:

$$H = H_x + H_y$$

with:

$$H_x = \frac{P_x^2}{2m} + V(X) \qquad H_y = \frac{P_y^2}{2m} + V(Y)$$

The potential energy $V(x)$ [or $V(y)$] is zero when x (or y) is included in the interval [0, a] and is infinite everywhere else.

a. Of the following sets of operators, which form a C.S.C.O.?

$$\{H\},\ \{H_x\},\ \{H_x, H_y\},\ \{H, H_x\}$$

b. Consider a particle whose wave function is:

$$\psi(x, y) = N \cos\frac{\pi x}{a} \cos\frac{\pi y}{a} \sin\frac{2\pi x}{a} \sin\frac{2\pi y}{a}$$

when $0 \leq x \leq a$ and $0 \leq y \leq a$, and is zero everywhere else (N is a constant).

α. What is the mean value $\langle H\rangle$ of the energy of the particle? If the energy H is measured, what results can be found, and with what probabilities?

β. The observable H_x is measured; what results can be found, and with what probabilities? If this measurement yields the result $\dfrac{\pi^2\hbar^2}{2ma^2}$, what will be the results of a subsequent measurement of H_y, and with what probabilities?

γ. Instead of performing the preceding measurements, one now performs a simultaneous measurement of H_x and P_y. What are the probabilities of finding:

$$E_x = \frac{9\pi^2\hbar^2}{2ma^2}$$

and:

$$p_0 \leq p_y \leq p_0 + \mathrm{d}p\ ?$$

14. Consider a physical system whose state space, which is three-dimensional, is spanned by the orthonormal basis formed by the three kets $|u_1\rangle$, $|u_2\rangle$, $|u_3\rangle$. In this basis, the Hamiltonian operator H of the system and the two observables A and B are written:

$$H = \hbar\omega_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad ; \quad A = a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad ; \quad B = b \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where ω_0, a and b are positive real constants.

The physical system at time $t = 0$ is in the state:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}|u_1\rangle + \frac{1}{2}|u_2\rangle + \frac{1}{2}|u_3\rangle$$

a. At time $t = 0$, the energy of the system is measured. What values can be found, and with what probabilities? Calculate, for the system in the state $|\psi(0)\rangle$, the mean value $\langle H\rangle$ and the root mean square deviation ΔH.

b. Instead of measuring H at time $t = 0$, one measures A; what results can be found, and with what probabilities? What is the state vector immediately after the measurement?

c. Calculate the state vector $|\psi(t)\rangle$ of the system at time t.

d. Calculate the mean values $\langle A\rangle(t)$ and $\langle B\rangle(t)$ of A and B at time t. What comments can be made?

e. What results are obtained if the observable A is measured at time t? Same question for the observable B. Interpret.

15. Interaction picture (It is recommended that Complement $\mathrm{F_{III}}$ and perhaps Complement $\mathrm{G_{III}}$ be read before this exercise is undertaken.)

Consider an arbitrary physical system. Denote its Hamiltonian by $H_0(t)$ and the corresponding evolution operator by $U_0(t, t')$:

$$\begin{cases} i\hbar\dfrac{\partial}{\partial t}U_0(t, t_0) = H_0(t)\ U_0(t, t_0) \\ U_0(t_0, t_0) = \mathbb{1} \end{cases}$$

Now assume that the system is perturbed in such a way that its Hamiltonian becomes:

$$H(t) = H_0(t) + W(t)$$

The state vector of the system in the "interaction picture", $|\psi_I(t)\rangle$, is defined from the state vector $|\psi_S(t)\rangle$ in the Schrödinger picture by:

$$|\psi_I(t)\rangle = U_0^\dagger(t, t_0)|\psi_S(t)\rangle$$

a. Show that the evolution of $|\psi_I(t)\rangle$ is given by:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi_I(t)\rangle = W_I(t)|\psi_I(t)\rangle$$

where $W_I(t)$ is the transform of operator $W(t)$ under the unitary transformation associated with $U_0^\dagger(t,t_0)$:

$$W_I(t) = U_0^\dagger(t,t_0)W(t)U_0(t,t_0)$$

Explain qualitatively why, when the perturbation $W(t)$ is much smaller than $H_0(t)$, the motion of the vector $|\psi_I(t)\rangle$ is much slower than that of $|\psi_S(t)\rangle$.

b. Show that the preceding differential equation is equivalent to the integral equation:

$$|\psi_I(t)\rangle = |\psi_I(t_0)\rangle + \frac{1}{i\hbar}\int_{t_0}^{t}\mathrm{d}t'\,W_I(t')|\psi_I(t')\rangle$$

where: $|\psi_I(t_0)\rangle = |\psi_S(t)\rangle$.

c. Solving this integral equation by iteration, show that the ket $|\psi_I(t)\rangle$ can be expanded in a power series in W of the form:

$$|\psi_I(t)\rangle = \left\{\mathbb{1} + \frac{1}{i\hbar}\int_{t_0}^{t}\mathrm{d}t'\,W_I(t') + \frac{1}{(i\hbar)^2}\int_{t_0}^{t}\mathrm{d}t'\,W_I(t')\int_{t_0}^{t'}\mathrm{d}t''\,W_I(t'') + \ \ldots\right\}|\psi_I(t_0)\rangle$$

16. Correlations between two particles

(It is recommended that the Complement E_{III} be read in order to answer question *e* of this exercise.)

Consider a physical system formed by two particles (1) and (2), of the same mass m, which do not interact with each other and which are both placed in an infinite potential well of width a (*cf.* Complement H_I, § 2-c). Denote by $H(1)$ and $H(2)$ the Hamiltonians of each of the two particles and by $|\varphi_n(1)\rangle$ and $|\varphi_q(2)\rangle$ the corresponding eigenstates of the first and second particle, of energies $\dfrac{n^2\pi^2\hbar^2}{2ma^2}$ and $\dfrac{q^2\pi^2\hbar^2}{2ma^2}$. In the state space of the global system, the basis chosen is composed of the states $|\varphi_n\varphi_q\rangle$ defined by:

$$|\varphi_n\varphi_q\rangle = |\varphi_n(1)\rangle \otimes \varphi_q(2)\rangle$$

a. What are the eigenstates and the eigenvalues of the operator $H = H(1)+H(2)$, the total Hamiltonian of the system? Give the degree of degeneracy of the two lowest energy levels.

b. Assume that the system, at time $t = 0$ is in the state:

$$|\psi(0)\rangle = \frac{1}{\sqrt{6}}|\varphi_1\varphi_1\rangle + \frac{1}{\sqrt{3}}|\varphi_1\varphi_2\rangle + \frac{1}{\sqrt{6}}|\varphi_2\varphi_1\rangle + \frac{1}{\sqrt{3}}|\varphi_2\varphi_2\rangle$$

α. What is the state of the system at time t?

β. The total energy H is measured. What results can be found, and with what probabilities?

γ. Same questions if, instead of measuring H, one measures $H(1)$.

c. α. Show that $|\psi(0)\rangle$ is a tensor product state. When the system is in this state, calculate the following mean values: $\langle H(1)\rangle$, $\langle H(2)\rangle$ and $\langle H(1)\,H(2)\rangle$. Compare $\langle H(1)\rangle\,\langle H(2)\rangle$ with $\langle H(1)\,H(2)\rangle$; how can this result be explained?

β. Show that the preceding results remain valid when the state of the system is the state $|\psi(t)\rangle$ calculated in *b*.

d. Now assume that the state $|\psi(0)\rangle$ is given by:

$$|\psi(0)\rangle = \frac{1}{\sqrt{5}}|\varphi_1\varphi_1\rangle + \sqrt{\frac{3}{5}}|\varphi_1\varphi_2\rangle + \frac{1}{\sqrt{5}}|\varphi_2\varphi_1\rangle$$

Show that $|\psi(0)\rangle$ cannot be put in the form of a tensor product. Answer for this case all the questions asked in *c*.

e. α. Write the matrix, in the basis of the vectors $|\varphi_n\varphi_p\rangle$, that represents the density operator $\rho(0)$ corresponding to the ket $|\psi(0)\rangle$ given in *b*. What is the density matrix $\rho(t)$ at time t? Calculate, at the instant $t = 0$, the partial traces:

$$\rho(1) = \mathrm{Tr}_2\rho \quad \text{and} \quad \rho(2) = \mathrm{Tr}_1\rho$$

Do the density operators ρ, $\rho(1)$ and $\rho(2)$ describe pure states? Compare ρ with $\rho(1) \otimes \rho(2)$; what is your interpretation?

β. Answer the same questions as in α, but choosing for $|\psi(0)\rangle$ the ket given in *d*.

The subject of the following exercises is the density operator: they therefore assume the concepts and results of Complement E_{III} to be known.

17. Let ρ be the density operator of an arbitrary system, where $|\chi_l\rangle$ and π_l are the eigenvectors and eigenvalues of ρ. Write ρ and ρ^2 in terms of the $|\chi_l\rangle$ and π_l. What do the matrices representing these two operators in the $\{|\chi_l\rangle\}$ basis look like – first, in the case where ρ describes a pure state and then, in the case of a statistical mixture of states? (Begin by showing that, in a pure case, ρ has only one non-zero diagonal element, equal to 1, while for a statistical mixture, ρ has several diagonal elements included between 0 and 1.) Show that ρ corresponds to a pure case if and only if the trace of ρ^2 is equal to 1.

18. Consider a system whose density operator is $\rho(t)$, evolving under the influence of a Hamiltonian $H(t)$. Show that the trace of ρ^2 does not vary over time. Conclusion: can the system evolve so as to be successively in a pure state and a statistical mixture of states?

19. Let (1) + (2) be a global system, composed of two subsystems (1) and (2). A and B denote two operators acting in the state space $\mathcal{E}(1) \otimes \mathcal{E}(2)$. Show that the two partial traces $\mathrm{Tr}_1\,\{AB\}$ and $\mathrm{Tr}_1\,\{BA\}$ are equal when A (or B) actually acts only in the space $\mathcal{E}(1)$, that is, when A (or B) can be written:

$$A = A(1) \otimes \mathbb{1}(2) \quad [\text{or} \quad B = B(1) \otimes \mathbb{1}(2)]$$

Application: if the operator H, the Hamiltonian of the global system, is the sum of two operators that act, respectively, only in $\mathcal{E}(1)$ and only in $\mathcal{E}(2)$:

$$H = H(1) + H(2),$$

calculate the variation $\dfrac{\mathrm{d}}{\mathrm{d}t}\rho(1)$ of the reduced density operator $\rho(1)$. Give the physical interpretation of the result obtained.

References

Exercise 5

Flügge (1.24), §§ 40 and 41; Landau and Lifshitz (1.19), § 22.

Exercise 10

Levine (12.3), Chap. 14; Eyring et al. (12.5), § 18 b

Exercise 15

See references of Complement G_{III}.

REVISITING ONE-DIMENSIONAL PROBLEMS

Now that we are more familiar with the mathematical formalism and the physical content of quantum mechanics, we can go in more detail into some of the results obtained in Chapter I. In the three complements that follow, we shall study in a general way the quantum properties of a particle subjected to a scalar potential[1] of arbitrary form, confining ourselves for simplicity to one-dimensional problems. We shall first focus on the bound stationary states of a particle, whose energies form a discrete spectrum (Complement M_{III}), and then treat the unbound states corresponding to an energy continuum (Complement N_{III}). In addition, we shall examine a special case that is very important because of its applications, particularly in solid state physics, that of a periodic potential (Complément O_{III}).

[1] The effects of a vector potentiel $\mathbf{A}$ will be studied later, in particular in Complement E_{VI}.

Complement M_{III}

Bound states in a "potential well" of arbitrary shape

1 Quantization of the bound state energies 359
2 Minimum value of the ground state energy 363

In complement H_I, we studied, for a special case (finite or infinite "square" well), the bound states of a particle in a potential well. We derived certain properties of these bound states: a discrete energy spectrum and a ground state energy greater than the classical minimum energy. These properties are, in fact, general, and have numerous physical consequences, as we shall show in this complement.

When the potential energy of a particle posesses a minimum (see Figure 1a), the particle is said to be placed in a "potential well"[1]. Before studying qualitatively the stationary states of a quantum particle in such a well, let us recall the corresponding motion of a classical particle. When its energy E_{cl} takes on the minimum possible value $E_{cl} = -V_0$. (where V_0 is the depth of the well), the particle is motionless at the point M_0 whose abscissa is x_0. In the case where $-V_0 < E_{cl} < 0$, the particle oscillates in the well, with an amplitude that increases with E_{cl}. Finally, when $E_{cl} > 0$, the particle does not remain in the well, but moves off towards infinity. The "bound states" of the classical particle therefore correspond to all negative energy values between $-V_0$ and 0.

For a quantum particle, the situation is very different. Well-defined energy states E are stationary states whose wave functions $\varphi(x)$ are solutions of the eigenvalue equation of the Hamiltonian H:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\varphi(x) = E\,\varphi(x) \tag{1}$$

Such a second-order differential equation has an infinite number of solutions, whatever the value chosen for E: if we pick arbitrary values of $\varphi(x)$ and its derivative at any given point, we can obtain φ for any other value of x. Equation (1) alone cannot, therefore, restrict the possible energy values. However, we shall show that if, in addition, we impose certain boundary conditions on $\varphi(x)$, only a certain number of values of E remain possible (quantization of energy levels).

1. Quantization of the bound state energies

We shall call "bound states of the particle" states whose wave functions $\varphi(x)$ satisfy the eigenvalue equation (1) and are *square-integrable* [indispensable if $\varphi(x)$ is actually to describe the physical state of a particle]. These are therefore stationary states, for which the position probability density $|\varphi(x)|^2$ takes on non-negligible values only in a limited region of space [for $\int_{-\infty}^{+\infty}\mathrm{d}x|\varphi(x)|^2$ to converge, $|\varphi(x)|^2$ must approach zero sufficiently

[1] The potential energy, of course, is only defined to within a constant. Following the usual convention, we set the potential equal to zero at infinity.

rapidly when $x \longrightarrow \pm\infty$]. Bound states remind us of classical motion where the particle oscillates inside the well without ever being able to emerge (energy E_{cl} negative, but greater than $-V_0$).

We shall see that in quantum mechanics, the fact that $\varphi(x)$ is required to be square-integrable implies that the possible energies form a discrete set of values which are also included between $-V_0$ and 0. To understand this, let us return to the potential shown in Figure 1a. For simplicity, we shall assume that $V(x)$ is identically equal to zero outside an interval $[x_1, x_2]$.

If $x < x_1$ (region I), $V(x) = 0$, and the solution to equation (1) can immediately be written:

– if $E > 0$:

$$\varphi_{\mathrm{I}}(x) = A\,\mathrm{e}^{ikx} + A'\,\mathrm{e}^{-ikx} \tag{2}$$

with:

$$k = \sqrt{\frac{2mE}{\hbar^2}} \tag{3}$$

– if $E < 0$:

$$\varphi_{\mathrm{I}}(x) = B\,\mathrm{e}^{\rho x} + B'\,\mathrm{e}^{-\rho x} \tag{4}$$

with:

$$\rho = \sqrt{-\frac{2mE}{\hbar^2}} \tag{5}$$

We are looking for a square-integrable solution; we must therefore eliminate the form (2) in which $\varphi_{\mathrm{I}}(x)$ is a superposition of plane waves of constant modulus which cause the integral:

$$\int_{-\infty}^{x_1} \mathrm{d}x |\varphi_{\mathrm{I}}(x)|^2 \tag{6}$$

to diverge. Only possibility (4) remains, and we obtain our first result: *the bound states of the particle all have a negative energy.* In (4), we cannot retain the term in $\mathrm{e}^{-\rho x}$, which diverges when $x \longrightarrow -\infty$. We are therefore left with:

$$\varphi_{\mathrm{I}}(x) = \mathrm{e}^{\rho x} \qquad \text{if} \quad x < x_1 \tag{7}$$

[We have omitted the proportionality factor B since the homogeneity of equation (1) allows us to define $\varphi(x)$ to within a multiplicative coefficient].

The value of $\varphi(x)$ in the interval $x_1 \leq x \leq x_2$ (region II) is obtained by extending $\varphi_{\mathrm{I}}(x)$: we must look for the solution of equation (1) which is equal to $\mathrm{e}^{\rho x_1}$ for $x = x_1$ and whose derivative at this point is equal to $\rho\,\mathrm{e}^{\rho x_1}$. The function $\varphi_{\mathrm{II}}(x)$ thus obtained depends on ρ and, of course, on the exact expression for $V(x)$. Nevertheless, since (1) is a second-order differential equation, $\varphi_{\mathrm{II}}(x)$ is determined uniquely by the preceding boundary conditions; it is, moreover, real (which enables us to trace curves such as those in Figures 1b, 1c and 1d).

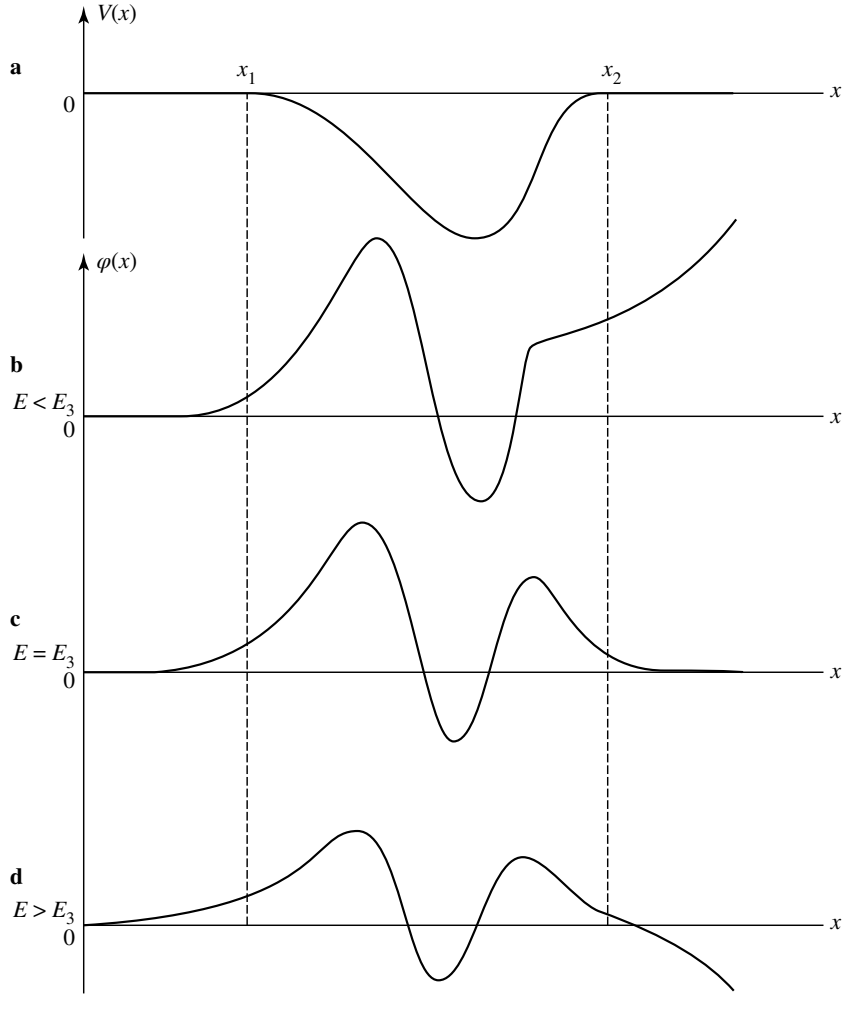


Figure 1: Potential well (fig. a) situated between the points $x = x_1$ and $x = x_2$. We choose a solution $\varphi(x)$ of the eigenvalue equation of H which, for $x < x_1$, approaches zero exponentially when $x \longrightarrow -\infty$. We then extend this solution to the entire x-axis. For an arbitrary energy value E, $\varphi(x)$ diverges like $\tilde{B}(E)\mathrm{e}^{\rho x}$ when $x \longrightarrow +\infty$: figure b represents the case where $\tilde{B}(E) > 0$; figure d, that where $\tilde{B}(E) < 0$. However, if the energy E is chosen so as to make $\tilde{B}(E) = 0$, $\varphi(x)$ approaches zero exponentially when $x \longrightarrow +\infty$ (fig. c), and $\varphi(x)$ is square-integrable.

All that now remains to be done is to obtain the solution when $x > x_2$ (region III); this solution can be written:

$$\varphi_{III}(x) = \tilde{B}e^{\rho x} + \tilde{B}'e^{-\rho x} \tag{8}$$

where $\tilde{B}$ and $\tilde{B}'$ are real constants determined by the two continuity conditions for $\varphi(x)$ and $d\varphi/dx$ at the point $x = x_2$. $\tilde{B}$ and $\tilde{B}'$ depend on ρ, as well as on the function $V(x)$.

We have therefore constructed a solution of equation (1), such as the one shown in Figure 1b. Is this solution square-integrable? We see from (8) that, in general, it is not, except when $\tilde{B}$ is zero (this special case is shown in Figure 1c). Now, for a given function $V(x)$, $\tilde{B}$ is a function of E through the intermediary of ρ. The only values of E for which a bound state exists are therefore solutions of the equation $\tilde{B}(E) = 0$. These solutions E_1, E_2, ... (*cf.* Fig. 2) form a discrete spectrum which, of course, depends on the potential $V(x)$ chosen (we shall see in the following section that all the energies E_i are greater than $-V_0$).

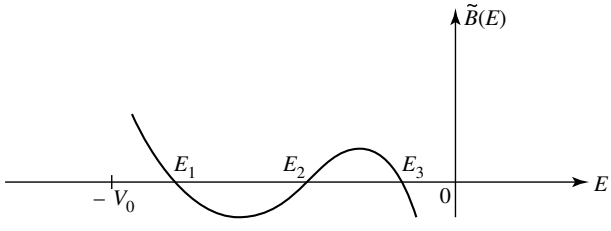


Figure 2: Graphical representation of the function $\tilde{B}(E)$. The zeros of $\tilde{B}(E)$ give the values of E for which $\varphi(x)$ is square-integrable (the situation in Figure 1c), that is, the energies E_1, E_2, E_3, ... of the bound states; all these energies are included between $-V_0$ and 0.

We thus arrive at the following result: *the bound state energy values possible for a particle placed in a potential well of arbitrary shape form a discrete set* (it is often said that the bound state energies are quantized). This result can be compared to the quantization of electromagnetic modes in a cavity. There is no analogue in classical mechanics, where, as we have seen, all energy values included between $-V_0$ and 0 are acceptable. In quantum mechanics, the lowest energy level E_1 is called the *ground state*, the energy level E_2 immediately above, the *first excited state*, the next energy level E_3, the second excited state, etc. The following schematic diagram is often associated with each of these states: inside the potential well representing $V(x)$, a horizontal line is drawn whose vertical position corresponds to the energy of the state and whose length gives an idea of the spatial extension of the wave function (this line actually covers the points of the axis which would be reached by a classical particle of the same energy). For the set of energy levels, we obtain a schematic diagram of the type shown of Figure 3.

As we saw in Chapter I, the phenomenon of energy quantization was one of the factors which led to the introduction of quantum mechanics. Discrete energy levels appear in a very large number of physical systems: atoms (*cf.* Chap. VII, hydrogen atom), the harmonic oscillator (*cf.* Chap. V), atomic nuclei, etc.

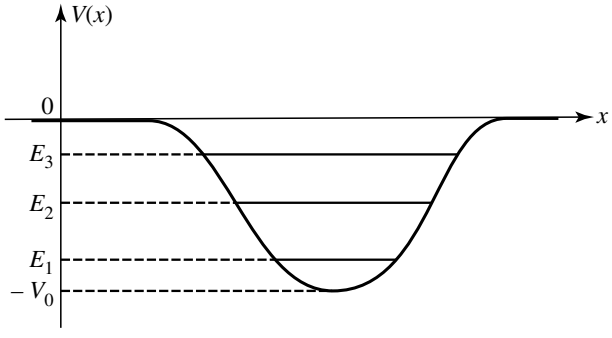


Figure 3: Schematic representation of the bound states of a particle in a potential well. For each of these stationary states, one draws a horizontal line whose ordinate is equal to the energy of the corresponding level. The ends of this line are the points of intersection with the curve representing the potential $V(x)$. The line is confined to the region of classical motion for the same energy, and gives an idea of the extension of the wave function.

2. Minimum value of the ground state energy

We now show that the energies E_1, E_2, etc... are all greater than the minimum value $-V_0$ of the potential energy $V(x)$. We shall see how this result can be easily understood using Heisenberg's uncertainty relation.

If $\varphi(x)$ is a solution of (1), we obtain, multiplying this equation by $\varphi^*(x)$ and integrating the relation thus obtained:

$$-\frac{\hbar^2}{2m}\int_{-\infty}^{+\infty} \mathrm{d}x\ \varphi^*(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi(x) + \int_{-\infty}^{+\infty} \mathrm{d}x\, V(x)\ |\varphi(x)|^2 = E\int_{-\infty}^{+\infty} \mathrm{d}x\ |\varphi(x)|^2 \tag{9}$$

For a bound state, the function $\varphi(x)$ can be normalized, and equation (9) can be written simply:

$$E = \langle T\rangle + \langle V\rangle \tag{10}$$

with:

$$\langle T\rangle = -\frac{\hbar^2}{2m}\int_{-\infty}^{+\infty} \mathrm{d}x\, \varphi^*(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi(x) = \frac{\hbar^2}{2m}\int_{-\infty}^{+\infty} \mathrm{d}x \left|\frac{\mathrm{d}}{\mathrm{d}x}\varphi(x)\right|^2 \tag{11}$$

[where we have performed an integration by parts and used the fact that $\varphi(x)$ goes to zero when $|x| \longrightarrow \infty$] and:

$$\langle V\rangle = \int_{-\infty}^{+\infty} \mathrm{d}x\, V(x)|\varphi(x)|^2 \tag{12}$$

Relation (10) shows simply that E is the sum of the mean value of the kinetic energy:

$$\langle T \rangle = \langle \varphi \left| \frac{P^2}{2m} \right| \varphi \rangle \tag{13}$$

and that of the potential energy:

$$\langle V \rangle = \langle \varphi | V(x) | \varphi \rangle \tag{14}$$

From relations (11) and (12), it follows immediately that:

$$\langle T \rangle > 0 \tag{15}$$

$$\langle V \rangle \geq \int_{-\infty}^{+\infty} \mathrm{d}x \; (-V_0) \; |\varphi(x)|^2 = -V_0 \tag{16}$$

Consequently:

$$E = \langle T \rangle + \langle V \rangle > \langle V \rangle \geq -V_0 \tag{17}$$

Since E is negative, as we showed in § 1, we see that, *as in classical mechanics, the bound state energies are always between* $-V_0$ *and* 0.

There exists, nevertheless, an important difference between the classical and quantum situations: while, in classical mechanics, the particle can have an energy equal to $-V_0$ (case of a particle at rest at M_0) or slightly greater than $-V_0$ (case of small oscillations), the same is not true in quantum mechanics, where the lowest possible energy is the energy E_1 of the ground state, which is necessarily greater than $-V_0$ (*cf.* Fig. 3). The Heisenberg uncertainty relations enable us to understand the physical origin of this result, as we now show.

If we try to construct a state of the particle for which the mean potential energy is as small as possible, we see from (12) that we must choose a wave function which is practically localized at the point M_0. The root mean square deviation ΔX is then very small, so ΔP is necessarily very large. Since:

$$\langle P^2 \rangle = (\Delta P)^2 + \langle P \rangle^2 \geq (\Delta P)^2 \tag{18}$$

the kinetic energy $\langle T \rangle = \langle P^2 \rangle / 2m$ is thus also very large. Therefore, if the potential energy of the particle approaches its minimum, the kinetic energy increases without bound. The wave function of the ground state corresponds to a compromise, for which the sum of these two energies is a minimum. The ground state of the quantum particle is thus characterized by a wave function that has a certain spatial extension (*cf.* Fig. 3), and its energy is necessarily greater than $-V_0$. Unlike the situation in classical mechanics, there exists no well-defined energy state in quantum mechanics where the particle is "at rest" at the bottom of the potential well.

Comment:

Since the energy of the bound states is included between $-V_0$ and 0, such states can exist only if the potential $V(x)$ takes on negative values in one or several regions of the x-axis. This is why we have chosen for this complement a potential

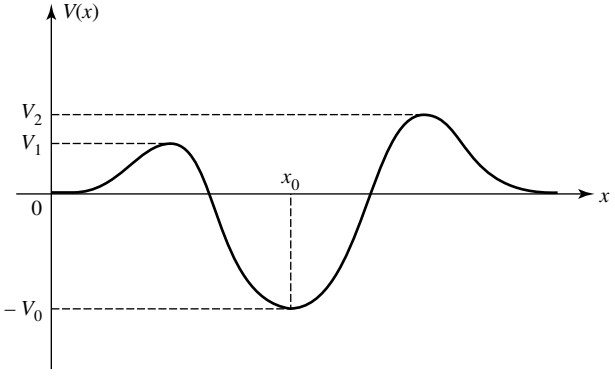


Figure 4: Potential well of depth $-V_0$ situated between two potential barriers of height V_1 and V_2 (assuming, for example, $V_1 \leq V_2$). Classically, there exist particle states whose energy is between $-V_0$ and V_1 that remain confined between the two barriers. In quantum mechanics, a particle whose energy is between 0 and V_1 can penetrate the barrier by the tunnel effect (§ 2-b-β of Complement H_I); consequently, the bound states always have energies between $-V_0$ and 0.

"well" like the one shown in Figure 1a (while in the following complement, we shall not confine ourselves to the case of a potential well).

However, there is nothing to prevent $V(x)$ from being positive for certain values of x; for example, the "well" can be surrounded by potential "barriers" as is shown in Figure 4 (we shall always assume the potential to be zero at infinity). In this case, certain classical motions of positive energy remain bounded, while in quantum mechanics, the same reasoning as above shows that the bound states always have an energy between $-V_0$ and 0. Physically, this difference arises from the fact that a potential barrier of finite height is never able to make a quantum particle turn back completely: the particle always has a non-zero probability of passing through by the tunnel effect.

References and suggestions for further reading:

Feynman III (1.2), § 16-6; Messiah (1.17), Chap. III, § II; Ayant and Belorizky (1.10), Chap. IV, §§ 1, 2, 3; Schiff (1.18), § 8.

Complement N_{III}

Unbound states of a particle in the presence of a potential well or barrier of arbitrary shape

1 Transmission matrix $M(k)$ 368
1-a Definition of $M(k)$ 368
1-b Properties of $M(k)$ 370
2 Transmission and reflection coefficients 372
3 Example 373

In complement M_{III}, we showed that bound states of a particle placed in a potential $V(x)$ have negative energies[1] and that they exist only if $V(x)$ is an attractive potential (a potential well which allows classical bounded motion). We had to reject positive energy values since they led to eigenfunctions $\varphi_k(x)$ of the Hamiltonian H which, at infinity, behaved like superpositions of non square-integrable exponentials $e^{\pm ikx}$. Nevertheless, we saw as early as Chapter I, that, by superposing such functions linearly, one can construct square-integrable wave functions $\psi(x)$ (wave packets) which can therefore represent the physical state of a particle. It is clear that, since the states thus obtained involve several values of k (that is, of the energy), they are no longer stationary states; the wave function $\psi(x)$ therefore evolves over time, propagating and becoming deformed. However, the fact that $\psi(x)$ is already expanded in terms of the eigenfunctions $\varphi_k(x)$ enables us to calculate this evolution very simply [as we did, for example, in complement J_I, where we used the properties of the $\varphi_k(x)$ to calculate the transmission and reflection coefficients of a potential barrier, the delay upon reflection, etc.]. This is why, despite the fact that each of the $\varphi_k(x)$ cannot alone represent a physical state, it is useful to study the positive energy eigenfunctions[2] of H, as we have already done, in complement H_I, for certain square potentials.

In this complement, we are going to study in a general way (confining ourselves, nevertheless, to one-dimensional problems) the effect of a potential $V(x)$ on the positive energy eigenfunctions $\varphi_k(x)$. We shall assume nothing about the shape of $V(x)$, which may present one or several barriers, wells, etc., except that $V(x)$ goes to zero outside a finite interval $[x_1, x_2]$ of the x-axis. We shall show that, in all cases, the effect of $V(x)$ on the functions $\varphi_k(x)$ can be described by a 2×2 matrix, $M(k)$, which possesses a certain number of general properties. We shall thus obtain various results that are independent of the shape of the potential $V(x)$ chosen. For example, we shall see that the transmission and reflection coefficients of a barrier (whether symmetrical or not) are the same for a particle coming from the left and for a particle of the same energy coming from the right. An additional aim of this complement N_{III} is to serve as the point of

[1] Recall that we chose the energy origin so as to make $V(x)$ zero at infinity.

[2] One might also consider studying the non square-integrable negative energy eigenfunctions of H (those whose energies do not belong to the discrete spectrum obtained in complement M_{III}). However, these functions diverge very rapidly (exponentially) at infinity, and one could not obtain square-integrable wave functions by linearly superposing them.

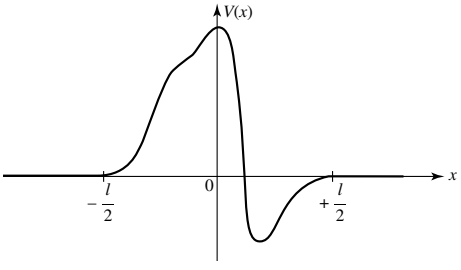


Figure 1: The potential $V(x)$ under consideration varies in an arbitrary way within the interval $-l/2 \leq x \leq l/2$ and goes to zero outside this interval.

departure for the calculations of complement O_{III}, in which we study the properties of a particle in a periodic potential $V(x)$.

1. Transmission matrix $M(k)$

1-a. Definition of $M(k)$

In a one-dimensional problem, consider a potential $V(x)$ which is zero outside an interval $[x_1,\ x_2]$ of length l, but which varies in an arbitrary way inside this interval (Fig. 1). We choose the x origin to be in the middle of the interval $[x_1,\ x_2]$, so as to have $V(x)$ vary only for $|x| < l/2$. The equation satisfied by every wave function $\varphi(x)$ associated with a stationary state of energy E is:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{2m}{\hbar^2}\left[E - V(x)\right]\right\}\varphi(x) = 0 \tag{1}$$

In the rest of this complement, we shall choose, to characterize the energy, the parameter k given by:

$$k = \sqrt{\frac{2mE}{\hbar^2}} \tag{2}$$

In the region $x < -l/2$, the function e^{ikx} satisfies equation (1); let us call $v_k(x)$ the solution of this equation that is identical to e^{ikx} for $x < -\dfrac{l}{2}$. When $x > +\dfrac{l}{2}$, $v_k(x)$ is necessarily a linear combination of two independent solutions e^{ikx} and e^{-ikx} of (1). This gives us:

$$\begin{cases} \text{if} \quad x < -\dfrac{l}{2}: & v_k(x) = \mathrm{e}^{ikx} \qquad (3a) \\ \text{if} \quad x > +\dfrac{l}{2}: & v_k(x) = F(k)\,\mathrm{e}^{ikx} + G(k)\,\mathrm{e}^{-ikx} \qquad (3b) \end{cases}$$

where $F(k)$ and $G(k)$ are coefficients which depend on k, as well as on the shape of the potential under study. Similarly, we can introduce the solution $v'_k(x)$, which, for $x < -l/2$, is equal to e^{-ikx}:

$$\begin{cases} \text{if} \quad x < -\dfrac{l}{2}: & v'_k(x) = \mathrm{e}^{-ikx} \qquad (4a) \\ \text{if} \quad x > +\dfrac{l}{2}: & v'_k(x) = F'(k)\,\mathrm{e}^{ikx} + G'(k)\,\mathrm{e}^{-ikx} \qquad (4b) \end{cases}$$

The most general solution $\varphi_k(x)$ of equation (1) (of second order in x), for a given value of E (that is, of k), is a linear combination of v_k and v'_k:

$$\varphi_k(x) = A\, v_k(x) + A'\, v'_k(x) \tag{5}$$

Relations (3a) and (4a) imply that:

$$\text{if} \quad x < -\frac{l}{2} : \qquad \varphi_k(x) = A\, \mathrm{e}^{ikx} + A'\, \mathrm{e}^{-ikx} \tag{6a}$$

while relations (3b) and (4b) yield:

$$\text{if} \quad x > +\frac{l}{2} : \qquad \varphi_k(x) = \tilde{A}\, \mathrm{e}^{ikx} + \tilde{A}'\, \mathrm{e}^{-ikx} \tag{6b}$$

with:

$$\begin{aligned} \tilde{A} &= F(k)\, A + F'(k)\, A' \\ \tilde{A}' &= G(k)\, A + G'(k)\, A' \end{aligned} \tag{7}$$

By definition, the matrix $M(k)$ is the 2×2 matrix:

$$M(k) = \begin{pmatrix} F(k) & F'(k) \\ G(k) & G'(k) \end{pmatrix} \tag{8}$$

which allows us to write relations (7) in the matrix form:

$$\begin{pmatrix} \tilde{A} \\ \tilde{A}' \end{pmatrix} = M(k) \begin{pmatrix} A \\ A' \end{pmatrix} \tag{9}$$

$M(k)$ therefore enables us to determine, given the behavior (6a) of the wave function to the left of the potential, its behavior (6b) to the right. We call $M(k)$ the "transmission matrix" of the potential.

Comment:

The current associated with a wave function $\varphi(x)$ is:

$$J(x) = \frac{\hbar}{2mi} \left[\varphi^*(x) \frac{\mathrm{d}\varphi}{\mathrm{d}x} - \varphi(x) \frac{\mathrm{d}\varphi^*}{\mathrm{d}x} \right] \tag{10}$$

Differentiating, we find:

$$\frac{\mathrm{d}}{\mathrm{d}x} J(x) = \frac{\hbar}{2mi} \left[\varphi^*(x) \frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} - \varphi(x) \frac{\mathrm{d}^2\varphi^*}{\mathrm{d}x^2} \right] \tag{11}$$

Taking (1) into account, we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}x} J(x) = 0 \tag{12}$$

Therefore, the current $J(x)$ associated with a stationary state is the same at all points of the x-axis. Note, moreover, that (12) is simply the one-dimensional analogue of the relation:

$$\text{div } \mathbf{J}(\mathbf{r}) = 0 \tag{13}$$

which is valid, according to relation (D-11) of Chapter III, for any stationary state of a particle moving in three-dimensional space. According to (12), the current $J_k(x)$ associated with $\varphi_k(x)$ can therefore be calculated for any x, choosing either the form (6a) or the form (6b) of $\varphi_k(x)$:

$$J_k(x) = \frac{\hbar k}{m}\left[|A|^2 - |A'|^2\right] = \frac{\hbar k}{m}\left[|\tilde{A}|^2 - |\tilde{A}'|^2\right] \tag{14}$$

1-b. Properties of $M(k)$

α. It is easy to show, using the fact that the function $V(x)$ is real, that if $\varphi(x)$ is a solution of equation (1), so is $\varphi^*(x)$. Now consider the function $v_k^*(x)$, which is a solution of (1); comparison of (3a) and (4a) shows that it is identical to $v_k'(x)$ when $x < -\dfrac{l}{2}$. We therefore have, for all x:

$$v_k^*(x) = v_k'(x) \tag{15}$$

Substituting relations (3b) and (4b) into this relation, we obtain:

$$F^*(k) = G'(k) \tag{16}$$
$$G^*(k) = F'(k) \tag{17}$$

It follows that the matrix $M(k)$ can be written in the simplified form:

$$M(k) = \begin{pmatrix} F(k) & G^*(k) \\ G(k) & F^*(k) \end{pmatrix} \tag{18}$$

β. We saw above [*cf.* (12)] that the probability current $J(x)$ does not depend on x for a stationary state. We must therefore have [*cf.* (14)]:

$$|A|^2 - |A'|^2 = |\tilde{A}|^2 - |\tilde{A}'|^2 \tag{19}$$

for any A and A'. Now relations (9) and (18) yield:

$$\begin{aligned} |\tilde{A}|^2 - |\tilde{A}'|^2 &= [F(k)A + G^*(k)A']\,[F^*(k)A^* + G(k)A'^*] \\ &\quad - [G(k)A + F^*(k)A']\,[G^*(k)A^* + F(k)A'^*] \\ &= \left[|F(k)|^2 - |G(k)|^2\right]\left[|A|^2 - |A'|^2\right] \end{aligned} \tag{20}$$

Condition (19) is therefore equivalent to:

$$|F(k)|^2 - |G(k)|^2 = \text{Det}\, M(k) = 1 \tag{21}$$

Comments:

(*i*) We have made no particular assumptions about the shape of the potential. If it is even, that is, if $V(x) = V(-x)$, the matrix $M(k)$ possesses an additional property: it can be shown that $G(k)$ is a pure imaginary.

(*ii*) Relations (6) show that A and $\tilde{A}'$ are the coefficients of "incoming" plane waves, i.e. waves associated with particles arriving respectively from $x = -\infty$ and $x = +\infty$ and moving towards the zone of influence of the potential (incident particles). On the other hand, $\tilde{A}$ and A' are the coefficients corresponding to "outgoing" waves, associated with particles moving away from the potential (transmitted or reflected particles). It is useful to introduce the matrix S, which allows us to calculate the amplitude of the outgoing waves in terms of that of the incoming waves:

$$\begin{pmatrix} \tilde{A} \\ A' \end{pmatrix} = S(k) \begin{pmatrix} A \\ \tilde{A}' \end{pmatrix} \tag{22}$$

$S(k)$ can easily be expressed in terms of the elements of the matrix $M(k)$, as we now show. The relations:

$$\tilde{A} = F(k)\, A + G^*(k)\, A' \tag{23a}$$

$$\tilde{A}' = G(k)\, A + F^*(k)\, A' \tag{23b}$$

imply that:

$$A' = \frac{1}{F^*(k)} \left[\tilde{A}' - G(k)\, A\right] \tag{24}$$

Substituting this relation into (23a), we obtain:

$$\tilde{A} = \frac{1}{F^*(k)} \left[\Big(F(k)F^*(k) - G(k)G^*(k)\Big) A + G^*(k)\tilde{A}'\right] \tag{25}$$

Taking (21) into account, we can then write the matrix $S(k)$:

$$S(k) = \frac{1}{F^*(k)} \begin{pmatrix} 1 & G^*(k) \\ -G(k) & 1 \end{pmatrix} \tag{26}$$

It is easy to verify, using (21) again, that:

$$S(k)\, S^\dagger(k) = S^\dagger(k)\, S(k) = 1 \tag{27}$$

$S(k)$ is therefore unitary. This matrix plays an important role in collision theory; we could have proved its unitary property from that of the evolution operator (*cf.* Complement $\mathrm{F_{III}}$), which simply expresses the conservation over time of the total probability of finding the particle somewhere on the Ox axis (norm of the wave function).

2. Transmission and reflection coefficients

To calculate the reflection and transmission coefficients for a particle encountering the potential $V(x)$, one should (as in complement J_I) construct a wave packet with the eigenfunctions of H which we have just studied. Consider, for example, an incident particle of energy E_i coming from the left. The corresponding wave packet is obtained by superposing functions $\varphi_k(x)$, for which we set $\tilde{A}' = 0$, with coefficients given by a function $g(k)$ which has a marked peak in the neighborhood of $k = k_i = \sqrt{2mE_i/\hbar^2}$. We shall not go into these calculations in detail here; they are analogous in every way to those of complement J_I. They show that the reflection and transmission coefficients are equal, respectively, to $|A'(k_i)/A(k_i)|^2$ and $|\tilde{A}(k_i)/A(k_i)|^2$.

Since $\tilde{A}' = 0$, relations (22) and (26) yield:

$$\begin{aligned} \tilde{A}(k) &= \frac{1}{F^*(k)}\, A(k) \\ A'(k) &= -\frac{G(k)}{F^*(k)}\, A(k) \end{aligned} \tag{28}$$

The reflection and transmission coefficients are therefore equal to:

$$R_1(k_i) = \left|\frac{A'(k_i)}{A(k_i)}\right|^2 = \left|\frac{G(k_i)}{F(k_i)}\right|^2 \tag{29a}$$

$$T_1(k_i) = \left|\frac{\tilde{A}(k_i)}{A(k_i)}\right|^2 = \frac{1}{|F(k_i)|^2} \tag{29b}$$

[it is easy to verify that condition (21) insures that $R_1(k_i) + T_1(k_i) = 1$].

If we now consider a particle coming from the right, we must take $A = 0$, which gives:

$$\begin{aligned} \tilde{A}(k) &= \frac{G^*(k)}{F^*(k)} \tilde{A}'(k) \\ A'(k) &= \frac{1}{F^*(k)}\, \tilde{A}'(k) \end{aligned} \tag{30}$$

The transmission and reflection coefficients are now equal to:

$$T_2(k_i) = \left|\frac{A'(k_i)}{\tilde{A}'(k_i)}\right|^2 = \frac{1}{|F(k_i)|^2} \tag{31a}$$

and:

$$R_2(k_i) = \left|\frac{\tilde{A}(k_i)}{\tilde{A}'(k_i)}\right|^2 = \left|\frac{G(k_i)}{F(k_i)}\right|^2 \tag{31b}$$

Comparison of (29) and (31) shows that $T_1(k_i) = T_2(k_i)$ and that $R_1(k_i) = R_2(k_i)$: for a given energy, the transparency of a barrier (whether symmetrical or not) is therefore always the same for particles coming from the right and from the left.

In addition, from (21) we have:

$$|F(k)| \geq 1 \tag{32}$$

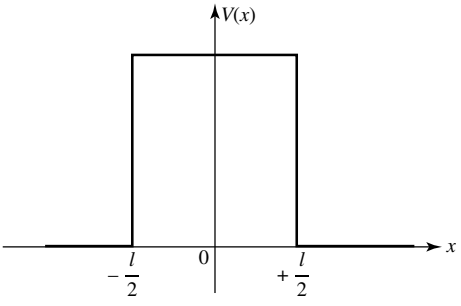


Figure 2: Square potential barrier.

If (32) becomes an equality, the reflection coefficient is zero and the transmission coefficient is equal to 1 (resonance). On the other hand, the inverse situation is not possible: since (21) imposes that $|F(k)| > |G(k)|$, one can never have $T = 0$ and $R = 1$ [except in the case where F and G tend simultaneously towards infinity]. Actually, such a situation can only occur for $k = 0$. To see this, divide the function $v_k(x)$ defined in (3) by $F(k)$. If $F(k)$ goes to infinity, the wave function will be identically zero on the left hand side, and hence necessarily, by extension, zero on the right hand side. However, this is impossible unless $k = 0$ and $F = -G$.

3. Example

Let us return to the square potentials studied in § 2-b of complement $\mathrm{H_I}$: in the region $-\frac{l}{2} < x < \frac{l}{2}$, $V(x)$ is equal[3] to a constant V_0 (see Figure 2, where V_0 has been chosen to be positive).

First, let us assume that E is smaller than V_0, and set:

$$\rho = \sqrt{\frac{2m}{\hbar^2}(V_0 - E)} \tag{33}$$

An elementary calculation analogous to the one in complement $\mathrm{H_I}$ yields:

$$M(k) = \begin{pmatrix} \left[\cosh\ \rho l + i\frac{k^2 - \rho^2}{2k\rho}\sinh\ \rho l\right] \mathrm{e}^{-ikl} & -i\frac{k_0^2}{2k\rho}\sinh\ \rho l \\ i\frac{k_0^2}{2k\rho}\sinh\ \rho l & \left[\cosh\ \rho l - i\frac{k^2 - \rho^2}{2k\rho}\sinh\ \rho l\right] \mathrm{e}^{ikl} \end{pmatrix} \tag{34}$$

with:

$$k_0 = \sqrt{\frac{2mV_0}{\hbar^2}} \tag{35}$$

(V_0 is necessarily positive here, since we have assumed $E < V_0$).

[3] In fact, we are considering here a barrier that is displaced relative to that of complement $\mathrm{H_I}$, since we are assuming it to be situated between $x = -l/2$ and $x = +l/2$, instead of between $x = 0$ and $x = l$.

If now we assume that $E > V_0$, we set:

$$k' = \sqrt{\frac{2m}{\hbar^2}(E - V_0)} \tag{36}$$

and:

$$k_0 = \sqrt{\varepsilon \frac{2mV_0}{\hbar^2}} \tag{37}$$

(where $\varepsilon = +1$ if $V_0 > 0$ and -1 if $V_0 < 0$). We thus obtain:

$$M(k) = \begin{pmatrix} \left[\cos\, k'l + i\frac{k^2 + k'^2}{2kk'} \sin\, k'l\right] \mathrm{e}^{-ikl} & -i\varepsilon \frac{k_0^2}{2kk'} \sin\, k'l \\ i\varepsilon \frac{k_0^2}{2kk'} \sin\, k'l & \left[\cos\, k'l - i\frac{k^2 - k'^2}{2kk'} \sin\, k'l\right] \mathrm{e}^{ikl} \end{pmatrix} \tag{38}$$

It is easy to verify that the matrices $M(k)$ written in (34) and (38) satisfy relations (16), (17) and (21).

References and suggestions for further reading:

Merzbacher (1.16), Chap. 6, §§ 5, 6 and 8; see also the references of complement M_{III}.

Complement O_{III}

Quantum properties of a particle in a one-dimensional periodic structure

1 Passage through several successive identical potential barriers 376

1-a Notation 377

1-b Matching conditions 377

1-c Iteration matrix $Q(\alpha)$ 379

1-d Eigenvalues of $Q(\alpha)$ 380

2 Discussion: the concept of an allowed or forbidden energy band 381

2-a Behavior of the wave function $\varphi_\alpha(x)$ 381

2-b Bragg reflection; possible energies for a particle in a periodic potential 382

3 Quantization of energy levels in a periodic potential; effect of boundary conditions 383

3-a Conditions imposed on the wave function 384

3-b Allowed energy bands: stationary states of the particle inside the lattice 386

3-c Forbidden bands: stationary states localized on the edges . . 390

In this complement, we are going to study the quantum properties of a particle placed in a potential $V(x)$ having a periodic structure. The functions $V(x)$ which we shall consider will not necessarily be periodic in the strict sense of the term; it suffices for them to have the shape of a periodic function in a finite region of the x-axis (Fig. 1), that is, to be the result of juxtaposing N times the same motif at regular intervals [$V(x)$ is truly periodic only in the limit $N \to \infty$].

Such periodic structures are encountered, for example, in the study of a linear molecule formed by N atoms (or groups of atoms) which are identical and equally spaced. They are also encountered in solid state physics, when one chooses a one-dimensional model in order to understand the disposition of the energy levels of an electron in a crystal. If N is very large (as in the case of a linear macromolecule or a macroscopic crystal), the potential $V(x)$ is given in a wide region of space by a periodic function, and the properties of the particle can be expected to be practically the same as they would be if $V(x)$ were really periodic. However, from a physical point of view, the limit of infinite N is never attained, and we shall be concerned here with the case where N is arbitrary.

To study the effect of the potential $V(x)$ on an eigenfunction $\varphi(x)$ of the Hamiltonian H, of eigenvalue E, we shall introduce a 2×2 matrix, the iteration matrix Q, which depends on E. We shall show that the behavior of $\varphi(x)$ is totally different depending on whether the eigenvalues of the iteration matrix are real or imaginary. Since these eigenvalues depend on the energy E chosen, we shall find it useful to distinguish between

domains of energy corresponding to real eigenvalues and those which lead to imaginary eigenvalues. The concept of an *allowed or forbidden energy band will thus be introduced.*

Comments:

(*i*) For the sake of convenience, we shall speak of a "potential barrier" to designate the motif which, repeated N times, gives the potential $V(x)$ (Fig. 1). However, this motif can also be a "potential well" or have an arbitrary shape.

(*ii*) Common usage in solid state physics reserves the letter k to designate a parameter that appears in the expression of the stationary wave functions, and which is not simply proportional to the square root of the energy. To conform to this usage, we shall henceforth use a notation slightly different from that of Complement N_{III}; we shall replace k by α, setting:

$$\alpha = \sqrt{\frac{2mE}{\hbar^2}} \tag{1}$$

and we shall not introduce the letter k until later (we shall see that k is directly related to the eigenvalues of the matrix Q when they are complex).

1. Passage through several successive identical potential barriers

Consider a potential $V(x)$ which is obtained by juxtaposing N barriers as in Figure 1: the first barrier is centered at $x = 0$, the second, at $x = l$, the third, at $x = 2l$, ..., the last at $x = (N-1)l$. We intend to study the behavior, during passage through this set of barriers, of an eigenfunction $\varphi_\alpha(x)$ which is a solution of the eigenvalue equation of H:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{2m}{\hbar^2}[E - V(x)]\right\}\varphi_\alpha(x) = 0 \tag{2}$$

where E and α are related by (1).

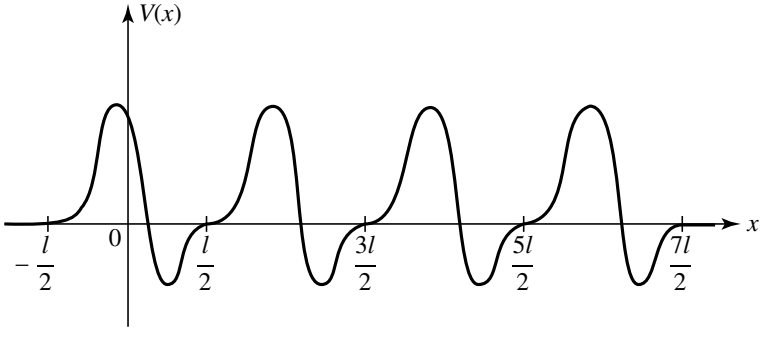


Figure 1: Potential $V(x)$ having a periodic structure obtained by juxtaposing N times the same motif ($N = 4$ in the figure).

1-a. Notation

To the left of the N barriers, that is, for $x \leq -\frac{l}{2}$, $V(x)$ is zero, and the general solution of equation (2) is:

$$\text{if} \quad x \leq -\frac{l}{2} \;:\; \varphi_\alpha(x) = A_0\, \mathrm{e}^{i\alpha x} + A'_0\, \mathrm{e}^{-i\alpha x} \tag{3a}$$

Consider, as in § 1-a of Complement $\mathrm{N_{III}}$, the two functions $v_k(x)$ and $v'_k(x)$ which here become $v_\alpha(x)$ and $v'_\alpha(x)$. In the region of the first barrier, centered at $x = 0$, the general solution of (2) is written:

$$\text{if} \quad -\frac{l}{2} \leq x \leq \frac{l}{2} \;:\; \varphi_\alpha(x) = A_1\, v_\alpha(x) + A'_1\, v'_\alpha(x) \tag{3b}$$

Similarly, in the region of the second barrier, centered at $x = l$, we obtain:

$$\text{if} \quad \frac{l}{2} \leq x \leq \frac{3l}{2} \;:\; \varphi_\alpha(x) = A_2\, v_\alpha(x-l) + A'_2\, v'_\alpha(x-l) \tag{3c}$$

and, more generally, in the region of the nth barrier, centered at $x = (n-1)l$:

$$\text{if} \quad (n-1)l - \frac{l}{2} \leq x \leq (n-1)l + \frac{l}{2} \;:\; \varphi_\alpha(x) = A_n\, v_\alpha\left[x-(n-1)l\right] + A'_n\, v'_\alpha\left[x-(n-1)l\right] \tag{3d}$$

Finally, to the right of the N barriers, that is, for $x \geq (N-1)l + \frac{l}{2}$, $V(x)$ is again zero, and we have:

$$\text{if} \quad x \geq (N-1)l + \frac{l}{2} \;:\; \varphi_\alpha(x) = C_0\, \mathrm{e}^{i\alpha[x-(N-1)l]} + C'_0\, \mathrm{e}^{-i\alpha[x-(N-1)l]} \tag{3e}$$

We must now match these various expressions for the wave function $\varphi_\alpha(x)$ at $x = -\frac{l}{2}, +\frac{l}{2}$, ..., $(N-1)l + \frac{l}{2}$. This is what we shall do in the following section.

1-b. Matching conditions

The functions v_α and v'_α depend on the form of the potential chosen. We shall show, however, that it is simple to calculate them, and their derivatives as well, at the two edges of each barrier, by using the results of Complement $\mathrm{N_{III}}$.

To do so, let us imagine that all but one of the barriers are removed, leaving, for example, the nth one, centered at $x = (n-1)l$. Solution (3d), always valid inside this barrier, must then be extended to the left and to the right by superposing plane waves. These waves are obtained by replacing, in formulas (6a) and (6b) of $\mathrm{N_{III}}$, x by $x-(n-1)l$ and k by α, and adding an index n to A, A', $\tilde{A}$, $\tilde{A}'$. Thus we have, if the nth barrier is isolated:

for $x \leq (n-1)l - \frac{l}{2}$:

$$A_n\, \mathrm{e}^{i\alpha[x-(n-1)l]} + A'_n\, \mathrm{e}^{-i\alpha[x-(n-1)l]} \tag{4}$$

for $x \geq (n-1)l + \dfrac{l}{2}$:

$$\tilde{A}_n \, \mathrm{e}^{i\alpha[x-(n-1)l]} + \tilde{A}'_n \, \mathrm{e}^{-i\alpha[x-(n-1)l]} \tag{5}$$

with:

$$\begin{pmatrix} \tilde{A}_n \\ \tilde{A}'_n \end{pmatrix} = M(\alpha) \begin{pmatrix} A_n \\ A'_n \end{pmatrix} \tag{6}$$

where, with the change in notation taken into account, $M(\alpha)$ is the matrix $M(k)$ introduced in Complement N_{III}. Consequently, at the left edge of the nth barrier, the function $\varphi_\alpha(x)$ defined in (3d) has the same value and the same derivative as the superposition of plane waves (4). Similarly, at the right edge of this barrier, it has the same value and the same derivative as (5). These results enable us to write simply the matching conditions in the periodic structure.

At the left edge of the first barrier (that is, at $x = -l/2$), it is sufficient to note that (3a) has the same value and the same derivative as $A_1 \, \mathrm{e}^{i\alpha x} + A'_1 \, \mathrm{e}^{-i\alpha x}$, which yields directly:

$$\begin{cases} A_0 = A_1 \\ A'_0 = A'_1 \end{cases} \tag{7}$$

(a result which was obvious from N_{III}).

At the right edge of the first barrier, which is the same as the left edge of the second one, we must write that $\tilde{A}_1 \, \mathrm{e}^{i\alpha x} + \tilde{A}'_1 \, \mathrm{e}^{-i\alpha x}$ and $A_2 \, \mathrm{e}^{i\alpha(x-l)} + A'_2 \, \mathrm{e}^{-i\alpha(x-l)}$ have the same value and the same derivative, which yields:

$$\begin{cases} A_2 = \tilde{A}_1 \, \mathrm{e}^{i\alpha l} \\ A'_2 = \tilde{A}'_1 \, \mathrm{e}^{-i\alpha l} \end{cases} \tag{8}$$

Similarly, at the junction of the nth and $(n+1)$th barriers $\left(x = nl - \dfrac{l}{2}\right)$, we obtain, setting equal the value and derivative of (5) and those of the expression obtained by replacing n by $n+1$ in (4):

$$\begin{cases} A_{n+1} = \tilde{A}_n \, \mathrm{e}^{i\alpha l} \\ A'_{n+1} = \tilde{A}'_n \, \mathrm{e}^{-i\alpha l} \end{cases} \tag{9}$$

Finally, at the right edge of the last barrier $\left(x = (N-1)l + \dfrac{l}{2}\right)$, we must write that (3e) has the same value and the same derivative as the expression obtained by replacing n by N in (5), which yields:

$$\begin{cases} C_0 = \tilde{A}_N \\ C'_0 = \tilde{A}'_N \end{cases} \tag{10}$$

1-c. Iteration matrix $Q(\alpha)$

Let us introduce the matrix $D(\alpha)$ defined by:

$$D(\alpha) = \begin{pmatrix} \mathrm{e}^{i\alpha l} & 0 \\ 0 & \mathrm{e}^{-i\alpha l} \end{pmatrix} \tag{11}$$

It enables us to write the matching condition (9) in the form:

$$\begin{pmatrix} A_{n+1} \\ A'_{n+1} \end{pmatrix} = D(\alpha) \begin{pmatrix} \tilde{A}_n \\ \tilde{A}'_n \end{pmatrix} \tag{12}$$

that is, taking (6) into account:

$$\begin{pmatrix} A_{n+1} \\ A'_{n+1} \end{pmatrix} = D(\alpha)M(\alpha) \begin{pmatrix} A_n \\ A'_n \end{pmatrix} \tag{13}$$

Iterating this equation and using (7), we obtain:

$$\begin{aligned} \begin{pmatrix} A_{n+1} \\ A_{n+1} \end{pmatrix} &= [D(\alpha)M(\alpha)]^n \begin{pmatrix} A_1 \\ A'_1 \end{pmatrix} \\ &= [D(\alpha)M(\alpha)]^n \begin{pmatrix} A_0 \\ A'_0 \end{pmatrix} \end{aligned} \tag{14}$$

Finally, the matching condition (10) can be transformed by using (6) and (14):

$$\begin{pmatrix} C_0 \\ C'_0 \end{pmatrix} = M(\alpha) \begin{pmatrix} A_N \\ A'_N \end{pmatrix} = M(\alpha) \left[D(\alpha)M(\alpha)\right]^{N-1} \begin{pmatrix} A_0 \\ A'_0 \end{pmatrix} \tag{15}$$

that is:

$$\begin{pmatrix} C_0 \\ C'_0 \end{pmatrix} = \underbrace{M(\alpha)D(\alpha)M(\alpha)D(\alpha) \ldots D(\alpha)M(\alpha)}_{N \text{ matrices } M(\alpha)} \begin{pmatrix} A_0 \\ A'_0 \end{pmatrix} \tag{16}$$

In this formula, which enables us to go from $\begin{pmatrix} A_0 \\ A'_0 \end{pmatrix}$ to $\begin{pmatrix} C_0 \\ C'_0 \end{pmatrix}$, a matrix $M(\alpha)$ is associated with each barrier, and another matrix $D(\alpha)$ with each interval between two successive barriers.

Relations (13) and (14) demonstrate the importance of the role played by the matrix:

$$Q(\alpha) = D(\alpha)M(\alpha) \tag{17}$$

which enters to the nth power when one goes from $\begin{pmatrix} A_1 \\ A'_1 \end{pmatrix}$ to $\begin{pmatrix} A_{n+1} \\ A'_{n+1} \end{pmatrix}$, that is, when one performs a translation through a distance nl along the periodic structure. For this

reason, we shall call $Q(\alpha)$ the "iteration matrix". Using formula (18) of Complement N_{III} and expression (11) for $D(\alpha)$, we obtain:

$$Q(\alpha) = \begin{pmatrix} e^{i\alpha l}F(\alpha) & e^{i\alpha l}G^*(\alpha) \\ e^{-i\alpha l}G(\alpha) & e^{-i\alpha l}F^*(\alpha) \end{pmatrix} \tag{18}$$

The calculation of $[Q(\alpha)]^n$ is facilitated if we change bases so as to make $Q(\alpha)$ diagonal; for this reason we shall study the eigenvalues of $Q(\alpha)$.

1-d. Eigenvalues of $Q(\alpha)$

Let λ be an eigenvalue of $Q(\alpha)$. The characteristic equation of the matrix (18) is written:

$$\left[e^{i\alpha l}F(\alpha) - \lambda\right]\left[e^{-i\alpha l}F^*(\alpha) - \lambda\right] - |G(\alpha)|^2 = 0 \tag{19}$$

that is, taking into account relation (21) of Complement N_{III}:

$$\lambda^2 - 2\lambda\, X(\alpha) + 1 = 0 \tag{20}$$

where $X(\alpha)$ is the real part of the complex number $e^{i\alpha l}F(\alpha)$:

$$X(\alpha) = \mathrm{Re}\left[e^{i\alpha l}F(\alpha)\right] = \frac{1}{2}\mathrm{Tr}\, Q(\alpha) \tag{21}$$

Recall [*cf.* Complement N_{III}, relation (21)] that the modulus of $F(\alpha)$ is greater than 1; the same is therefore true of $e^{i\alpha l}F(\alpha)$.

The discriminant of the second-degree equation (20) is:

$$\Delta' = [X(\alpha)]^2 - 1 \tag{22}$$

Two cases may then arise:

(i) If the energy E is such that:

$$|X(\alpha)| \leq 1 \tag{23}$$

(for example, if, in Figure 2, α is between α_0 and α_1), one can set:

$$X(\alpha) = \cos[k(\alpha)l] \tag{24}$$

with:

$$0 \leq k(\alpha) \leq \frac{\pi}{l} \tag{25}$$

A simple calculation then shows that the eigenvalues of $Q(\alpha)$ are given by:

$$\lambda = e^{\pm ik(\alpha)l} \tag{26}$$

There are therefore two eigenvalues, which are complex conjugates and whose modulus is equal to 1.

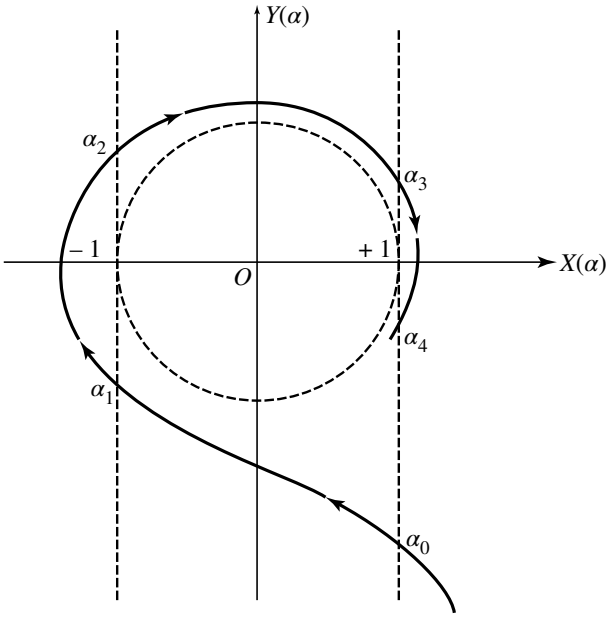


Figure 2: Variation with respect to α of the complex number $e^{i\alpha l}F(\alpha) = X(\alpha) + iY(\alpha)$. Since $|F(\alpha)| > 1$, the curve obtained in the complex plane falls outside the circle of unit radius centered at O. The following discussion shows that if $|X(\alpha)|$ is less than 1, that is, if the value of α chosen gives a point of the curve which is between the two vertical dashed lines of the figure, the corresponding energy falls in an "allowed band"; in the opposite case, it falls in a "forbidden band".

(*ii*) If, on the other hand, the energy E gives a value of α such that:

$$|X(\alpha)| > 1 \tag{27}$$

(for example, if, in Figure 2, α is between α_1 and α_2), one sets:

$$X(\alpha) = \varepsilon \cosh[\rho(\alpha)l] \tag{28}$$

with:

$$\rho(\alpha) \geq 0 \tag{29}$$

and $\varepsilon = +1$ if $X(\alpha)$ is positive, $\varepsilon = -1$ if $X(\alpha)$ is negative. We then find:

$$\lambda = \varepsilon\, e^{\pm\rho(\alpha)l} \tag{30}$$

In this case, both eigenvalues of $Q(\alpha)$ are real, and they are each other's inverse.

2. Discussion: the concept of an allowed or forbidden energy band

2-a. Behavior of the wave function $\varphi_\alpha(x)$

To apply (14), we begin by calculating the two column matrices $\Lambda_1(\alpha)$ and $\Lambda_2(\alpha)$ associated with the eigenvectors of $Q(\alpha)$ and corresponding respectively to the eigenval-

ues λ_1 and λ_2. We then decompose the column matrix $\binom{A_1}{A_1'}$ into the form:

$$\begin{pmatrix} A_1 \\ A_1' \end{pmatrix} = c_1(\alpha)\Lambda_1(\alpha) + c_2(\alpha)\Lambda_2(\alpha) \tag{31}$$

which enables us to obtain directly:

$$\begin{pmatrix} A_n \\ A_n' \end{pmatrix} = \lambda_1^{n-1} c_1(\alpha)\Lambda_1(\alpha) + \lambda_2^{n-1} c_2(\alpha)\Lambda_2(\alpha) \tag{32}$$

It is clear from this expression that the behavior of the wave function is very different depending on whether $|X(\alpha)|$ is smaller or greater than 1 in the energy domain of the wave function. In the first case, formula (26) shows that the effect of traversing successive barriers is expressed in (32) by a phase shift in the components of the column matrix $\binom{A_n}{A_n'}$ onto $\Lambda_1(\alpha)$ and $\Lambda_2(\alpha)$. The behavior of $\varphi_\alpha(x)$ here recalls that of a superposition of imaginary exponentials. On the other hand, if the energy is such that $|X(\alpha)| > 1$, formula (30) indicates that only one of the two eigenvalues (for example, λ_1) has a modulus greater than 1. For n sufficiently large, we have, as a result:

$$\begin{pmatrix} A_n \\ A_n' \end{pmatrix} \simeq \varepsilon^{n-1}\, \mathrm{e}^{(n-1)\rho(\alpha)l}\, c_1(\alpha)\Lambda_1(\alpha) \tag{33}$$

A_n and A_n' therefore increase exponentially with n [except in the special case where $c_1(\alpha) = 0$]; the wave function $\varphi_\alpha(x)$ then increases in modulus as it traverses the successive potential barriers, and its behavior recalls that of a superposition of real exponentials.

2-b. Bragg reflection; possible energies for a particle in a periodic potential

Depending on whether $\varphi_\alpha(x)$ behaves like a superposition of real or imaginary exponentials, the resulting phenomena can reasonably be expected to be different.

Let us evaluate, for example, the transmission coefficient $T_N(\alpha)$ of the set of N identical barriers. For these N barriers, relation (15) shows that the matrix $M(\alpha)[Q(\alpha)]^{N-1}$ plays a role analogous to the one played by $M(\alpha)$ for a single barrier. Now, according to relation (29b) of Complement $\mathrm{N_{III}}$, the transmission coefficient $T(\alpha)$ is expressed in terms of the element of this matrix which is placed in the first row and the first column [the inverse of $T_N(\alpha)$ is equal to the square of the modulus of this element]. What happens if the energy E of the particle is chosen so as to make the eigenvalues of $Q(\alpha)$ real, that is, given by (30)? When N becomes sufficiently large, the eigenvalue $\lambda_1 = \varepsilon\, \mathrm{e}^{\rho(\alpha)l}$ becomes dominant, and the matrix $[Q(\alpha)]^{N-1}$ increases exponentially with N [as can also be seen from relation (33)]. Consequently, the transmission coefficient decreases exponentially:

$$T_N(\alpha) \propto \mathrm{e}^{-2N\rho(\alpha)l} \tag{34}$$

In this case, for large values of N, the set of N potential barriers reflects the particle practically without fail. This is explained by the fact that the waves scattered by the different potential barriers interfere totally destructively for the transmitted wave, and constructively for the reflected wave. This phenomenon can therefore be likened to *Bragg reflection*. Note, moreover, that this destructive interference for the transmitted wave can be produced even if the energy E is greater than the height of the barrier (a case where, in classical mechanics, the particle is transmitted).

Nevertheless, if the transmission coefficient of an isolated barrier is very close to 1, we have $|F(\alpha)| \simeq 1$ [for example, in Figure 2, $|F(\alpha)| \longrightarrow 1$ if α, that is, the energy, approaches infinity]. The point representing the complex number $e^{i\alpha l}F(\alpha)$ is then very close to the circle of unit radius centered at O. Figure 2 shows that the regions of the energy axis where $|X(\alpha)| > 1$, that is, where total reflection occurs, are very narrow and can practically be seen as isolated energy values. Physically, this is explained by the fact that, if the energy E of the incident particle is much larger than the amplitude of variation of the potential $V(x)$, its momentum is well-defined, as is the associated wavelength. The Bragg condition $l = n\dfrac{\lambda}{2}$ (where n is an integer) then gives well-defined energy values.

If, on the other hand, the energy E of the particle falls in a domain where the eigenvalues are of modulus 1 as in (26), the elements of the matrix $[Q(\alpha)]^{N-1}$ no longer approach infinity when N does. Under these conditions, the transmission coefficient $T_N(\alpha)$ does not approach zero when the number of barriers is increased. We are again dealing with a purely quantum mechanical phenomenon, related to the wave-like nature of the wave function, which enables it to propagate in the regular periodic potential structure without being exponentially attenuated. Note especially that the transmission coefficient $T_N(\alpha)$ is very different from the product of the individual transmission coefficients of the barriers taken separately (this product approaches zero when $N \longrightarrow \infty$ since all the factors are smaller than 1).

The quantization of energy levels for a particle placed in a series of identical and evenly spaced potential wells (i.e. a periodic potential $V(x)$) is another interesting problem, particularly in solid state physics. It will be studied in detail in § 3, but we can already guess the form of the spectrum of possible energies. If we assume that the energy of the particle is such that $|X(\alpha)| > 1$, equation (33) shows that the coefficients A_n and A'_n become infinite when $n \longrightarrow \infty$. It is clear that this possibility must be rejected, since it means that the wave function does not remain bounded. The corresponding energies are therefore forbidden; hence, the name of *forbidden bands* given to the energy domains for which $|X(\alpha)| > 1$. On the other hand, if the energy of the particle is such that $|X(\alpha)| < 1$, A_n and A'_n remain bounded when $n \longrightarrow \infty$; the corresponding regions of the energy axis are called *allowed bands*. To sum up, the energy spectrum is composed of finite intervals inside which all the energies are acceptable, separated by regions all of whose energies are forbidden.

3. Quantization of energy levels in a periodic potential; effect of boundary conditions

Consider a particle of mass m placed in the potential $V(x)$ shown in Figure 3. In the region $-\dfrac{l}{2} \leq x \leq Nl + \dfrac{l}{2}$, $V(x)$ has the form of a periodic function, composed of a series of $N+1$ successive barriers of height V_0, centered at $x = 0,\ l,\ 2\,l,\ ...,\ Nl$, Outside this region, $V(x)$ undergoes arbitrary variations over distances comparable to l, then becomes equal to a positive constant value V_e. In what follows, the region $[0, Nl]$ will be called "inside the lattice" and the limiting regions $x \simeq -\dfrac{l}{2}$ and $x \simeq Nl + \dfrac{l}{2}$, "ends (or edges) of the lattice".

Physically, such a function $V(x)$ can represent the potential seen by an electron in a linear molecule or in a crystal (in a one-dimensional model). The potential wells at $x = \dfrac{l}{2},\ \dfrac{3l}{2},\ ...$

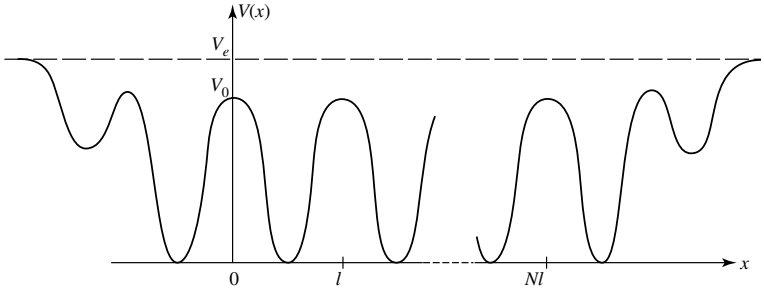


Figure 3: Variation with respect to x of the potential seen by an electron in a "one-dimensional crystal" and on its edges. Inside the crystal, the potential has a periodic structure; $V(x)$ is maximum between the ions (barriers at $x = 0$, l, $2l$, ...) and minimum at the positions of the ions (wells at $x = l/2$, $3l/2$, ...). On the edges of the crystal, $V(x)$ varies in a more or less complicated way over a distance comparable to l, then rapidly approaches a constant value V_e.

then correspond to the attraction of the electron by the various ions. Far from the crystal (or the molecule), the electron is not subjected to any attractive forces, which is why $V(x)$ rapidly becomes constant outside the region $-\frac{l}{2} \leq x \leq Nl + \frac{l}{2}$.

The potential $V(x)$ that we have chosen fits perfectly into the framework of Complement M_{III} (apart from a change in the energy origin). We already know, therefore, that the bound states of the particle form a discrete spectrum of energies, all less than V_e. However, the potential $V(x)$ picked here also presents the remarkable peculiarity of having a periodic structure of the type of those considered in § 1 above; relying on the results of this section, we shall show that the conclusions of Complement M_{III} take on a special form in this case. For example, in Complement M_{III} we stressed the fact that it is the boundary conditions [$\varphi(x) \longrightarrow 0$ when $x \longrightarrow \pm\infty$] that introduce the quantization of the energy levels. The boundary conditions of the problem we are studying here, that is, the variation of the potential at the edges of the lattice, might thus be expected to play a critical role in determining the possible energies. Actually, this is not the case: we shall see that these energies depend practically only on the values of $V(x)$ in the region where it is periodic, and not on the edge effects (on condition, of course, that the number of potential wells is sufficiently large). In addition, we shall verify the result obtained intuitively in § 2-b, showing that most of the possible energies are grouped in allowed energy bands. Only a few stationary states, localized near the edges, depend on a critical manner on the variation of $V(x)$ in this region and can have an energy which falls in a forbidden band.

We shall therefore proceed essentially as in Complement M_{III}, first examining precisely the conditions imposed on the wave function $\varphi_\alpha(x)$ of a stationary state.

3-a. Conditions imposed on the wave function

In the region where $V(x)$ is periodic, relation (3d) gives the form of the wave function $\varphi_\alpha(x)$; the coefficients A_n and A'_n are determined from (32). To write (32) more explicitly, let

us set:

$$c_1(\alpha)\Lambda_1(\alpha) = \begin{pmatrix} f_1(\alpha) \\ f_1'(\alpha) \end{pmatrix}$$
$$c_2(\alpha)\Lambda_2(\alpha) = \begin{pmatrix} f_2(\alpha) \\ f_2'(\alpha) \end{pmatrix} \tag{35}$$

We then obtain:

$$A_n = f_1(\alpha)\,\lambda_1^{n-1} + f_2(\alpha)\,\lambda_2^{n-1}$$
$$A_n' = f_1'(\alpha)\,\lambda_1^{n-1} + f_2'(\alpha)\,\lambda_2^{n-1} \tag{36}$$

Now let us examine the boundary conditions on the wave function $\varphi_\alpha(x)$. First of all, to the left, far from the lattice, $V(x)$ is equal to V_e and $\varphi_\alpha(x)$ is written in the form:

$$\varphi_\alpha(x) = B\,\mathrm{e}^{\mu(\alpha)x} \tag{37a}$$

with:

$$\mu(\alpha) = \sqrt{\frac{2m}{\hbar^2}(V_e - E)} \tag{37b}$$

(we eliminate the solution in $\mathrm{e}^{-\mu(\alpha)x}$, which diverges when $x \longrightarrow -\infty$). The probability current associated with the function (37) is zero (*cf.* Complement $\mathrm{B_{III}}$). Now, for a stationary state, this current is independent of x [*cf.* Complement $\mathrm{N_{III}}$, relation (12)]; it therefore remains zero at all x, even inside the lattice. According to relation (14) of Complement $\mathrm{N_{III}}$, the coefficients A_n and A_n' therefore necessarily have the same modulus. Thus, if we choose to express the boundary conditions on the left as relations between the coefficients A_1 and A_1' [that is, by writing that the expression for $\varphi_\alpha(x)$ for $-\dfrac{l}{2} \leq x \leq \dfrac{l}{2}$ is the extension of the wave function (37)], we find a relation of the form:

$$\frac{A_1}{A_1'} = \mathrm{e}^{i\chi(\alpha)} \tag{38a}$$

$\chi(\alpha)$ is a real function of α (and therefore of the energy E) which depends on the precise behavior of $V(x)$ at the left-hand edge of the lattice [in what follows, we shall not need the exact expression for this function $\chi(\alpha)$; the essential point is that the boundary conditions on the left have the form (38a)].

The same type of reasoning can obviously be applied on the right ($x \longrightarrow +\infty$). The boundary conditions are written:

$$\frac{A_{N+1}}{A_{N+1}'} = \mathrm{e}^{i\chi'(\alpha)} \tag{38b}$$

where the real function $\chi'(\alpha)$ depends on the behavior of $V(x)$ on the right-hand edge of the lattice.

To sum up, we can say that the quantization of the energy levels can be obtained in the following manner:

– we start with two coefficients A_1 and A_1' that satisfy (38a); this ensures that the function $\varphi_\alpha(x)$ will remain bounded when $x \longrightarrow -\infty$. Since $\varphi_\alpha(x)$ is defined to within a constant factor, we can choose, for example:

$$A_1 = \mathrm{e}^{i\chi(\alpha)/2}$$
$$A_1' = \mathrm{e}^{-i\chi(\alpha)/2} \tag{39}$$

– we then calculate, using (36), the coefficients A_n and A'_n so as to extend the wave function chosen throughout all the crystal. Note that the condition (39) implies that $\varphi_\alpha(x)$ is real (*cf.* Complement N_{III}, § 1-b); calculation of A_n and A'_n must therefore yield:

$$A'_n = A^*_n \tag{40}$$

– finally, we write that the coefficients A_{N+1} and A'_{N+1} satisfy (38b), a relation ensuring that $\varphi_\alpha(x)$ will remain bounded when $x \longrightarrow +\infty$. In fact, relation (40) shows that the ratio A_{N+1}/A'_{N+1} is automatically a complex number of unit modulus; condition (38b) therefore amounts to an equality between the phases of two complex numbers. We thus obtain a real equation in α, which has a certain number of real solutions giving the allowed energies.

We are going to apply this method, distinguishing between two cases: real eigenvalues of $Q(\alpha)$ [the case where $|X(\alpha)| > 1$] and imaginary ones [the case where $|X(\alpha)| < 1$].

3-b. Allowed energy bands: stationary states of the particle inside the lattice

First assume that the energy E is in a domain where $|X(\alpha)| < 1$.

α. *Form of the quantization equation*

Taking (26) into account, relations (36) become:

$$\begin{cases} A_n = f_1(\alpha)\, e^{i(n-1)k(\alpha)l} + f_2(\alpha)\, e^{-i(n-1)k(\alpha)l} \\ A'_n = f'_1(\alpha)\, e^{i(n-1)k(\alpha)l} + f'_2(\alpha)\, e^{-i(n-1)k(\alpha)l} \end{cases} \tag{41}$$

We also know that the choice (39) of A_1 and A'_1 implies that $A'_n = A^*_n$ for all n. Now, it is easy to show that relations (41) yield two complex conjugate numbers only if:

$$\begin{aligned} f^*_1(\alpha) &= f'_2(\alpha) \\ f^*_2(\alpha) &= f'_1(\alpha) \end{aligned} \tag{42}$$

Condition (38b) can then be written:

$$\frac{f_1(\alpha)\, e^{2iNk(\alpha)l} + f_2(\alpha)}{f^*_2(\alpha)\, e^{2iNk(\alpha)l} + f^*_1(\alpha)} = e^{i\chi'(\alpha)} \tag{43}$$

This equation in α gives the quantization of the energy levels. To solve it, let us set:

$$\Theta(\alpha) = \mathrm{Arg}\left\{ \frac{f^*_1(\alpha)\, e^{i\chi'(\alpha)/2} - f_2(\alpha)\, e^{-i\chi'(\alpha)/2}}{f_1(\alpha)\, e^{-i\chi'(\alpha)/2} - f^*_2(\alpha)\, e^{i\chi'(\alpha)/2}} \right\} \tag{44}$$

[$\Theta(\alpha)$ can, in principle, be calculated from $\chi(\alpha)$, $\chi'(\alpha)$ and the matrix $Q(\alpha)$]. Equation (43) can then be written simply:

$$e^{2iNk(\alpha)l} = e^{i\Theta(\alpha)} \tag{45}$$

The energy levels are therefore given by:

$$k(\alpha) = \frac{\Theta(\alpha)}{2Nl} + p\frac{\pi}{Nl} \tag{46}$$

with:

$$p = 0,\ 1,\ 2,\ ...,\ (N-1) \tag{47}$$

[the other values of p must be excluded, since condition (25) forces $k(\alpha)$ to vary within an interval of width π/l]. We can already see that if N is very large, we can write equation (46) in the simplified form:

$$k(\alpha) \simeq p\frac{\pi}{Nl} \tag{48}$$

β. Graphical solution; locating the energy levels

If we substitute definition (24) of $k(\alpha)$ into (46), we obtain an equation in α that gives the allowed energies. To solve it graphically, let us begin by tracing the curve that represents the function $X(\alpha) = \mathrm{Re}[e^{i\alpha l} F(\alpha)]$. Because of the imaginary exponential $e^{i\alpha l}$, we expect this curve to have an oscillatory behavior, of the type shown in Figure 4a. Since $|F(\alpha)|$ is greater than 1 [*cf.* Complement $\mathrm{N_{III}}$, relation (32)], the amplitude of the oscillation is greater than 1, and the curve intersects the two straight lines $X(\alpha) = \pm 1$ at certain values α_0, α_1, α_2, ... of the variable α. We then eliminate all regions of the x-axis, bounded by these values, where the condition $|X(\alpha)| < 1$ is not satisfied. Using the set of arcs of curves thus obtained for $X(\alpha)$, we must represent the function:

$$k(\alpha) = \frac{1}{l} \operatorname{Arc\,cos}\, X(\alpha) \tag{49}$$

Taking into account the form of the Arc cosine function (*cf.* Fig. 5), we are led to the curve whose shape is shown in Figure 4b. Equation (46) indicates that the energy levels correspond to the intersections of this curve with the curves representing the functions $\dfrac{\Theta(\alpha)}{2Nl} + p\dfrac{\pi}{Nl}$, that is, if $N \gg 1$, with the horizontal lines whose equations are $y = p\dfrac{\pi}{Nl}$ (with $p = 0, 1, 2, ..., N-1$).

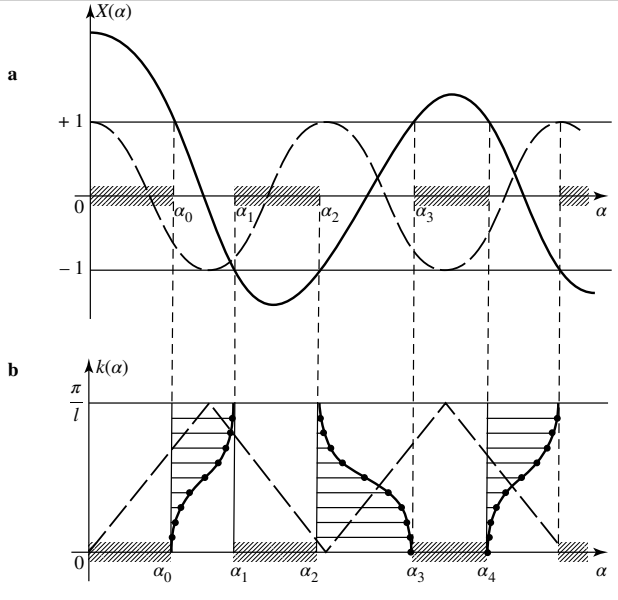


Figure 4: Variation with respect to α of $X(\alpha) = \mathrm{Re}[F(\alpha)\, e^{i\alpha l}]$ (see Fig. 2) and of $k(\alpha) = \dfrac{1}{l}\operatorname{Arc\,cos}\, X(\alpha)$. The values of α (that is, of the energy E) associated with stationary states are obtained (if $N \gg 1$) by cutting the curve which represents $k(\alpha)$ with the horizontal lines whose equations are $y = p\pi/Nl$ $(p = 0, 1, 2, ... N-1)$. The allowed bands are thus revealed (intervals $\alpha_0 \leq \alpha \leq \alpha_1$, etc.); each includes N very close levels. The forbidden bands are represented by the shaded areas ($\alpha_1 < \alpha < \alpha_2$, etc...). The dashed-line curves correspond to the special case where $V(x) = 0$ (a free particle).

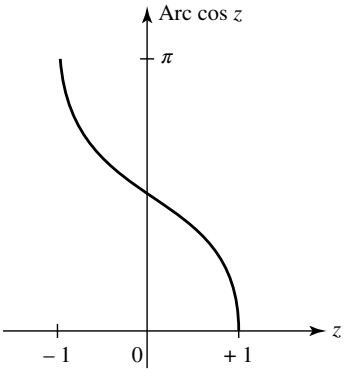


Figure 5: The Arc cosine function.

We thus obtain groups of N levels, associated with equidistant values of $k(\alpha)$ and situated in the allowed bands defined by $\alpha_0 \leq \alpha \leq \alpha_1,\ \alpha_2 \leq \alpha \leq \alpha_3$, etc. Between these allowed bands are the forbidden bands (we shall examine their properties in § c).

If we consider a particular allowed band, we can locate each level according to the value of $k(\alpha)$ which corresponds to it. This leads to choosing k as the variable and considering α and, consequently, E as functions $\alpha(k)$ and $E(k)$ of k. The variation of α with respect to k is given directly by the curve of Figure 4b, so it suffices to evaluate the function $\dfrac{\hbar^2\alpha^2}{2m}$ to obtain the energy $E(k)$. The corresponding curve has the shape shown in Figure 6.

Comment:

It is clear from Figure 4-b that, to a given value of k, correspond several values of α and therefore of the energy; this is why several arcs appear in Figure 6. Nevertheless,

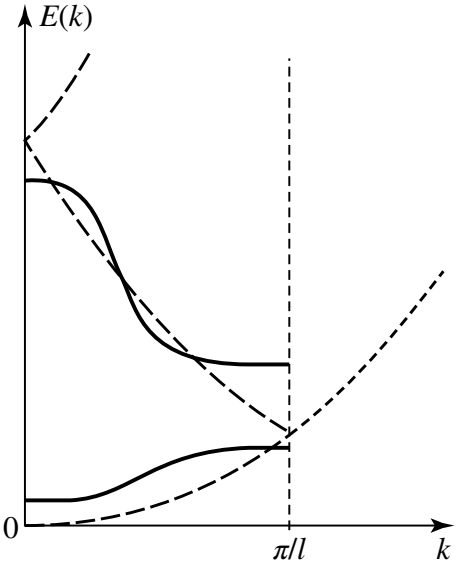


Figure 6: Variation of the energy with respect to the parameter k. The solid lines correspond to the energies for the first two allowed bands (the values of k which give the energy levels being equidistant inside the interval $0 \leq k \leq \pi/l$). The dashed lines correspond to the special case where the potential $V(x)$ is zero (a free particle); the allowed bands are then contiguous, and there are no forbidden bands.

if, within a given allowed band, $X(\alpha)$ increases steadily from -1 to $+1$ (or decreases steadily from $+1$ to -1), only one energy level corresponds to each value of k for this band, and *this band includes N energy levels.*

γ. *Discussion*

The preceding calculations show how, when we go from $N = 1$ to very high values of N, we move gradually from a set of discrete levels to allowed energy bands. Rigorously, these bands are formed by discrete levels, but their separation is so small for a macroscopic lattice that they practically constitute a continuum. When k is taken as a parameter, the *density of states* (the number of possible energies per unit interval of k) *is constant* and equal to Nl/π. This property, which is very useful, explains why k is generally chosen as the variable.

An important point appears in going from (46) to (48): when N is large, the edge effects of the lattice, which enter only through the intermediary of the functions $\chi(\alpha)$, $\chi'(\alpha)$ and, in (46), $\Theta(\alpha)$, no longer play any role; only the form of the periodic potential inside the lattice is important in determining the possible energies.

It is interesting to consider the two following limiting cases:

(*i*) If $V(x) = 0$ (free particle), we have:

$$\begin{cases} F(\alpha) = 1 \\ X(\alpha) = \cos \alpha l \end{cases} \tag{50}$$

and we obtain:

$$\begin{aligned} &\text{if} \quad 0 \leq \alpha \leq \frac{\pi}{l} : \qquad k(\alpha) = \alpha \\ &\text{if} \quad \frac{\pi}{l} \leq \alpha \leq \frac{2\pi}{l} : \qquad k(\alpha) = \frac{2\pi}{l} - \alpha \\ &\text{etc...} \end{aligned} \tag{51}$$

(the corresponding broken line is shown in Figure 4b as a dashed line). Relation (50) shows that the condition $|X(\alpha)| \leq 1$ is always satisfied: as we know, forbidden bands do not exist for a free particle.

Figure 6 thus enables us to see the effect of the potential $V(x)$ on the curve $E(k)$. When forbidden bands appear, the curves representing the energy become deformed so as to have horizontal tangents for $k = 0$ and $k = \pi/l$ (edges of the band). Unlike what happens for a free particle, there exists a point of inflection for each band where the energy varies linearly with k.

(*ii*) If the transmission coefficient $T(\alpha)$ is practically zero, we have [*cf.* Complement $\mathrm{N_{III}}$, equations (29) and (21)]:

$$\begin{cases} |F(\alpha)| \gg 1 \\ |G(\alpha)| \gg 1 \end{cases} \tag{52}$$

In Figure 2, the point representing the complex number $e^{i\alpha l}F(\alpha)$ is very far from the origin. We thus see in this figure that the regions of the α-axis where $|X(\alpha)| < 1$ are extremely narrow. The allowed bands therefore shrink if the transmission coefficient of the elementary barrier decreases; in the limit of zero transmission, they reduce to individual levels in an isolated well. Inversely, as soon as the tunnel effect allows the particle to pass from one well to the next one, *each of the discrete levels of the well gives rise to an energy band, whose width increases as the transmission coefficient grows.* We shall return to this property in Complement $\mathrm{F_{XI}}$.

3-c. Forbidden bands: stationary states localized on the edges

α. Form of the equations; energy levels

Let us now assume that α belongs to a domain where $|X(\alpha)| > 1$. According to (30), relations (36) can then be written:

$$\begin{cases} A_n = \varepsilon^{n-1}\left[f_1(\alpha)\,\mathrm{e}^{(n-1)\rho(\alpha)l} + f_2(\alpha)\,\mathrm{e}^{-(n-1)\rho(\alpha)l}\right] \\ A'_n = \varepsilon^{n-1}\left[f'_1(\alpha)\,\mathrm{e}^{(n-1)\rho(\alpha)l} + f'_2(\alpha)\,\mathrm{e}^{-(n-1)\rho(\alpha)l}\right] \end{cases} \tag{53}$$

The fact that $A'_n = A^*_n$ for all n means that we must have:

$$\begin{cases} f'_1(\alpha) = f^*_1(\alpha) \\ f'_2(\alpha) = f^*_2(\alpha) \end{cases} \tag{54}$$

The quantization condition (38b) then takes on the form:

$$\frac{A_{N+1}}{A'_{N+1}} = \frac{f_1(\alpha) + f_2(\alpha)\,\mathrm{e}^{-2N\rho(\alpha)l}}{f^*_1(\alpha) + f^*_2(\alpha)\,\mathrm{e}^{-2N\rho(\alpha)l}}\mathrm{e}^{i\chi'(\alpha)} \tag{55}$$

that is:

$$\mathrm{e}^{-2N\rho(\alpha)l} = L(\alpha) \tag{56}$$

where the real function $L(\alpha)$ is defined by:

$$L(\alpha) = -\frac{f^*_1(\alpha)\,\mathrm{e}^{i\chi'(\alpha)/2} - f_1(\alpha)\,\mathrm{e}^{-i\chi'(\alpha)/2}}{f^*_2(\alpha)\,\mathrm{e}^{i\chi'(\alpha)/2} - f_2(\alpha)\,\mathrm{e}^{-i\chi'(\alpha)/2}} \tag{57}$$

Consider the case where $N \gg 1$; we then have $\mathrm{e}^{-2N\rho(\alpha)l} \simeq 0$, and equation (56) reduces to:

$$L(\alpha) = 0 \tag{58}$$

The energy levels situated in the forbidden bands are therefore given by the zeros of the function $L(\alpha)$ (*cf.* Fig. 7). N enters neither into (57) nor into (58), so the number of these levels does not depend on N (unlike the number of levels situated in an allowed band). Consequently, when $N \gg 1$, it can be said that practically all the levels are grouped in the allowed bands.

β. Discussion

The situation here is radically different from the one encountered in § b: the number N, that is, the length of the lattice, plays no role (provided, nevertheless, that it is sufficiently large); on the other hand, definition (57) of $L(\alpha)$ shows that the functions $\chi(\alpha)$ and $\chi'(\alpha)$ play an essential role in the problem. Since we already know that these functions depend on the behavior of $V(x)$ on the edges of the lattice, we expect to obtain states localized in these regions.

This is indeed the case. Equations (57) and (58) offer two possibilities:

(*i*) if $f_1(\alpha) \neq 0$, the fact that $L(\alpha) = 0$ requires that:

$$\frac{f_1(\alpha)}{f^*_1(\alpha)} = \frac{f_1(\alpha)}{f'_1(\alpha)} = \mathrm{e}^{i\chi'(\alpha)} \tag{59}$$

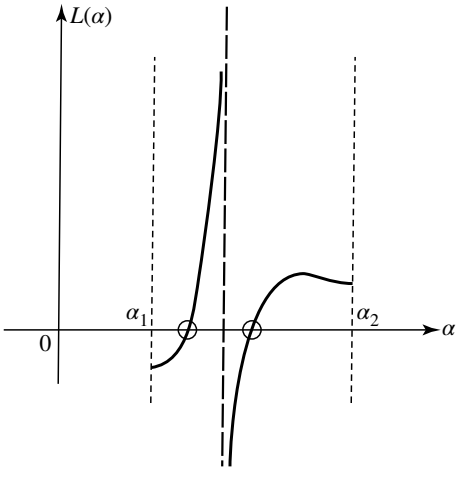


Figure 7: Variation of $L(\alpha)$ with respect to α in a forbidden band. The zeros of $L(\alpha)$ give the stationary states which are localized on the edges of the lattice.

Let us return to definition (35) of $f_1(\alpha)$ and $f_1'(\alpha)$; we see that relation (59) shows that the wave function constructed from the first eigenvector of $Q(\alpha)$ satisfies the boundary conditions on the right. This is easy to understand: if we start at $x = 0$ with an arbitrary wave function which satisfies the boundary conditions on the left, the matrix $\begin{pmatrix} A_1 \\ A_1' \end{pmatrix}$ has components on the two eigenvectors of $Q(\alpha)$; the coefficients A_{N+1} and A'_{N+1} are then (if $N \gg 1$) essentially given by (33), which expresses the fact that the column matrix $\begin{pmatrix} A_{N+1} \\ A'_{N+1} \end{pmatrix}$ is proportional to the column matrix of the first eigenvector of $Q(\alpha)$.

Note that since the eigenvalue $\lambda_1(\alpha)$ is greater than 1, the wave function grows exponentially when x increases. The stationary state given by the first eigenvector of $Q(\alpha)$ is therefore localized at the right end of the lattice.

(*ii*) if $f_1(\alpha) = 0$, (54) gives $f_1'(\alpha) = 0$, and definitions (35) imply that $c_1(\alpha) = 0$: the corresponding stationary state is associated with the second eigenvector of $Q(\alpha)$. Aside from the fact that this state is localized at the left end of the lattice, the conclusions obtained in (*i*) remain valid.

References and suggestions for further reading:

Merzbacher (1.16), Chap. 6, § 7; Flügge (1.24), §§ 28 and 29; Landau and Lifshitz (1.19), § 104; see also solid state physics texts (section 13 of the bibliography).

Chapter IV

Application of the postulates to simple cases: spin 1/2 and two-level systems

A Spin 1/2 particle: quantization of the angular momentum . 394
A-1 Experimental demonstration 394
A-2 Theoretical description 399
B Illustration of the postulates in the case of a spin 1/2 . . . 401
B-1 Actual preparation of the various spin states 401
B-2 Spin measurements 404
B-3 Evolution of a spin 1/2 particle in a uniform magnetic field . 409
C General study of two-level systems 411
C-1 Outline of the problem 412
C-2 Static aspect: effect of coupling on the stationary states of the system 413
C-3 Dynamical aspect: oscillation of the system between the two unperturbed states 418

In this chapter, we intend to illustrate the postulates of quantum mechanics, which we stated and discussed in Chapter III. We shall apply them to simple concrete cases, in which the dimension of the state space is finite (equal to two). The interest of these examples is not confined to their mathematical simplicity, which will allow a better understanding of the postulates and their consequences. It is also based on their physical importance: they exhibit typically quantum mechanical behavior which can be verified experimentally.

In §§ A and B, we study the spin 1/2 case (which we shall take up again in more detail in Chapter IX). First, we describe (§ A-1) a fundamental experiment that revealed

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

the quantization of a simple physical quantity, the angular momentum. We shall see that the component along Oz of the angular momentum (or magnetic moment) of a neutral paramagnetic atom can take on only certain values, which belong to a *discrete* set. Thus, for a silver atom in its ground state, there are only two possible values ($+\hbar/2$ and $-\hbar/2$) for the component S_z of its angular momentum: a silver atom in the ground state is said to be a spin 1/2 *particle*. In § A-2, we indicate how quantum mechanics describes the "spin variables" of such a particle. In situations where one can dispense with a quantum treatment of the "external variables" $\mathbf{r}$ and $\mathbf{p}$, the state of the particle ("spin state space") has only two dimensions. We shall then (§ B) be able to illustrate and discuss the quantum mechanical postulates in this particularly simple case: we shall first see how to prepare silver atoms in any desired arbitrary spin state, in a real experiment. We shall then show how the measurement of the physical values of the spin on such silver atoms enables us to verify the quantum mechanical postulates experimentally. By integrating the corresponding Schrödinger equation, we shall study the evolution of a spin 1/2 particle in a uniform magnetic field (Larmor precession). Finally, in § C, we shall begin the study of *two-level systems*. Although these systems are not generally spin 1/2 particles, their study leads to calculations very similar to those developed in §§ A and B. We shall treat in detail the effect of an external perturbation on the stationary states of a two-level system and use this very simple model to point out important physical effects.

A. Spin 1/2 particle: quantization of the angular momentum

A-1. Experimental demonstration

First of all, we are going to describe and analyze the Stern-Gerlach experiment, which demonstrated the quantization of the components of an angular momentum (sometimes called "space quantization").

A-1-a. The Stern-Gerlach apparatus

The experiment consists of studying the deflection of a beam of neutral paramagnetic atoms (in this case, silver atoms) in a highly inhomogeneous magnetic field. The apparatus used is shown schematically[1] in Figure 1.

Silver atoms contained in a furnace E, heated to a high temperature, leave through a small opening and propagate in a straight line in the high vacuum existing inside the whole apparatus. A collimating slit F selects those atoms whose velocity is parallel to a particular direction that we shall choose for the Oy axis. The atomic beam thus constructed traverses the gap of an electromagnet A before condensing on a plate P.

Let us describe the characteristics of the magnetic field $\mathbf{B}$ produced by the electromagnet A. This magnetic field has a plane of symmetry (which we shall designate by yOz) that contains the initial direction Oy of the atomic beam. In the air-gap, it is the same at all points situated on any given line parallel to Oy (the edges of the electromagnet are parallel to Oy, and we neglect edge effects). $\mathbf{B}$ has no component along Oy. Its largest component is along Oz; it varies strongly with z: in Figure 1-b, the field lines are much closer together close to the north pole than close to the south pole of the magnet.

[1] We only indicate the most important characteristics of this equipment. A more detailed description of the experimental technique can be found in a textbook on atomic physics.

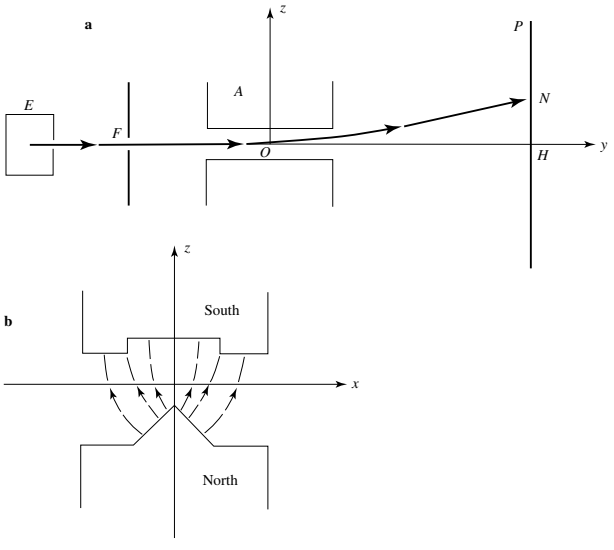


Figure 1: Schematic diagram of the Stern-Gerlach experiment. Figure a shows the trajectory of a silver atom emitted from the high-temperature furnace E. This atom is deflected by the gradient of the magnetic field created by the electromagnet A and then condenses at N on plate P.
Figure b shows a cross section in the xOz plane of the electromagnet A; the lines of force of the magnetic field are shown in dashed lines. B_z has been assumed to be positive and $\partial B_z/\partial z$, negative. Consequently, the trajectory of figure a corresponds to a negative component $\mathscr{M}_z$ of the magnetic moment, that is, to a positive component of $\mathscr{S}_z$ (γ is negative for a silver atom).

Of course, since the magnetic field has a conserved flux ($\mathrm{div}\mathbf{B} = 0$), it must also have a component along Ox which varies with the distance x from the plane of symmetry.

A-1-b. Classical calculation of the deflection[2]

Note, first, that the silver atoms, being neutral, are not subjected to the Lorentz force. On the other hand, they possess a permanent magnetic moment $\boldsymbol{\mathscr{M}}$ (they are paramagnetic atoms); the resulting forces are derived from the potential energy:

$$W = -\boldsymbol{\mathscr{M}} \cdot \mathbf{B} \tag{A-1}$$

The existence, for an atom, of an electronic magnetic moment $\boldsymbol{\mathscr{M}}$ and an angular momentum $\boldsymbol{\mathscr{S}}$ is due to two causes: the motion of the electrons about the nucleus (the corresponding rotation of the charges being responsible for the appearance of an orbital magnetic moment) and the intrinsic angular momentum, or spin, (*cf.* Chapter IX) of the electrons, with which is associated a spin magnetic moment. It can be shown (as we shall

[2]We only give here an outline of the calculation.

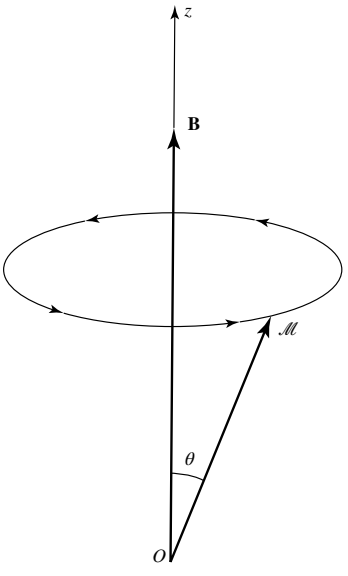


Figure 2: The silver atom possesses a magnetic moment $\mathscr{M}$ and an angular momentum $\mathscr{S}$ which are proportional. Consequently, the effect of a uniform maguetic field **B** *is to cause $\mathscr{M}$ to turn about* **B** *with a constant angular velocity (Larmor precession).*

assume here without proof) that, for a given atomic level, $\mathscr{M}$ and $\mathscr{S}$ are proportional[3]:

$$\mathscr{M} = \gamma \mathscr{S} \tag{A-2}$$

The proportionality constant γ is called the *gyromagnetic ratio* of the level under consideration.

Before the atoms traverse the electromagnet, the magnetic moments of the silver atoms that form the atomic beam are oriented randomly (isotropically). Let us study the action of the magnetic field on one of these atoms, whose magnetic moment $\mathscr{M}$ has a given direction at the entrance of the air-gap. From expression (A-1) for the potential energy, it is easy to deduce that the resultant of the forces exerted on the atom is:

$$\mathbf{F} = \boldsymbol{\nabla}(\mathscr{M} \cdot \mathbf{B}) \tag{A-3}$$

(this resultant would be equal to zero if the field **B** were uniform), and that their total moment relative to the position of the atom is:

$$\boldsymbol{\Gamma} = \mathscr{M} \times \mathbf{B} \tag{A-4}$$

The angular momentum theorem can be written:

$$\frac{\mathrm{d}\mathscr{S}}{\mathrm{d}t} = \boldsymbol{\Gamma} \tag{A-5}$$

[3]In the case of silver atoms in the ground state (like those of the beam), the angular momentum $\mathscr{S}$ is simply equal to the spin of the outer electron, which is therefore solely responsible for the existence of the magnetic moment $\mathscr{M}$. This is because the outer electron has a zero orbital angular momentum, and the resultant orbital and spin angular momenta of the inner electrons are also zero. Moreover the experimental conditions realized in practice are such that effects linked to the spin of the nucleus are negligible. This is why the silver atom in the ground state, like the electron, has a spin 1/2.

that is:

$$\frac{\mathrm{d}\boldsymbol{\mathscr{S}}}{\mathrm{d}t} = \gamma\, \boldsymbol{\mathscr{S}} \times \mathbf{B} \tag{A-6}$$

The atom thus behaves like a gyroscope (Fig. 2): $\mathrm{d}\boldsymbol{\mathscr{S}}/\mathrm{d}t$ is perpendicular to $\boldsymbol{\mathscr{S}}$, and the angular momentum turns about the magnetic field, the angle θ between $\boldsymbol{\mathscr{S}}$ and $\mathbf{B}$ remaining constant. The rotational angular velocity is equal to the product of the gyromagnetic ratio γ and the modulus of the magnetic field. The components of $\boldsymbol{\mathscr{M}}$ which are perpendicular to the magnetic field therefore oscillate around zero, the component parallel to $\mathbf{B}$ remaining constant.

To calculate the force $\mathbf{F}$ [formula (A-3)], we can, to a very good approximation, neglect in the potential energy W the terms proportional to $\mathscr{M}_x$ and $\mathscr{M}_y$ and take $\mathscr{M}_z$ to be constant. This is because the frequency of oscillation due to the rotation of $\boldsymbol{\mathscr{M}}$ is so great that only the time-averaged values of $\mathscr{M}_x$ and $\mathscr{M}_y$ can play a role in W, and these are both zero. Consequently, it is as if the atom were submitted to the sole force:

$$\mathbf{F}' = \boldsymbol{\nabla}(\mathscr{M}_z B_z) = \mathscr{M}_z\, \boldsymbol{\nabla} B_z \tag{A-7}$$

In addition, the components of $\boldsymbol{\nabla} B_z$ along Ox and Oy are zero: $\partial B_z/\partial y = 0$ because the magnetic field is independent of y (§ A-1-a above), and $\partial B_z/\partial x = 0$ at all points of the plane of symmetry yOz. *The force on the atom is therefore parallel to Oz and proportional to* $\mathscr{M}_z$. Since it is this force that produces the deflection HN of the atom (Fig. 1), HN is proportional to $\mathscr{M}_z$ (and hence, to $\mathscr{S}_z$). Consequently, measuring HN is equivalent to measuring $\mathscr{M}_z$ or $\mathscr{S}_z$.

Since, at the entrance to the air-gap, the moments of the various atoms of the beam are distributed isotropically (all values of $\mathscr{M}_z$ included between $|\boldsymbol{\mathscr{M}}|$ and $-|\boldsymbol{\mathscr{M}}|$ are found), we expect the beam to form a single pattern, symmetrical with respect to H, on the plate P. The upper bound N_1 and the lower bound N_2 of this pattern correspond in principle to the maximum value $|\boldsymbol{\mathscr{M}}|$ and minimum value $-|\boldsymbol{\mathscr{M}}|$ of $\mathscr{M}_z$. In fact, the dispersion of the velocities and the finite width of the slit F cause the atoms having a given value of $\mathscr{M}_z$ to condense, not at the same point, but in a spot centered about the deflection corresponding to the average velocity.

A-1-c. Results and conclusions

The results of the experiment (performed for the first time in 1922 by Stern and Gerlach) are in complete contradiction with the preceding predictions.

We do not observe a single spot centered at H, but *two spots* (Fig. 3) centered at the points N_1 and N_2, symmetrical with respect to H (the width of these two spots is due to the dispersion of the velocities, and to the width of the slit F). The predictions of classical mechanics are therefore shown to be invalidated by the experiment.

Now let us see how these experimental results can be interpreted. Of the physical quantities associated with a silver atom, some correspond to its external degrees of freedom (that is, are functions of its position $\mathbf{r}$ and its linear momentum $\mathbf{p}$), and others, to its internal degrees of freedom (also called spin degrees of freedom) $\boldsymbol{\mathscr{M}}$ or $\boldsymbol{\mathscr{S}}$.

Let us first show that, under these experimental conditions, it is not necessary to treat the external degrees of freedom quantum mechanically. To do this, we shall verify that it is possible, in order to describe the motion of the silver atoms, to construct wave

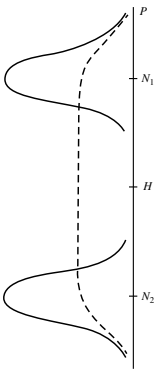


Figure 3: Spots observed on the plate P in the Stern-Gerlach experiment. The magnetic moments $\mathscr{M}$ of the atoms emitted from the furnace E are distributed randomly in all directions of space, so classical mechanics predicts that a measurement of $\mathscr{M}_z$ can yield with equal probability all values included between $+|\mathscr{M}|$ and $-|\mathscr{M}|$. One should therefore observe only one large spot centered in H (dashed lines in the figure). In reality, the result of the experiment is completely different: two spots, centered at N_1 and N_2, are observed. This means that a measurement of $\mathscr{M}_z$ can yield only two possible results (quantization of the measurement result).

packets whose width Δz and momentum dispersion Δp_z are negligible. Δz and Δp_z must satisfy the Heisenberg relation:

$$\Delta z.\Delta p_z \gtrsim \hbar \tag{A-8}$$

Numerically, the mass M of a silver atom is equal to 1.8×10^{-25} kg. Δz and the velocity uncertainty $\Delta v_z = \Delta p_z / M$ must be such that:

$$\Delta z.\Delta v_z \gtrsim \frac{\hbar}{M} \simeq 10^{-9} \quad \text{M.K.S.A.} \tag{A-9}$$

Now what are the lengths and velocities involved in the problem? The width of the slit F is about 0.1 mm and the separation N_1N_2 of the two spots, that is several millimeters. The distance over which the magnetic field varies appreciably can be deduced from the values of the field in the middle of the air-gap ($B \simeq 10^4$ gauss) and its gradient ($\frac{\partial B}{\partial z} \simeq 10^5$ gauss/cm), which yields $B/\frac{\partial B}{\partial z} \simeq 1$ mm. In addition, the velocity of the silver atoms leaving a furnace at an absolute temperature of 1000 °K is of the order of 500 m/s. However well-defined the beam is, the dispersion of the velocities along Oz is not much less than several meters per second. It is then easy to find uncertainties Δz and Δv_z, which, while satisfying (A-9), are negligible on the scale of the experiment being considered. As far as the external variables $\mathbf{r}$ and $\mathbf{p}$ of each atom are concerned, it is therefore not necessary to resort to quantum mechanics. It is possible to reason in terms of quasi-pointlike wave packets moving along classical trajectories. Consequently, it is correct to claim that measurement of the deflection HN constitutes a measurement of $\mathscr{M}_z$ or $\mathscr{S}_z$.

The results of the experiment thus lead us necessarily to the following conclusion: if we measure the component $\mathscr{S}_z$ of the intrinsic angular momentum of a silver atom in its ground state, we can find only one or the other of two values corresponding to the deflections HN_1 and HN_2. We are therefore obliged to reject the classical image of a vector $\mathscr{S}$ whose angle θ with the magnetic field can take on any value: $\mathscr{S}_z$ is a quantized physical quantity whose discrete spectrum includes only two eigenvalues. When we study the quantum theory of angular momentum (Chap. VI), we shall see that these eigenvalues are $+\hbar/2$ and $-\hbar/2$; we shall assume this here and say that the spin of the silver atom in its ground state is 1/2.

A-2. Theoretical description

We are now going to show how quantum mechanics describes the degrees of freedom of a silver atom, that is, of a spin 1/2 particle.

We do not yet possess all the necessary elements for the presentation of a deductive and rigorous theory of the spin 1/2 particle. Such a study will be developed in Chapter IX, in the framework of the general theory of angular momentum. We shall therefore be forced here to assume without proof a small number of results which will be proved later, in Chapter IX. Such a point of view is justified by the fact that the essential goal of the present chapter is to show the reader how to handle the quantum mechanical formalism in a simple and concrete case, and not to focus on the angular momentum aspect of the spin 1/2. The idea is to give precise examples of kets and observables, to show how physical predictions can be extracted from them and how to distinguish clearly between the various stages of an experiment (preparation, evolution, measurement).

We saw in Chapter III that with every measurable physical quantity must be associated, in quantum mechanics, an observable, that is, a Hermitian operator whose eigenvectors can form a basis in the state space. We must therefore define the state space and the observables corresponding to the components of $\boldsymbol{\mathscr{S}}$ ($\mathscr{S}_x$, $\mathscr{S}_y$, $\mathscr{S}_z$ and, more generally, $\mathscr{S}_{\mathbf{u}} = \boldsymbol{\mathscr{S}} \cdot \mathbf{u}$, where $\mathbf{u}$ is an arbitrary unit vector), which we know from § A-1 to be measurable.

A-2-a. The observable S_z and the spin state space

With $\mathscr{S}_z$ we must associate an observable S_z which has, according to the results of the experiment described in § A-1 above, two eigenvalues, $+\hbar/2$ and $-\hbar/2$. We shall assume (see Chap. IX) that these two eigenvalues are not degenerate, and we shall denote by $|+\rangle$ and $|-\rangle$ the corresponding orthonormal eigenvectors:

$$\begin{cases} S_z|+\rangle = +\dfrac{\hbar}{2}\,|+\rangle \\ S_z|-\rangle = -\dfrac{\hbar}{2}\,|-\rangle \end{cases} \tag{A-10}$$

with:

$$\begin{cases} \langle +|+\rangle = \langle -|-\rangle = 1 \\ \langle +|-\rangle = 0 \end{cases} \tag{A-11}$$

S_z alone therefore forms a C.S.C.O., and the spin state space is the two-dimensional space $\mathcal{E}_s$ spanned by its eigenvectors $|+\rangle$ and $|-\rangle$. The fact that these eigenvectors form a basis of $\mathcal{E}_s$ is expressed by the closure relation:

$$|+\rangle\langle +| \;+\; |-\rangle\langle -| = \mathbb{1} \tag{A-12}$$

The most general (normalized) vector of $\mathcal{E}_s$ is a linear superposition of $|+\rangle$ and $|-\rangle$:

$$|\psi\rangle = \alpha|+\rangle + \beta|-\rangle \tag{A-13}$$

with:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{A-14}$$

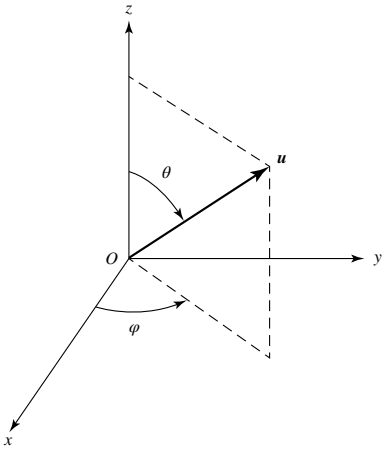


Figure 4: Definition of the polar angles θ and φ characterizing a unit vector $\mathbf{u}$.

In the $|+\rangle, |-\rangle$ basis, the matrix representing S_z is diagonal and is written:

$$(S_z) = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A-15}$$

A-2-b. The other spin observables

With the $\mathscr{S}_x$ and $\mathscr{S}_y$ components of $\mathscr{S}$ will be associated the observables S_x and S_y. The operators S_x and S_y must be represented in the $\{|+\rangle, |-\rangle\}$ basis by 2×2 Hermitian matrices.

We shall see in Chapter VI that in quantum mechanics, the three components of an angular momentum do not commute with each other but satisfy well-defined commutation relations. This will enable us to show that, in the case of a spin 1/2, with which we are concerned here, the matrices representing S_x and S_y in the basis of the eigenvectors $|+\rangle$ and $|-\rangle$ of S_z are the following:

$$(S_x) = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{A-16}$$

$$(S_y) = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{A-17}$$

For the moment, we shall assume this result.

As for the $\mathscr{S}_{\mathbf{u}}$ component of $\boldsymbol{\mathscr{S}}$ along the unit vector $\mathbf{u}$, characterized by the polar angles θ and φ (Fig. 4), it is written:

$$\mathscr{S}_{\mathbf{u}} = \boldsymbol{\mathscr{S}} \cdot \mathbf{u} = \mathscr{S}_x \sin\theta\cos\varphi + \mathscr{S}_y \sin\theta\sin\varphi + \mathscr{S}_z \cos\theta \tag{A-18}$$

Using (A-15), (A-16) and (A-17), we easily find the matrix that represents the corresponding observable $S_{\mathbf{u}} = \mathbf{S}.\ \mathbf{u}$ in the $\{|+\rangle, |-\rangle\}$ basis:

$$\begin{aligned}(S_u) &= (S_x)\sin\theta\cos\varphi + (S_y)\sin\theta\sin\varphi + (S_z)\cos\theta \\ &= \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta\ \mathrm{e}^{-\mathrm{i}\varphi} \\ \sin\theta\ \mathrm{e}^{\mathrm{i}\varphi} & -\cos\theta \end{pmatrix}\end{aligned} \tag{A-19}$$

In what follows, we shall need to know the eigenvalues and eigenvectors of the observables S_x, S_y and S_u. The calculations that enable us to obtain them from the matrices (A-16), (A-17) and (A-19) are not difficult. We shall only present the results here.

The S_x, S_y and S_u operators have the same eigenvalues, $+\hbar/2$ and $-\hbar/2$, as S_z. This result could have been expected, since it is always possible to rotate the Stern-Gerlach apparatus as a whole so as to make the axis defined by the magnetic field parallel either to Ox, to Oy, or to $\mathbf{u}$. Since all directions of space have the same properties, the phenomena observed on the plate of the apparatus must be unchanged under such rotations: the measurement of $\mathscr{S}_x$, $\mathscr{S}_y$ or $\mathscr{S}_u$ can therefore yield only one of two results: $+\hbar/2$ or $-\hbar/2$.

As for the eigenvectors of S_x, S_y and S_u, we shall denote them respectively by $|\pm\rangle_x$, $|\pm\rangle_y$ and $|\pm\rangle_u$ (the sign in the ket is that of the corresponding eigenvalue). Their expansions on the basis of eigenvectors $|\pm\rangle$ of S_z is written:

$$|\pm\rangle_x = \frac{1}{\sqrt{2}}\left[|+\rangle \pm |-\rangle\right] \tag{A-20}$$

$$|\pm\rangle_y = \frac{1}{\sqrt{2}}\left[|+\rangle \pm i|-\rangle\right] \tag{A-21}$$

$$\begin{cases} |+\rangle_u = \cos\dfrac{\theta}{2}\ \mathrm{e}^{-i\varphi/2}|+\rangle + \sin\dfrac{\theta}{2}\ \mathrm{e}^{i\varphi/2}|-\rangle & \text{(A-22a)} \\ |-\rangle_u = -\sin\dfrac{\theta}{2}\ \mathrm{e}^{-i\varphi/2}|+\rangle + \cos\dfrac{\theta}{2}\ \mathrm{e}^{i\varphi/2}|-\rangle & \text{(A-22b)} \end{cases}$$

B. Illustration of the postulates in the case of a spin 1/2

Using the formalism that we have just described, we are now going to apply the postulates of quantum mechanics to a certain number of experiments on silver atoms which can actually be performed with the Stern-Gerlach apparatus. We shall thus be able to discuss the consequences of these postulates in a concrete case.

B-1. Actual preparation of the various spin states

In order to make predictions about the result of a measurement, we must know the state of the system (here, the spin of a silver atom) immediately before the measurement. We are going to see how to prepare a beam of silver atoms so that they are all in a given spin state.

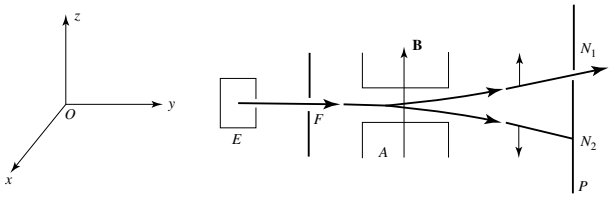


Figure 5: When we pierce a hole in the plate P at the position of the spot N_1, the atoms that pass through this hole are all in the spin state $|+\rangle$. The Stern-Gerlach apparatus is then acting like a polarizer.

B-1-a. Preparation of the states $|+\rangle$ and $|-\rangle$

Let us assume that we pierce a hole in the plate P of the apparatus represented in Figure 1-a, at the position of the spot centered at N_1 (Fig. 3). The atoms which are deflected downward continue to condense about N_2, while some of those which are deflected upward pass through the plate P (Fig. 5). Each of the atoms of the beam which propagates to the right of the plate is a physical system on which we have just performed a measurement of the observable S_z, the result being $+\hbar/2$. According to the fifth postulate of Chapter III, this atom is in the eigenstate corresponding to this result, that is, in the state $|+\rangle$ (since S_z alone constitutes a C.S.C.O., the measurement result suffices to determine the state of the system after this measurement). The device in Figure 5 thus produces a beam of atoms which are all in the spin state $|+\rangle$. This device acts like an "atomic polarizer", since it acts the same way on atoms as an ordinary polarizer does on photons.

Of course, if we pierced the plate around N_2 and not around N_1, we would obtain a beam all of whose atoms would be in the spin state $|-\rangle$.

B-1-b. Preparation of the states $|\pm\rangle_x$, $|\pm\rangle_y$, $|\pm\rangle_u$

The observable S_x also constitutes a C.S.C.O. since none of its eigenvalues is degenerate. To prepare one of its eigenstates, we must simply select, after a measurement of S_x, the atoms for which this measurement has yielded the corresponding eigenvalue. In practice, if we rotate the apparatus of Figure 5 through an angle of $+\pi/2$ about Oy, we obtain a beam of atoms whose spin state is $|+\rangle_x$ (Fig. 6).

This method can be generalized: by placing the Stern-Gerlach apparatus so that the axis of the magnetic field is parallel to an arbitrary unit vector $\mathbf{u}$, and piercing the plate either at N_1 or at N_2, we can prepare silver atoms in the spin state $|+\rangle_u$ or $|-\rangle_u$[4].

B-1-c. Preparation of the most general state

We indicated above that the most general (normalized) ket of the spin state space is of the form:

$$|\psi\rangle = \alpha|+\rangle + \beta|-\rangle \tag{B-1}$$

[4]The direction of the atomic beam is no longer necessarily Oy, but this is not important in what concerns us here.

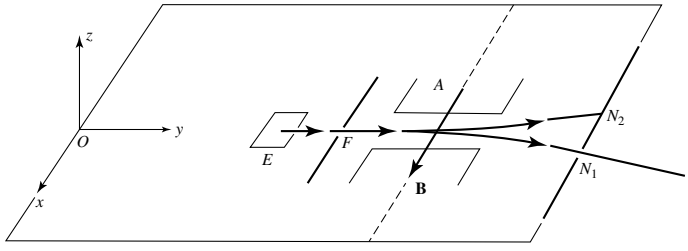


Figure 6: When the apparatus of Figure 5 is rotated through 90° about Oy, we obtain a polarizer which prepares atoms in the spin state $|+\rangle_x$.

---

with:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{B-2}$$

Is it possible to prepare atoms whose spin state is described by the corresponding ket $|\psi\rangle$?

We are going to show that there exists, for all $|\psi\rangle$, a unit vector $\mathbf{u}$ such that $|\psi\rangle$ is collinear with the ket $|+\rangle_{\mathbf{u}}$. We therefore choose two complex numbers α and β which satisfy relation (B-2) but which are arbitrary in every other respect. Taking (B-2) into account, we find that there necessarily exists an angle θ such that:

$$\begin{cases} \cos\dfrac{\theta}{2} = & |\alpha| \\ \sin\dfrac{\theta}{2} = & |\beta| \end{cases} \tag{B-3}$$

If, in addition, we impose:

$$0 \leq \theta \leq \pi \tag{B-4}$$

the equation $\tan\frac{\theta}{2} = \left|\frac{\beta}{\alpha}\right|$ determines θ uniquely. We already know that only the difference of the phases of α and β enters into the physical predictions. Let us therefore set:

$$\varphi = \text{Arg } \beta - \text{Arg } \alpha \tag{B-5}$$

$$\chi = \text{Arg } \beta + \text{Arg } \alpha \tag{B-6}$$

We thus have:

$$\begin{aligned} \text{Arg } \beta = & \quad \frac{1}{2}\chi + \frac{1}{2}\varphi \\ \text{Arg } \alpha = & \quad \frac{1}{2}\chi - \frac{1}{2}\varphi \end{aligned} \tag{B-7}$$

With this notation, the ket $|\psi\rangle$ can be written:

$$|\psi\rangle = \mathrm{e}^{i\frac{\chi}{2}} \left[\cos\frac{\theta}{2}\ \mathrm{e}^{-i\frac{\varphi}{2}}|+\rangle + \sin\frac{\theta}{2}\ \mathrm{e}^{i\frac{\varphi}{2}}|-\rangle\right] \tag{B-8}$$

If we compare this expression with formula (A-22a), we see that $|\psi\rangle$ differs from the ket $|+\rangle_u$ (which corresponds to the unit vector $\mathbf{u}$ characterized by θ and φ) only by the phase factor $e^{i\chi/2}$, which has no physical significance.

Consequently, to prepare silver atoms in the state $|\psi\rangle$, it suffices to place the Stern-Gerlach apparatus (with its plate pierced at N_1) so that its axis is directed along the vector $\mathbf{u}$ whose polar angles are determined from α and β by (B-3) and (B-5).

B-2. Spin measurements

We saw in § A that a Stern-Gerlach apparatus enables us to *measure* the component of the angular momentum $\mathscr{S}$ of silver atoms along a given axis. We have just pointed out, in § B-1, that an apparatus of the same type can be used to *prepare* an atomic beam in a given spin state. Consequently, if we place two Stern-Gerlach magnets one after the other, we can verify experimentally the predictions of the postulates. The first apparatus acts like a "polarizer": the beam coming out of it is composed of a large number of silver atoms all in the same spin state. This beam then enters the second apparatus, which is used to measure a specified component of the angular momentum $\mathscr{S}$: this is, as it were, the "analyzer" (note the analogy with the optical experiment described in § A-3 of Chapter I). We shall assume in this section that the spin state of the atoms of the beam does not evolve between the time they leave the "polarizer" and the time they enter the "analyzer", that is, between the preparation and the measurement. It would be easy to forgo this hypothesis, by using the Schrödinger equation to determine the spin evolution between the moment of preparation and the moment of measurement.

B-2-a. First experiment

Let us choose the axes of the two apparatuses parallel to Oz (Fig. 7). The first one prepares the atoms in the state $|+\rangle$ and the second one measures $\mathscr{S}_z$. What is observed on the plate of the second apparatus?

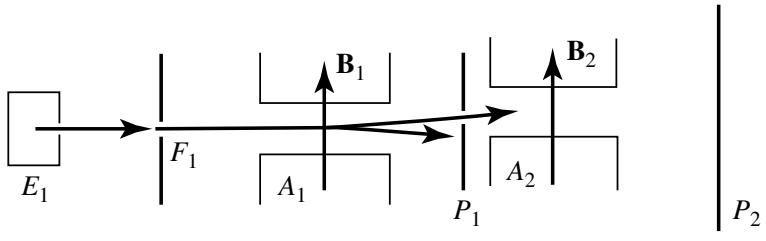


Figure 7: The first apparatus (a "source" composed of the furnace E_1 and the slit F_1, plus a "polarizer" formed by the magnet A_1 and the pierced plate P_1) prepares the atoms in the state $|+\rangle$. The second one (an "analyzer" composed of the magnet A_2 and the plate P_2) measures the component $\mathscr{S}_z$. The result obtained is certain ($+\hbar/2$).

Since the state of the system under study is an eigenstate of the observable S_z which we want to measure, the postulates indicate that the measurement result is *certain*: we find, without fail, the corresponding eigenvalue $(+\hbar/2)$. Consequently, all the atoms of

the beam must condense into a single spot on the plate of the second apparatus, at the position corresponding to $+\hbar/2$.

This is indeed what is observed experimentally: all the atoms strike the second plate in the vicinity of N_1, none hitting near N_2.

B-2-b. Second experiment

Now let us place the axis of the first apparatus along the unit vector $\mathbf{u}$, with polar angles θ, $\varphi = \pi$ ($\mathbf{u}$ is therefore contained in the xOz plane). The axis of the second apparatus remains parallel to Oz (Fig. 8). According to (A-22a), the spin state of the atoms when they leave the "polarizer" is (we ignore an irrelevant i factor multiplying whole ket):

$$|\psi\rangle = -\cos\frac{\theta}{2}\,|+\rangle + \sin\frac{\theta}{2}\,|-\rangle \tag{B-9}$$

The "analyzer" measures $\mathscr{S}_z$ on these atoms. What are the results?

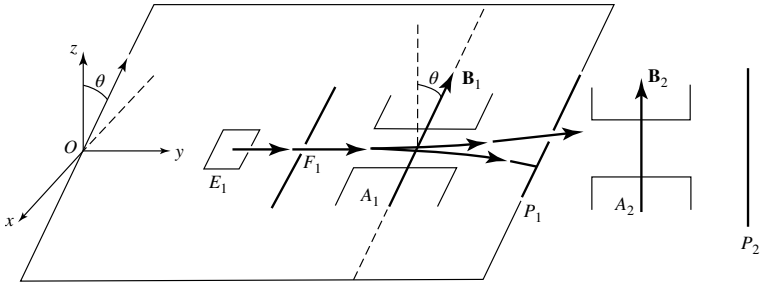


Figure 8: The first apparatus prepares the spins in the state $|+\rangle_u$ ($\mathbf{u}$ is the unit vector of the xOz plane that makes an angle θ with Oz). The second one measures the $\mathscr{S}_z$ component. The possible results are $+\hbar/2$ (probability $\cos^2\theta/2$) and $-\hbar/2$ (probability $\sin^2\theta/2$).

This time, we find that certain atoms condense at N_1, and others at N_2, although they have all been prepared in the same way: there is an indeterminacy in the behavior of each of the atoms taken individually. The postulate of spectral decomposition merely enables us to predict the probability of each atom's appearance at N_1 or N_2. Since (B-9) gives the expansion of the spin state of an atom in terms of the eigenstates of the observable being measured, we can calculate directly that these probabilities are, respectively, $\cos^2\theta/2$ and $\sin^2\theta/2$. Thus, when enough atoms have condensed on the plate, we observe that the intensity of the spots at N_1 and N_2 corresponds to numbers of atoms which are proportional, respectively, to $\cos^2\theta/2$ and $\sin^2\theta/2$.

Comment:

For any value of the angle θ (except exactly 0 or π), it is therefore always possible to find the two results $+\hbar/2$ and $-\hbar/2$ in a measurement of S_z. This prediction

may seem to a certain extent paradoxical. For example, if θ is very small, the spin at the exit of the first apparatus points in a direction which is practically Oz, and yet one can find $-\hbar/2$ as well as $+\hbar/2$ in a measurement of S_z (while, in classical mechanics, the result would be $(\hbar\cos\theta)/2 \simeq \hbar/2$). Nevertheless, the smaller θ, the smaller the probability of finding $-\hbar/2$. Moreover, we shall see later [formula (B-11)] that the mean value of the results which would be obtained in a large number of identical experiments is $\langle S_z\rangle = \frac{\hbar}{2}\cos\theta$, which corresponds to the classical result.

B-2-c. Third experiment

Let us take a "polarizer" positioned as in § B-2-b, so as to prepare atoms in the state (B-9), and let us rotate the "analyzer" until its axis is directed along Ox, so that it measures the $\mathscr{S}_x$ component of the angular momentum.

To calculate the predictions of the postulates in this case, we must expand the state (B-9) in terms of the eigenstates of the observable S_x [formula (A-20)]. We find:

$$\begin{cases} {}_x\langle +|\psi\rangle = & \frac{1}{\sqrt{2}}(\cos\frac{\theta}{2} + \sin\frac{\theta}{2}) = & \cos(\frac{\pi}{4} - \frac{\theta}{2}) \\ {}_x\langle -|\psi\rangle = & \frac{1}{\sqrt{2}}(\cos\frac{\theta}{2} - \sin\frac{\theta}{2}) = & \sin(\frac{\pi}{4} - \frac{\theta}{2}) \end{cases} \tag{B-10}$$

The probability of finding the eigenvalue $+\hbar/2$ of S_x is therefore $\cos^2\left(\frac{\pi}{4} - \frac{\theta}{2}\right)$ and that of finding $-(\hbar/2)$, $\sin^2\left(\frac{\pi}{4} - \frac{\theta}{2}\right)$.

It is possible to verify these predictions by measuring the intensity of the two spots on the plate situated at the exit of the second Stern-Gerlach apparatus.

Comment:

The fact that it is $\left(\frac{\pi}{4} - \frac{\theta}{2}\right)$ that enters in here is not at all surprising: in § B-2-b, the angle between the axes of the two apparatus was θ; it became $\left(\frac{\pi}{2} - \theta\right)$ after rotation of the second apparatus.

B-2-d. Mean values

In the situation of § B-2-b, we find experimentally that, of a great number $\mathcal{N}$ of atoms, $\mathcal{N}\cos^2\theta/2$ arrive at N_1 and $\mathcal{N}\sin^2\theta/2$, at N_2. The measurement of $\mathscr{S}_z$ therefore yields $+\hbar/2$ for each of the first group and $-\hbar/2$ for each of the second. If we calculate the mean value of these results, we obtain:

$$\begin{aligned} \langle S_z\rangle &= \frac{1}{\mathcal{N}}\left[\frac{\hbar}{2}\times\mathcal{N}\cos^2\frac{\theta}{2} - \frac{\hbar}{2}\times\mathcal{N}\sin^2\frac{\theta}{2}\right] \\ &= \frac{\hbar}{2}\cos\theta \end{aligned} \tag{B-11}$$

It is easy to verify from formulas (B-9) and (A-10) that this is indeed the value of the matrix element $\langle\psi|S_z|\psi\rangle$.

Similarly, the average of the measurement results obtained in the experiment of § B-2-c is equal to:

$$\begin{aligned}\langle S_x\rangle &= \frac{1}{\mathcal{N}}\left[\frac{\hbar}{2}\times\mathcal{N}\cos^2\left(\frac{\pi}{4}-\frac{\theta}{2}\right)-\frac{\hbar}{2}\times\mathcal{N}\sin^2\left(\frac{\pi}{4}-\frac{\theta}{2}\right)\right]\\ &= \frac{\hbar}{2}\sin\theta\end{aligned} \tag{B-12}$$

To calculate the matrix element $\langle\psi|S_x|\psi\rangle$, we can use the matrix (A-16) which represents S_x in the $\{|+\rangle,|-\rangle\}$ basis. In this same basis, the ket $|\psi\rangle$ is represented by the column vector $\begin{pmatrix}\cos\theta/2\\ \sin\theta/2\end{pmatrix}$, and the bra $\langle\psi|$ by the corresponding row vector. We therefore have:

$$\begin{aligned}\langle\psi|S_x|\psi\rangle &= \frac{\hbar}{2}\begin{pmatrix}\cos\theta/2 & \sin\theta/2\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\cos\theta/2\\ \sin\theta/2\end{pmatrix}\\ &= \frac{\hbar}{2}\sin\theta\end{aligned} \tag{B-13}$$

The mean value of $\mathscr{S}_x$ is indeed equal to the matrix element, in the state $|\psi\rangle$, of the associated observable S_x.

It is interesting to note that if we were dealing with a *classical* angular momentum of modulus $\hbar/2$ directed along the axis of the "polarizer", its components along Ox and Oz would be precisely $(\hbar/2)\sin\theta$ and $(\hbar/2)\cos\theta$. More generally, if we calculate [using the same technique as in (B-13)] the mean values of S_x, S_y and S_z in the state $|+\rangle_u$ [formula (A-22a)], we find:

$$\begin{aligned}{}_u\langle+|S_x|+\rangle_u &= \frac{\hbar}{2}\sin\theta\cos\varphi\\ {}_u\langle+|S_y|+\rangle_u &= \frac{\hbar}{2}\sin\theta\sin\varphi\\ {}_u\langle+|S_z|+\rangle_u &= \frac{\hbar}{2}\cos\theta\end{aligned} \tag{B-14}$$

These mean values are equal to the components of a classical angular momentum of modulus $\hbar/2$ oriented along the vector $\mathbf{u}$ whose polar angles are θ and φ. Therefore, we can also establish here a relation between classical mechanics and quantum mechanics through the mean values. However, we must not lose sight of the fact that *a measurement of $\mathscr{S}_x$, for example, on a given atom will never yield $\frac{\hbar}{2}\sin\theta\cos\varphi$*: the only results which can be found are $+\hbar/2$ and $-\hbar/2$. Only in taking the average of values obtained in a large number of identical measurements (same state of the system, here $|+\rangle_u$, and same observable measured, here S_x) do we obtain $\frac{\hbar}{2}\sin\theta\cos\varphi$.

Comment:

It is useful to consider again at this stage the problem of external degrees of freedom (position, momentum).

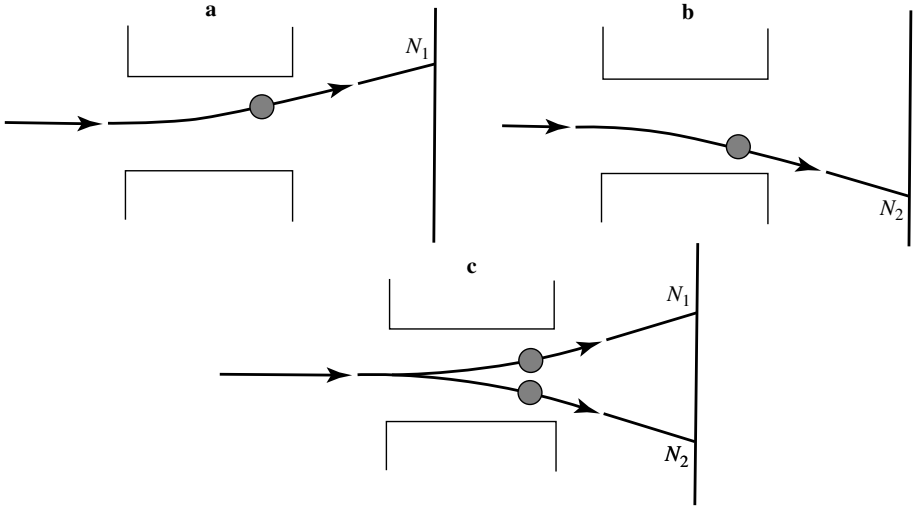


Figure 9: When the spin is in the state $|+\rangle$ (fig. a) or $|-\rangle$ (fig. b), the center of the wave packet follows a well-defined trajectory which can be calculated classically. When the spin state is a linear superposition of $|+\rangle$ and $|-\rangle$, the wave packet splits into two parts and it is no longer possible to say that the atom follows a classical trajectory (despite the fact that the spread of each of the packets is much smaller than the characteristic dimensions of the problem).

When a silver atom enters the second Stern-Gerlach apparatus in the spin state $|\psi\rangle$ given by (B-9), we have just seen that it is impossible to predict with certainty whether it will condense at N_1 or N_2. It seems difficult to reconcile this indeterminacy with the idea of a perfectly well-determined classical trajectory, given the initial state of the system.

In fact, this is not a real paradox. To say that the external degrees of freedom can be treated classically means only that it is possible to form wave packets which are much smaller than all the dimensions of the problem. It does not necessarily mean, as we shall see, that the particle itself follows a classical trajectory.

Let us first consider a silver atom which enters the apparatus in the initial spin state $|+\rangle$. The wave function which describes the external degrees of freedom of this particle is a wave packet whose spread is very small and whose center follows the classical trajectory of Figure 9-a. Similarly, if the silver atom enters with the spin state $|-\rangle$, the center of the wave packet associated with it follows the classical trajectory of Figure 9-b.

If we now consider an atom which enters with the spin state $|\psi\rangle$ of formula (B-9), the corresponding initial state is a well-defined linear superposition of the two preceding initial states. Since the Schrödinger equation is linear, the wave function of the particle at a subsequent instant (Fig. 9-c) is a linear superposition of the two wave packets of Figures 9-a and 9-b. The particle therefore has a certain probability amplitude of being in one or the other of these two wave packets. We see that it does not follow a classical trajectory at all, unlike what happens to the centers of the two wave packets. Upon arrival on the screen, the wave function has non-zero values in two different regions, each

very localized, around the points N_1 and N_2. The particle can therefore appear either near N_1 or near N_2, and we cannot predict with certainty at which of these two points the appearance will occur.

Note that the two wave packets of Figure 9-c do not represent two different particles; they represent only one particle, whose wave function has two parts, each of which is very localized about a different point. The two wave packets, moreover, have a well-defined phase relation because they arise from the same initial wave packet, split into two under the influence of the gradient of $\mathbf{B}$. We could recombine them to form one wave packet again by removing the screen (that is, by not performing the measurement) and by submitting them to a new field gradient, whose sign would be the opposite of the first one.

B-3. Evolution of a spin 1/2 particle in a uniform magnetic field

B-3-a. The interaction Hamiltonian and the Schrödinger equation

Consider a silver atom in a *uniform* magnetic field $\mathbf{B}_0$, and choose the Oz axis along $\mathbf{B}_0$. The classical potential energy of the magnetic moment $\boldsymbol{\mathscr{M}} = \gamma \boldsymbol{\mathscr{S}}$ of this atom is then:

$$W = -\boldsymbol{\mathscr{M}} \cdot \mathbf{B}_0 = -\mathscr{M}_z B_0 = -\gamma B_0 \mathscr{S}_z \tag{B-15}$$

where B_0 is the modulus of the magnetic field. Let us set:

$$\omega_0 = -\gamma B_0 \tag{B-16}$$

It is easy to see that ω_0 has the dimensions of the inverse of a time, that is, of an angular velocity.

Since we are quantizing only the internal degrees of freedom of the particle, $\mathscr{S}_z$ must be replaced by the operator S_z, and the classical energy (B-15) becomes an operator: it is the Hamiltonian H which describes the evolution of the spin of the atom in the field $\mathbf{B}_0$:

$$H = \omega_0 S_z \tag{B-17}$$

Since this operator is time-independent, solving the corresponding Schrödinger equation amounts to solving the eigenvalue equation of H. We see that the eigenvectors of H are those of S_z:

$$\begin{aligned} H|+\rangle &= +\frac{\hbar\omega_0}{2}\ |+\rangle \\ H|-\rangle &= -\frac{\hbar\omega_0}{2}\ |-\rangle \end{aligned} \tag{B-18}$$

There are therefore two energy levels, $E_+ = +\hbar\omega_0/2$ and $E_- = -\hbar\omega_0/2$ (Fig. 10). Their separation $\hbar\omega_0$ is proportional to the magnetic field; they define a single "Bohr frequency":

$$\nu_{+-} = \frac{1}{h}(E_+ - E_-) = \frac{\omega_0}{2\pi} \tag{B-19}$$

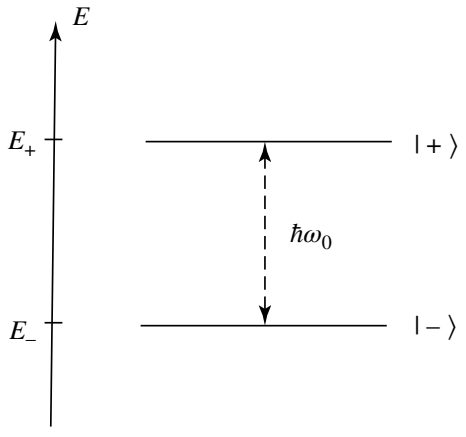


Figure 10: Energy levels of a spin 1/2, of gyromagnetic ratio γ, placed in a magnetic field B_0 parallel to Oz; ω_0 is defined by $\omega_0 = -\gamma B_0$.

Comment:

(*i*) If the field $\mathbf{B}_0$ is parallel to the unit vector $\mathbf{u}$ whose polar angles are θ and φ, relation (B-17) must be replaced by:

$$H = \omega_0 S_u \tag{B-20}$$

where $S_u = \mathbf{S} \cdot \mathbf{u}$ is the component of $\mathbf{S}$ along $\mathbf{u}$.

(*ii*) For the silver atom, γ is negative; ω_0 is therefore positive, according to (B-16). This explains the arrangement of the levels in Figure 10.

B-3-b. Larmor precession

Let us assume that, at time $t = 0$, the spin is in the state:

$$|\psi(0)\rangle = \cos\frac{\theta}{2}\ \mathrm{e}^{-i\varphi/2}|+\rangle + \sin\frac{\theta}{2}\ \mathrm{e}^{i\varphi/2}|-\rangle \tag{B-21}$$

(we showed in § B-1-c that any state could be put in this form). To calculate the state $|\psi(t)\rangle$ at an arbitrary instant $t > 0$, we apply the rule (D-54) given in Chapter III. In expression (B-21), $|\psi(0)\rangle$ is already expanded in terms of the eigenstates of the Hamiltonian, and we therefore obtain:

$$|\psi(t)\rangle = \cos\frac{\theta}{2}\ \mathrm{e}^{-i\frac{\varphi}{2}}\ \mathrm{e}^{-iE_+\frac{t}{\hbar}}|+\rangle + \sin\frac{\theta}{2}\ \mathrm{e}^{i\frac{\varphi}{2}}\ \mathrm{e}^{-iE_-\frac{t}{\hbar}}|-\rangle \tag{B-22}$$

or, using the values of E_+ and E_-:

$$|\psi(t)\rangle = \cos\frac{\theta}{2}\ \mathrm{e}^{-i(\varphi+\omega_0 t)/2}|+\rangle + \sin\frac{\theta}{2}\ \mathrm{e}^{i(\varphi+\omega_0 t)/2}|-\rangle \tag{B-23}$$

The presence of the magnetic field $\mathbf{B}_0$ therefore introduces a phase shift, proportional to the time, between the coefficients of the kets $|+\rangle$ and $|-\rangle$.

Comparing expression (B-23) for $|\psi(t)\rangle$ with that for the eigenket $|+\rangle_u$ of the observable $\mathbf{S}\cdot\mathbf{u}$ [formula (A-22a)], we see that the direction $\mathbf{u}(t)$ along which the spin component is $+\hbar/2$ with certainty is defined by the polar angles:

$$\begin{cases}\theta(t)=\theta\\ \varphi(t)=\varphi+\omega_0 t\end{cases} \tag{B-24}$$

The angle between $\mathbf{u}(t)$ and Oz (the direction of the magnetic field $\mathbf{B}_0$) therefore remains constant, but $\mathbf{u}(t)$ revolves about Oz at an angular velocity of ω_0 (proportional to the magnetic field). Thus, we find in quantum mechanics the phenomenon which we described for a classical magnetic moment in § A-1-b, and which bears the name of *Larmor precession*.

From expression (B-17) for the Hamiltonian, it is obvious that the observable S_z is a constant of the motion. It can be verified from (B-23) that the probabilities of obtaining $+\hbar/2$ or $-\hbar/2$ in a measurement of this observable are time-independent. Since the modulus of $e^{\pm i(\varphi+\omega_0 t)/2}$ is equal to 1, these probabilities are equal, respectively, to $\cos^2\theta/2$ and $\sin^2\theta/2$. The mean value of S_z is also time-independent:

$$\langle\psi(t)|S_z|\psi(t)\rangle=\frac{\hbar}{2}\cos\theta \tag{B-25}$$

On the other hand, S_x and S_y do not commute with H [it is easy to show this by using the matrices representing S_x, S_y and S_z, which are given in (A-15), (A-16) and (A-17)]. Thus, formulas (B-14) here become:

$$\begin{aligned}\langle\psi(t)|S_x|\psi(t)\rangle &= \frac{\hbar}{2}\sin\theta\,\cos(\varphi+\omega_0 t)\\ \langle\psi(t)|S_y|\psi(t)\rangle &= \frac{\hbar}{2}\sin\theta\,\sin(\varphi+\omega_0 t)\end{aligned} \tag{B-26}$$

In these expressions, we again find the single Bohr frequency $\omega_0/2\pi$ of the system. Moreover, the mean values of S_x, S_y and S_z behave like the components of a classical angular momentum of modulus $\hbar/2$ undergoing Larmor precession.

C. General study of two-level systems

The simplicity of the calculations presented in § B derives from the fact that the state space has only two dimensions.

There exist numerous other cases in physics which, to a first approximation, can be treated just as simply. Consider, for example, a physical system having two states whose energies are close together and very different from those of all other states of the system. Assume that we want to evaluate the effect of an external perturbation (or of internal interactions previously neglected) on these two states. When the intensity of the perturbation is sufficiently weak, it can be shown (*cf.* Chap. XI) that its effect on the two states can be calculated, to a first approximation, by ignoring all the other energy levels of the system. All the calculations can then be perfomed in a two-dimensional subspace of the state space.

In this section, we shall study certain general properties of two-level systems (which are not necessarily spin 1/2 particles). Such a study is interesting because it enables us, using a mathematically simple model, to bring out some general and important physical ideas (quantum resonance, oscillation between two levels, etc...).

C-1. Outline of the problem

C-1-a. Notation

Consider a physical system whose state space is two-dimensional (as we have already pointed out, this is usually an approximation: under certain conditions, we can confine ourselves to a two-dimensional subspace of the state space). For a basis, we choose the system of the two eigenstates $|\varphi_1\rangle$ and $|\varphi_2\rangle$ of the Hamiltonian H_0 whose eigenvalues are, respectively, E_1 and E_2:

$$\begin{aligned} H_0|\varphi_1\rangle &= E_1\ |\varphi_1\rangle \\ H_0|\varphi_2\rangle &= E_2\ |\varphi_2\rangle \end{aligned} \tag{C-1}$$

This basis is orthonormal:

$$\langle\varphi_i|\varphi_j\rangle = \delta_{ij} \quad ; \quad i,j = 1,2 \tag{C-2}$$

Assume that we want to take into account an external perturbation, or interactions internal to the system, initially neglected in H_0. The Hamiltonian becomes:

$$H = H_0 + W \tag{C-3}$$

The eigenstates and eigenvalues of H will be denoted by $|\psi_\pm\rangle$ and $E_\pm$:

$$\begin{aligned} H|\psi_+\rangle &= E_+|\psi_+\rangle \\ H|\psi_-\rangle &= E_-|\psi_-\rangle \end{aligned} \tag{C-4}$$

H_0 is often called the unperturbed Hamiltonian and W, the perturbation or coupling. We shall assume here that W is time-independent. In the $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ basis of eigenstates of H_0 (called unperturbed states), W is represented by a Hermitian matrix:

$$(W) = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \tag{C-5}$$

W_{11} and W_{22} are real. Moreover:

$$W_{12} = (W_{21})^* \tag{C-6}$$

In the absence of coupling, E_1 and E_2 are the possible energies of the system, and the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are stationary states (if the system is placed in one of these two states, it remains there indefinitely). The problem consists of evaluating the modifications that appear when the coupling W is introduced.

C-1-b. Consequences of the coupling

α. E_1 and E_2 are no longer the possible energies of the system

A measurement of the energy of the system can yield only one of the two eigenvalues E_+ and E_- of H, which generally differ from E_1 and E_2.

The first problem is to calculate E_+ and E_- in terms of E_1, E_2 and the matrix elements W_{ij} of W. This amounts to studying the effect of the coupling on the energy levels.

β. $|\varphi_1\rangle$ *and* $|\varphi_2\rangle$ *are no longer stationary states*

Since $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are not generally eigenstates of the total Hamiltonian H, they are no longer stationary states. If, for example, the system at time $t = 0$ is in the state $|\varphi_1\rangle$, there is a certain probability $\mathscr{P}_{12}(t)$ of finding it in the state $|\varphi_2\rangle$ at time t: W therefore induces transitions between the two unperturbed states. Hence the name "coupling" (between $|\varphi_1\rangle$ and $|\varphi_2\rangle$) given to W.

This dynamic aspect of the effect of W constitutes the second problem with which we shall be concerned.

Comment:

In Complement C_{IV}, the two problems we have just cited are considered by introducing the concept of a fictitious spin. It can indeed be shown that the Hamiltonian H to be diagonalized has the same form as that of a spin 1/2 placed in a static magnetic field $\mathbf{B}$, whose components B_x, B_y and B_z are simply expressed in terms of E_1, E_2 and the matrix elements W_{ij}. In other words, with every two-level system (not necessarily a spin 1/2), can be associated a spin 1/2 (called a fictitious spin) placed in a static field $\mathbf{B}$ and described by a Hamiltonian whose form is identical to that of the initial two level system. All the results related to two-level systems which we are going to establish in this section can be interpreted in a simple geometric way in terms of magnetic moment, Larmor precession, and the various concepts introduced in §§ A and B of this chapter in connection with spin 1/2 particles. This geometrical interpretation is developed in Complement C_{IV}.

C-2. Static aspect: effect of coupling on the stationary states of the system

C-2-a. Expressions for the eigenstates and eigenvalues of H

In the $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ basis, the matrix representing H is written:

$$(H) = \begin{pmatrix} E_1 + W_{11} & W_{12} \\ W_{21} & E_2 + W_{22} \end{pmatrix} \tag{C-7}$$

The diagonalization of matrix (C-7) presents no problems (it is performed in detail in Complement B_{IV}). We find the eigenvalues:

$$\begin{aligned} E_+ &= \frac{1}{2}(E_1 + W_{11} + E_2 + W_{22}) + \frac{1}{2}\sqrt{(E_1 + W_{11} - E_2 - W_{22})^2 + 4|W_{12}|^2} \\ E_- &= \frac{1}{2}(E_1 + W_{11} + E_2 + W_{22}) - \frac{1}{2}\sqrt{(E_1 + W_{11} - E_2 - W_{22})^2 + 4|W_{12}|^2} \end{aligned} \tag{C-8}$$

(it can be verified that if $W = 0$, E_+ and E_- are identical[5] to E_1 and E_2). The

[5] If $E_1 > E_2$, E_+ approaches E_1 and E_- approaches E_2 when W approaches zero. On the other hand, if $E_1 < E_2$, E_+ approaches E_2 and E_- approaches E_1.

eigenvectors associated with E_+ and E_- are written:

$$|\psi_+\rangle = \cos\frac{\theta}{2}\,\mathrm{e}^{-i\varphi/2}|\varphi_1\rangle + \sin\frac{\theta}{2}\,\mathrm{e}^{i\varphi/2}|\varphi_2\rangle \tag{C-9a}$$

$$|\psi_-\rangle = -\sin\frac{\theta}{2}\,\mathrm{e}^{-i\varphi/2}|\varphi_1\rangle + \cos\frac{\theta}{2}\,\mathrm{e}^{i\varphi/2}|\varphi_2\rangle \tag{C-9b}$$

where the angles θ and φ are defined by:

$$\tan\,\theta = \frac{2|W_{12}|}{E_1 + W_{11} - E_2 - W_{22}} \quad \text{with :} \quad 0 \leq \theta < \pi \tag{C-10}$$

$$W_{21} = |W_{21}|\,\mathrm{e}^{i\varphi} \tag{C-11}$$

C-2-b. Discussion

α. Graphical representation of the effect of coupling

All the interesting effects which we shall discuss later arise from the fact that the perturbation W possesses non-diagonal matrix elements $W_{12} = W_{21}^*$ (if $W_{12} = 0$, the eigenstates of H are the same as those of H_0, the new eigenvalues being simply $E_1 + W_{11}$ and $E_2 + W_{22}$). To simplify the discussion, we shall therefore assume from now on that the matrix (W) is purely non-diagonal[6], that is, that $W_{11} = W_{22} = 0$. Formulas (C-8) and (C-10) then become:

$$\begin{aligned} E_+ &= \frac{1}{2}(E_1 + E_2) + \frac{1}{2}\sqrt{(E_1 - E_2)^2 + 4|W_{12}|^2} \\ E_- &= \frac{1}{2}(E_1 + E_2) - \frac{1}{2}\sqrt{(E_1 - E_2)^2 + 4|W_{12}|^2} \end{aligned} \tag{C-12}$$

$$\tan\,\theta = \frac{2|W_{12}|}{E_1 + E_2} \qquad 0 \leq \theta < \pi \tag{C-13}$$

We now study the effect of the coupling W on the energies E_+ and E_- in terms of the values of E_1 and E_2. Assume that W_{12} is fixed and introduce the two parameters:

$$\begin{aligned} E_m &= \frac{1}{2}(E_1 + E_2) \\ \Delta &= \frac{1}{2}(E_1 - E_2) \end{aligned} \tag{C-14}$$

We see from (C-12) that the variation of E_+ and E_- with respect to E_m is extremely simple: changing E_m reduces to shifting the origin along the energy axis. Moreover, it can be verified from (C-9), (C-10) and (C-11) that the vectors $|\psi_+\rangle$ and $|\psi_-\rangle$ do not depend on E_m. We are therefore concerned only with the influence of the parameter Δ. Let us show on the same graph, in terms of Δ, the four energies E_1, E_2, E_+ and E_-. We obtain for E_1 and E_2 two straight lines of slope $+1$ and -1 (shown in dashed lines in Figure 11). Substituting (C-14) into (C-12), we find:

$$E_+ = E_m + \sqrt{\Delta^2 + |W_{12}|^2} \tag{C-15}$$

[6] If W_{11} and W_{22} are non-zero, we simply set: $\widetilde{E}_1 = E_1 + W_{11}$, $\widetilde{E}_2 = E_2 + W_{22}$. All the results obtained in this section then remain valid if we replace E_1 and E_2 by $\widetilde{E}_1$ and $\widetilde{E}_2$.

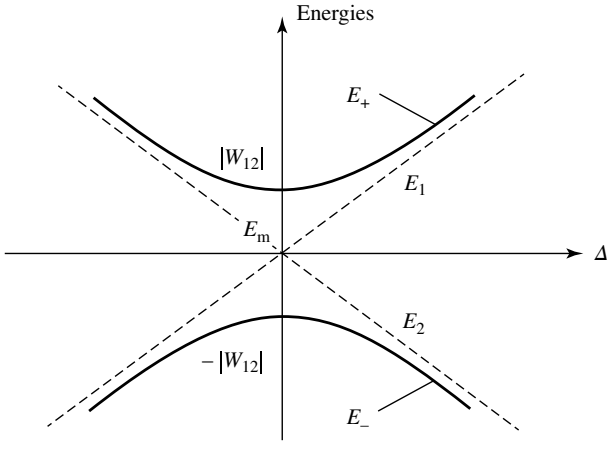


Figure 11: Variation of the energies E_+ and E_- as a function of the energy difference $\Delta = (E_1 - E_2)/2$. In the absence of coupling, the levels cross at the origin (dashed straight lines). Under the effect of the non-diagonal coupling W, the two perturbed levels "repel each other" and we obtain an "anti-crossing": the curves giving E_+ and E_- in terms of Δ are branches of a hyperbola (solid lines in the figure) whose asymptotes are the unperturbed levels.

$$E_- = E_m - \sqrt{\Delta^2 + |W_{12}|^2} \tag{C-16}$$

When Δ varies, E_+ and E_- describe the two branches of a hyperbola which is symmetrical with respect to the coordinate axes and whose asymptotes are the two straight lines associated with the unperturbed levels; the minimum separation between the two branches is $2|W_{12}|$ (solid lines in Figure 11)[7].

β. Effect of the coupling on the energy levels

In the absence of coupling, the energies E_1 and E_2 of the two levels "cross" at $\Delta = 0$. It is clear from Figure 11 that under the effect of coupling, the two levels "repel each other"– that is, the energy values move further away from each other. The diagram in solid lines in Figure 11 is often called, for this reason, an *anti-crossing* diagram.

Also, we see that, for any Δ, we always have:

$$|E_+ - E_-| > |E_1 - E_2|$$

This is a result that appears rather often in other domains of physics (for example, in electrical circuit theory): *the coupling separates the normal frequencies.*

[7]It is clear from Figure 11 why, when $W \longrightarrow 0$:

$E_+ \longrightarrow E_1, E_- \longrightarrow E_2$ if $E_1 > E_2$

$E_+ \longrightarrow E_2, E_- \longrightarrow E_1$ if $E_1 < E_2$.

Near the asymptotes, that is, for $|\Delta| \gg |W_{12}|$, formulas (C-15) and (C-16) can be written in the form of a limited power series expansion in $|\frac{W_{12}}{\Delta}|$:

$$\begin{aligned} E_+ &= E_m + \Delta\Big(1 + \frac{1}{2}\left|\frac{W_{12}}{\Delta}\right|^2 + ...\Big) \\ E_- &= E_m - \Delta\Big(1 + \frac{1}{2}\left|\frac{W_{12}}{\Delta}\right|^2 + ...\Big) \end{aligned} \tag{C-17}$$

On the other hand, at the center of the hyperbola, for $E_2 = E_1$ ($\Delta = 0$), formulas (C-15) and (C-16) yield:

$$\begin{aligned} E_+ &= E_m + |W_{12}| \\ E_- &= E_m - |W_{12}| \end{aligned} \tag{C-18}$$

Therefore, *the effect of the coupling is much more important when the two unperturbed levels have the same energy.* The effect is then of first order, as can be seen from (C-18), while it is of second order when $\Delta \gg |W_{12}|$ [formulas (C-17)].

γ. *Effect of the coupling on the eigenstates*

When (C-14) is used, formula (C-13) becomes:

$$\tan\ \theta = \frac{|W_{12}|}{\Delta} \tag{C-19}$$

It follows that, when $\Delta \ll |W_{12}|$ (strong coupling), $\theta \simeq \pi/2$; on the other hand, when $\Delta \gg |W_{12}|$ (weak coupling), $\theta \simeq 0$ (assuming $\Delta \geq 0$).

At the center of the hyperbola, when $E_2 = E_1$ ($\Delta = 0$), we have:

$$\begin{aligned} |\psi_+\rangle &= \frac{1}{\sqrt{2}}\left[e^{-i\varphi/2}|\varphi_1\rangle + e^{i\varphi/2}|\varphi_2\rangle\right] \\ |\psi_-\rangle &= \frac{1}{\sqrt{2}}\left[-e^{-i\varphi/2}|\varphi_1\rangle + e^{i\varphi/2}|\varphi_2\rangle\right] \end{aligned} \tag{C-20}$$

while near the asymptotes (that is, for $\Delta \gg |W_{12}|$), we have, to first order in $|W_{12}|/\Delta$:

$$\begin{aligned} |\psi_+\rangle &= e^{-i\varphi/2}\left[|\varphi_1\rangle + e^{i\varphi}\frac{|W_{12}|}{2\Delta}|\varphi_2\rangle + ...\right] \\ |\psi_-\rangle &= e^{i\varphi/2}\left[|\varphi_2\rangle - e^{-i\varphi}\frac{|W_{12}|}{2\Delta}|\varphi_1\rangle + ...\right] \end{aligned} \tag{C-21}$$

In other words, for a weak coupling ($E_1 - E_2 \gg |W_{12}|$), the perturbed states differ very slightly from the unperturbed states. We see from (C-21) that to within a global phase factor $e^{-i\varphi/2}$, $|\psi_+\rangle$ is equal to the state $|\varphi_1\rangle$ slightly "contaminated" by a small contribution from the state $|\varphi_2\rangle$. On the other hand, for a strong coupling ($E_1 - E_2 \ll |W_{12}|$), formulas (C-20) indicate that the states $|\psi_+\rangle$ and $|\psi_-\rangle$ are very different from the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, since they are linear superpositions of them with coefficients of the same modulus.

Thus, like the energies, the eigenstates undergo significant modifications in the neighborhood of the point where the two unperturbed states cross.

C-2-c. Important application: the phenomenon of quantum resonance

When $E_1 = E_2 = E_m$, the corresponding energy of H_0 is two-fold degenerate. As we have just seen, the coupling W_{12} lifts this degeneracy and, in particular, gives rise to a level whose energy is lowered by $|W_{12}|$. In other words, if the ground state of a physical system is two-fold degenerate (and sufficiently far from all the other levels), any (purely non-diagonal) coupling between the two corresponding states lowers the energy of the ground state of the system, which thus becomes more stable.

As a first example of this phenomenon, we shall cite the resonance stabilization of the benzene molecule C_6H_6. Experiments show that the six carbon atoms are situated at the vertices of a regular hexagon, and we would expect the ground state to include three double bonds between neighboring carbon atoms. Figures 12-a and 12-b represent two possible dispositions of these bonds. The nuclei are assumed here to be fixed because of their high masses. Thus, the electronic states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, associated with Figures 12-a and 12-b respectively, are different. If the structure of Figure 12-a were the only one possible, the ground state of the electronic system would have an energy of $E_m = \langle\varphi_1|H|\varphi_1\rangle$, where H is the Hamiltonian of the electrons in the potential created by the nuclei. But the bonds can also be placed as shown in Figure 12-b. By symmetry, we have $\langle\varphi_2|H|\varphi_2\rangle = \langle\varphi_1|H|\varphi_1\rangle$, and we could conclude that the ground state of the molecule is doubly degenerate. However, the non-diagonal matrix element $\langle\varphi_2|H|\varphi_1\rangle$ of the Hamiltonian H is not zero. This coupling between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ gives rise to two distinct levels, one of which has an energy lower than E_m. The benzene molecule is therefore more stable than we would have expected. Moreover, in its true ground state, the configuration of the molecule cannot be represented either by Figure 12-a or by Figure 12-b: this state is a linear superposition of $|\varphi_1\rangle$ and $|\varphi_2\rangle$ [the coefficients of this superposition having, as in (C-20), the same modulus]. This is what is symbolized by the double arrow of Figure 12, commonly used by chemists.

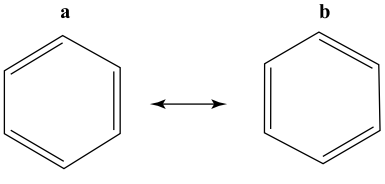


Figure 12: Two possible configurations of the double bonds in a benzene molecule.

Another example is that of the (ionized) molecule H_2^+ , composed of two protons p_1 and p_2 and one electron. The two protons, because of their large masses, can be considered to be fixed. Let us call R the distance between them and $|\varphi_1\rangle$ and $|\varphi_2\rangle$, the states where the electron is localized around p_1 or around p_2, its wave function being that of the hydrogen atom it would form with p_1 or p_2 (Fig. 13). As above, the diagonal elements $\langle\varphi_1|H|\varphi_1\rangle$ and $\langle\varphi_2|H|\varphi_2\rangle$ of the Hamiltonian are equal because of symmetry; we shall denote them by $E_m(R)$. The two states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are not, however, stationary states, since the matrix element $\langle\varphi_1|H|\varphi_2\rangle$ is not zero. Here again, we obtain an energy level lower than $E_m(R)$ and, in the ground state, the wave function of the electron is a linear combination of those of Figures 13-a and 13-b. The electron is thus no longer localized about one of the two protons alone, and it is this delocalization which, by

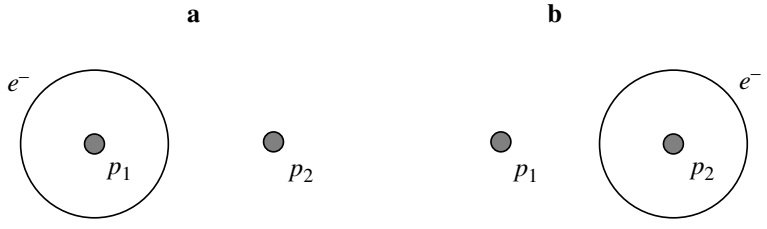


Figure 13: In the H_2^+ ion, the electron could logically be localized either around the proton p_1 (fig. a) or around the proton p_2 (fig. b). In the ground state of the ion, the wave function of the electron is a linear superposition of the wave functions associated with figures a and b. Its probability of presence is symmetrical with respect to the plane bisecting p_1p_2.

lowering its potential energy, is responsible for the chemical bond[8].

C-3. Dynamical aspect: oscillation of the system between the two unperturbed states

C-3-a. Evolution of the state vector

Let:

$$|\psi(t)\rangle = a_1(t)|\varphi_1\rangle + a_2(t)|\varphi_2\rangle \tag{C-22}$$

be the state vector of the system at the instant t. The evolution of $|\psi(t)\rangle$ in the presence of the coupling W is given by the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = (H_0 + W)|\psi(t)\rangle \tag{C-23}$$

Let us project this equation onto the basis vectors $|\varphi_1\rangle$ and $|\varphi_2\rangle$. We obtain, using (C-5) (where we have set $W_{11} = W_{22} = 0$) and (C-22):

$$\begin{aligned} i\hbar\frac{\mathrm{d}}{\mathrm{d}t}a_1(t) &= E_1\ a_1(t) + W_{12}\ a_2(t) \\ i\hbar\frac{\mathrm{d}}{\mathrm{d}t}a_2(t) &= W_{21}\ a_1(t) + E_2\ a_2(t) \end{aligned} \tag{C-24}$$

If $|W_{12}| \neq 0$, these equations form a linear system of homogeneous coupled differential equations. The classical method of solving such a system reduces, in fact, to the application of rule (D-54) of Chapter III: look for the eigenvectors $|\psi_+\rangle$ (eigenvalue E_+) and $|\psi_-\rangle$ (eigenvalue E_-) of the operator $H = H_0 + W$ [whose matrix elements are the coefficients of equations (C-24)], and decompose $|\psi(0)\rangle$ in terms of $|\psi_+\rangle$ and $|\psi_-\rangle$:

$$|\psi(0)\rangle = \lambda|\psi_+\rangle + \mu|\psi_-\rangle \tag{C-25}$$

[8] A more elaborate study of the ionized molecule H_2^+ will be presented in Complement G_{XI}.

(where λ and μ are fixed by the initial conditions). We then have:

$$|\psi(t)\rangle = \lambda\ \mathrm{e}^{-iE_+t/\hbar}|\psi_+\rangle + \mu\ \mathrm{e}^{-iE_-t/\hbar}|\psi_-\rangle \tag{C-26}$$

[which enables us to obtain $a_1(t)$ and $a_2(t)$ by projecting $|\psi(t)\rangle$ onto $|\varphi_1\rangle$ and $|\varphi_2\rangle$].

It can be shown that a system whose state vector is the vector $|\psi(t)\rangle$ given in (C-26) oscillates between the two unperturbed states $|\varphi_1\rangle$ and $|\varphi_2\rangle$. To see this, we shall assume that the system at time $t = 0$ is in the state $|\varphi_1\rangle$:

$$|\psi(0)\rangle = |\varphi_1\rangle \tag{C-27}$$

and calculate the probability $\mathscr{P}_{12}(t)$ of finding it in the state $|\varphi_2\rangle$ at time t.

C-3-b. Calculation of $\mathscr{P}_{12}(t)$: Rabi's formula

As in (C-25), let us expand the ket $|\psi(0)\rangle$ given in (C-27) on the $\{|\psi_+\rangle, |\psi_-\rangle\}$ basis. Inverting formulas (C-9), we obtain:

$$|\psi(0)\rangle = |\varphi_1\rangle = \mathrm{e}^{i\varphi/2}\left[\cos\frac{\theta}{2}|\ \psi_+\rangle - \sin\frac{\theta}{2}\ |\psi_-\rangle\right] \tag{C-28}$$

from which we deduce, using (C-26):

$$|\psi(t)\rangle = \mathrm{e}^{i\varphi/2}\left[\cos\frac{\theta}{2}\ \mathrm{e}^{-iE_+t/\hbar}|\psi_+\rangle - \sin\frac{\theta}{2}\ \mathrm{e}^{-iE_-t/\hbar}|\psi_-\rangle\right] \tag{C-29}$$

The probability amplitude of finding the system at time t in the state $|\varphi_2\rangle$ is then written:

$$\begin{aligned}\langle\varphi_2|\psi(t)\rangle &= \mathrm{e}^{i\varphi/2}\left[\cos\frac{\theta}{2}\ \mathrm{e}^{-iE_+t/\hbar}\langle\varphi_2|\psi_+\rangle - \sin\frac{\theta}{2}\ \mathrm{e}^{-iE_-t/\hbar}\langle\varphi_2|\psi_-\rangle\right]\\ &= \mathrm{e}^{i\varphi/2}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\ [\mathrm{e}^{-iE_+t/\hbar} - \mathrm{e}^{-iE_-t/\hbar}]\end{aligned} \tag{C-30}$$

which enables us to calculate $\mathscr{P}_{12}(t) = |\langle\varphi_2|\psi(t)\rangle|^2$. We thus find:

$$\begin{aligned}\mathscr{P}_{12}(t) &= \frac{1}{2}\sin^2\theta\left[1 - \cos\left(\frac{E_+ - E_-}{\hbar}t\right)\right]\\ &= \sin^2\theta\sin^2\left(\frac{E_+ - E_-}{2\hbar}t\right)\end{aligned} \tag{C-31}$$

or, using expressions (C-12) and (C-13):

$$\mathscr{P}_{12}(t) = \frac{4|W_{12}|^2}{4|W_{12}|^2 + (E_1 - E_2)^2}\sin^2\left[\sqrt{4|W_{12}|^2 + (E_1 - E_2)^2}\ \frac{t}{2\hbar}\right] \tag{C-32}$$

Formula (C-32) is sometimes called Rabi's formula.

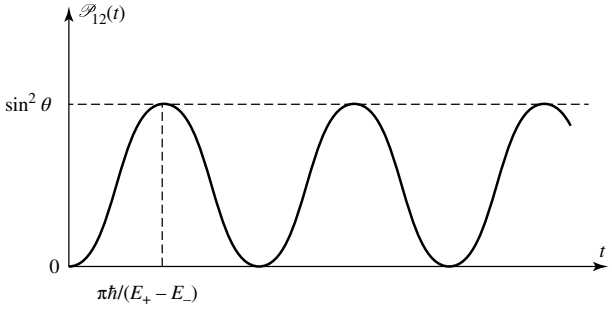


Figure 14: Variation with respect to time of the probability $\mathscr{P}_{12}(t)$ of finding the system in the state $|\varphi_2\rangle$ when it was initially in the state $|\varphi_1\rangle$. When the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ have the same unperturbed energy, the probability $\mathscr{P}_{12}(t)$ can attain the value 1.

C-3-c. Discussion

Relation (C-31) shows that the probability $\mathscr{P}_{12}(t)$ oscillates over time with a frequency of $(E_+ - E_-)/h$, which is simply the unique Bohr frequency of the system. $\mathscr{P}_{12}(t)$ varies between zero and a maximum value which, according to (C-31), is equal to $\sin^2\theta$. This maximum value is attained for all values of t such that $t = (2k+1)h/2(E_+ - E_-)$, with $k = 0, 1, 2, \ldots$ (Fig. 14).

The oscillation frequency $(E_+ - E_-)/h$, as well as the maximum value $\sin^2\theta$ of $\mathscr{P}_{12}(t)$, are functions of $|W_{12}|$ and $E_1 - E_2$, whose main features we are now going to describe.

When $E_1 = E_2$, $(E_+ - E_-)/h$ is equal to $2|W_{12}|/h$, and $\sin^2\theta$ takes on its greatest possible value, that is, 1: at certain times, $t = (2k+1)\pi\hbar/2|W_{12}|$, the system (which started from the state $|\varphi_1\rangle$) is in the state $|\varphi_2\rangle$. Therefore, any coupling between two states of equal energy causes the system to oscillate completely from one state to the other with a frequency proportional to the coupling[9].

When $E_1 - E_2$ increases, so does $(E_+ - E_-)/h$, while $\sin^2\theta$ decreases. For a weak coupling ($E_1 - E_2 \gg |W_{12}|$), $E_+ - E_-$ differs very little from $E_1 - E_2$, and $\sin^2\theta$ becomes very small. This last result is not surprising since, in the case of a weak coupling, the state $|\varphi_1\rangle$ is very close to the stationary state $|\psi_+\rangle$ [*cf.* formulas (C-21)]: the system, having started in the state $|\varphi_1\rangle$, evolves very little over time.

C-3-d. Example of oscillation between two states

Let us return to the example of the H_2^+ molecule. We shall assume that, at a certain time, the electron is localized about proton p_1: it is, for example, in the state shown in Figure 13-a. According to the results of the preceding section, we know that it will oscillate between the two protons with a frequency equal to the Bohr frequency associated

[9] The same phenomenon is found in other domains of physics. Consider, for example, two identical pendulums (1) and (2), suspended from the same support and having the same frequency. Let us assume that at time $t = 0$, pendulum (1) is set in motion. The coupling is ensured by their common support. We then know (*cf.* Complement H_V) that, after a certain time (which decreases if the coupling is increased), we arrive at a situation where only pendulum (2) oscillates, with the initial amplitude of pendulum (1). Then the motion is transferred back to pendulum (1), and so on.

with the two stationary states $|\psi_+\rangle$ and $|\psi_-\rangle$ of the molecule. To this oscillation of the electron between the two states, represented in 13-a and 13-b, corresponds an oscillation of the mean value of the electric dipole moment of the molecule (the dipole moment is non-zero when the electron is localized about one of the two protons, and changes sign depending on whether the proton involved is p_1 or p_2). Thus we see concretely how, when the molecule is not in a stationary state, an oscillating electric dipole moment can appear. Such an oscillating dipole can exchange energy with an electromagnetic wave of the same frequency. Consequently, this frequency must appear in the absorption and emission spectrum of the H_2^+ ion.

Other examples of oscillations between two states are discussed in Complements F_{IV}, G_{IV} and H_{IV}.

References and suggestions for further reading:

The Stern-Gerlach experiment: original article (3.8); Cagnac and Pebay-Peyroula (11.2), Chap. X; Eisberg and Resnick (1.3), § 8-3; Bohm (5.1), §§ 22.5 and 22.6; Frisch (3.13).

Two-level systems: Feynman III (1.2), Chaps. 6, 10 and 11; Valentin (16.1), Annexe XII; Allen and Eberly (15.8), particularly Chap. 3.

A_{IV} : THE PAULI MATRICES **B_{IV} : DIAGONALISATION OF A 2×2 HERMITIAN MATRIX**	Technical study of 2 × 2 matrices; simple, and important for solving numerous quantum mechanical problems.
C_{IV} : FICTITIOUS SPIN 1/2 ASSOCIATED WITH A TWO-LEVEL SYSTEM	Establishes the close relation that exists between §§ B and C of Chapter IV; supplies a simple geometrical interpretation of the properties of two-level systems (easy, but not indispensable for what follows).
D_{IV} : SYSTEM OF TWO SPIN 1/2 PARTICLES	Simple illustration of the tensor product and of the postulates of quantum mechanics (can be considered to be a set of worked exercises).
E_{IV} : SPIN 1/2 DENSITY MATRIX	Illustration, in the case of spin 1/2 particles, of the concepts introduced in Complement E_{III}.
F_{IV} : SPIN 1/2 PARTICLE IN A STATIC MAGNETIC FIELD AND A ROTATING FIELD: MAGNETIC RESONANCE	Study of a very important physical phenomenon with many applications: magnetic resonance. Can be studied later.
G_{IV} : A SIMPLE MODEL OF THE AMMONIA MOLECULE	Example of a physical system whose study can be reduced, in a first approximation, to that of a two-level system; moderately difficult.
H_{IV} : COUPLING BETWEEN A STABLE STATE AND AN UNSTABLE STATE	Study of the influence of coupling between two levels with different lifetimes; easy, but requires the concepts introduced in Complement K_{III}.
J_{IV} : EXERCISES	

Complement A_{IV}

The Pauli matrices

1 **Definition; eigenvalues and eigenvectors** 425
2 **Simple properties** 426
3 **A convenient basis of the 2×2 matrix space** 427

In § A-2 of Chapter IV, we introduced the matrices representing the three components S_x, S_y and S_z of a spin S in the $|+\rangle, |-\rangle$ basis (eigenvectors of S_z). In quantum mechanics, it is often convenient to introduce the dimensionless operator $\boldsymbol{\sigma}$, proportional to $\mathbf{S}$, and given by:

$$\mathbf{S} = \frac{\hbar}{2}\,\boldsymbol{\sigma} \tag{1}$$

The matrices representing the three components of $\boldsymbol{\sigma}$ in the $\{|+\rangle, |-\rangle\}$ basis are called the "Pauli matrices".

1. Definition; eigenvalues and eigenvectors

Let us go back to equations (A-15), (A-16) and (A-17) of Chapter IV. Using (1), we see that the definition of the Pauli matrices is:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2}$$

These are Hermitian matrices, all three of which have the same characteristic equation:

$$\lambda^2 - 1 = 0 \tag{3}$$

The eigenvalues of σ_x, σ_y and σ_z are therefore:

$$\lambda = \pm 1 \tag{4}$$

which is consistent with the fact that those of S_x, S_y and S_z are $\pm\hbar/2$.

It is easy to obtain, from definition (2), the eigenvectors of σ_x, σ_y and σ_z. They are the same, respectively, as those of S_x, S_y and S_z, already introduced in § A-2 of Chapter IV:

$$\begin{aligned} \sigma_x|\pm\rangle_x &= \pm|\pm\rangle_x \\ \sigma_y|\pm\rangle_y &= \pm|\pm\rangle_y \\ \sigma_z|\pm\rangle &= \pm|\pm\rangle \end{aligned} \tag{5}$$

with:

$$|\pm\rangle_x = \frac{1}{\sqrt{2}}\,[|+\rangle \pm |-\rangle]$$
$$|\pm\rangle_y = \frac{1}{\sqrt{2}}\,[|+\rangle \pm i|-\rangle] \tag{6}$$

2. Simple properties

It is easy to see from definition (2) that the Pauli matrices verify the relations:

$$\mathrm{Det}(\sigma_j) = -1 \quad j = x, y \text{ or } z \tag{7}$$
$$\mathrm{Tr}(\sigma_j) = 0 \tag{8}$$
$$(\sigma_x^2) = (\sigma_y^2) = (\sigma_z^2) = I \quad (\text{where } I \text{ is the } 2 \times 2 \text{ unit matrix}) \tag{9}$$
$$\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z \tag{10}$$

as well as the equations that can be deduced from (10) by cyclic permutation of x, y and z.

Equations (9) and (10) are sometimes condensed into the form:

$$\sigma_j\sigma_k = i\sum_l \varepsilon_{jkl}\ \sigma_l + \delta_{jk}\ I \tag{11}$$

where ε_{jkl} is antisymmetric with respect to the interchange of any two of its indices. It is equal to:

$$\varepsilon_{jkl} = \begin{cases} 0 & \text{if} \quad \text{the indices} \quad j,k,l \quad \text{are not all different} \\ 1 & \text{if} \quad j,k,l \quad \text{is an even permutation of} \quad x,y,z \\ -1 & \text{if} \quad j,k,l \quad \text{is an odd permutation of} \quad x,y,z \end{cases} \tag{12}$$

From (10), we immediately conclude:

$$[\sigma_x, \sigma_y] = 2i\sigma_z \tag{13}$$

(and the relations obtained by cyclic permutation). This yields:

$$[S_x, S_y] = i\hbar S_z$$
$$[S_y, S_z] = i\hbar S_x \tag{14}$$
$$[S_z, S_x] = i\hbar S_y$$

We shall see later (*cf.* Chap. VI) that equations (14) are characteristic of an angular momentum.

We also see from (10) that:

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0 \tag{15}$$

(the σ_i matrices are said to anticommute with each other) and that, taking (9) into account:

$$\sigma_x\sigma_y\sigma_z = i\ I \tag{16}$$

Finally, let us mention an identity which is sometimes useful in quantum mechanics. If $\mathbf{A}$ and $\mathbf{B}$ denote two vectors whose components are numbers (or operators which commute with all operators acting in the two-dimensional spin state space):

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B}\ I + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B}) \tag{17}$$

This identity can be demonstrated as follows. Using formula (11) and the fact that $\mathbf{A}$ and $\mathbf{B}$ commute with $\boldsymbol{\sigma}$, we can write:

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) &= \sum_{j,k}\sigma_j A_j\ \sigma_k B_k \\ &= \sum_{j,k} A_j B_k \left[\sum_l \varepsilon_{jkl}\ \sigma_l + \delta_{jk} I\right] \\ &= i\sum_l \sigma_l \left[\sum_{j,k}\varepsilon_{jkl}\ A_j B_k\right] + \sum_j A_j B_j\ I\end{aligned} \tag{18}$$

In the second term, we recognize the scalar product $\mathbf{A}\cdot\mathbf{B}$. In addition, it is easy to see from (12) that $\sum_{j,k}\varepsilon_{jkl}A_jB_k$ is the lth component of the vector product $\mathbf{A}\times\mathbf{B}$. This proves (17). Note that if $\mathbf{A}$ and $\mathbf{B}$ do not commute, they must appear in the same order on both sides of the identity.

3. A convenient basis of the 2×2 matrix space

Consider an arbitrary 2×2 matrix:

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22}\end{pmatrix} \tag{19}$$

It can always be written as a linear combination of the four matrices:

$$I,\ \sigma_x,\ \sigma_y,\ \sigma_z \tag{20}$$

since, using (2), we can immediately verify that:

$$M = \frac{m_{11}+m_{22}}{2}\ I + \frac{m_{11}-m_{22}}{2}\ \sigma_z + \frac{m_{12}+m_{21}}{2}\ \sigma_x + i\ \frac{m_{12}-m_{21}}{2}\ \sigma_y \tag{21}$$

Therefore, any 2×2 matrix can be put in the form:

$$M = a_0\ I + \mathbf{a}\cdot\boldsymbol{\sigma} \tag{22}$$

where the coefficients a_0, a_x, a_y and a_z are complex numbers.

Comparing (21) and (22), we see that M is Hermitian if and only if the coefficients a_0 and $\mathbf{a}$ are real. These coefficients can be expressed formally in terms of the matrix M in the following manner:

$$a_0 = \frac{1}{2}\mathrm{Tr}\{M\} \tag{23a}$$

$$\mathbf{a} = \frac{1}{2}\mathrm{Tr}\{M\boldsymbol{\sigma}\} \tag{23b}$$

These formulas can easily be proven from (8), (9) and (10).

Complement B_{IV}

Diagonalization of a 2×2 Hermitian matrix

1 Introduction 429
2 Changing the eigenvalue origin 429
3 Calculation of the eigenvalues and eigenvectors 430
3-a Angles θ and φ 430
3-b Eigenvalues of K 431
3-c Eigenvalues of H 431
3-d Normalized eigenvectors of H 432

1. Introduction

In quantum mechanics, one must often diagonalize 2×2 matrices. When we need only the eigenvalues, it is very easy to solve the characteristic equation since it is of second degree. In principle, the calculation of the normalized eigenvectors is also extremely simple; however, if it is performed clumsily, it can lead to expressions that are unnecessarily complicated and difficult to handle.

The goal of this complement is to present a simple method of calculation which is applicable in all cases. After having changed the origin of the eigenvalues, we introduce the angles θ and φ, defined in terms of the matrix elements, which enable us to write the normalized eigenvectors in a simple easy-to-use form. The angles θ and φ also have an interesting physical interpretation in the study of two-level systems, as we shall see in Complement C_{IV}.

2. Changing the eigenvalue origin

Consider the Hermitian matrix:

$$(H) = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \tag{1}$$

H_{11} and H_{22} are real. Moreover:

$$H_{12} = H_{21}^* \tag{2}$$

The matrix (H) therefore represents, in an orthonormal basis $\{|\varphi_1\rangle, |\varphi_2\rangle\}$, a certain Hermitian operator[1] H.

[1] We use the letter H because the Hermitian operator which we are trying to diagonalize is often a Hamiltonian. Nevertheless, the calculations presented in this complement can obviously be applied to any 2×2 Hermitian matrix.

Using the half-sum and half-difference of the diagonal elements H_{11} and H_{22}, we can write (H) in the following way:

$$(H) = \begin{pmatrix} \frac{1}{2}(H_{11}+H_{22}) & 0 \\ 0 & \frac{1}{2}(H_{11}+H_{22}) \end{pmatrix} + \begin{pmatrix} \frac{1}{2}(H_{11}-H_{22}) & H_{12} \\ H_{21} & -\frac{1}{2}(H_{11}-H_{22}) \end{pmatrix} \quad (3)$$

It follows that the operator H itself can be decomposed into:

$$H = \frac{1}{2}(H_{11}+H_{22})\mathbb{1} + \frac{1}{2}(H_{11}-H_{22})K \quad (4)$$

where $\mathbb{1}$ is the identity operator and K is the Hermitian operator represented in the $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ basis by the matrix:

$$(K) = \begin{pmatrix} 1 & \frac{2H_{12}}{H_{11}-H_{22}} \\ \frac{2H_{21}}{H_{11}-H_{22}} & -1 \end{pmatrix} \quad (5)$$

It is clear from (4) that H and K have the same eigenvectors. Let $|\psi_\pm\rangle$ be these eigenvectors, and $E_\pm$ and $\kappa_\pm$, the corresponding eigenvalues for H and K:

$$H|\psi_\pm\rangle = E_\pm|\psi_\pm\rangle \quad (6)$$
$$K|\psi_\pm\rangle = \kappa_\pm|\psi_\pm\rangle \quad (7)$$

From (4), we immediately conclude that:

$$E_\pm = \frac{1}{2}(H_{11}+H_{22}) + \frac{1}{2}(H_{11}-H_{22})\kappa_\pm \quad (8)$$

Actually, the first matrix appearing on the right-hand side of (3) plays a minor role: we could make it disappear by choosing $(H_{11}+H_{22})/2$ for the eigenvalue origin[2].

3. Calculation of the eigenvalues and eigenvectors

3-a. Angles θ and φ

Let θ and φ be the angles defined in terms of the matrix elements H_{ij} by:

$$\tan\theta = \frac{2|H_{21}|}{H_{11}-H_{22}} \qquad \text{with} \quad 0 \le \theta < \pi \quad (9)$$

[2]Furthermore, this new origin is the same, whatever the basis $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ initially chosen, since $H_{11}+H_{22} = \mathrm{Tr}(H)$ is invariant under a change of orthonormal bases.

$$H_{21} = |H_{21}|\ \mathrm{e}^{i\varphi} \qquad \text{with} \quad 0 \leq \varphi < 2\pi \tag{10}$$

φ is the argument of the complex number H_{21}. According to (2), we have $|H_{12}| = |H_{21}|$ and:

$$H_{12} = |H_{12}|\ \mathrm{e}^{-i\varphi} \tag{11}$$

If we use (9), (10) and (11), the matrix (K) becomes:

$$(K) = \begin{pmatrix} 1 & \tan\,\theta\ \mathrm{e}^{-i\varphi} \\ \tan\,\theta\ \mathrm{e}^{i\varphi} & -1 \end{pmatrix} \tag{12}$$

3-b. Eigenvalues of K

The characteristic equation of the matrix (12):

$$\mathrm{Det}[(K) - \kappa I] = \kappa^2 - 1 - \tan^2\theta = 0 \tag{13}$$

directly yields the eigenvalues κ_+ and κ_- of (K):

$$\kappa_+ = +\frac{1}{\cos\theta} \tag{14a}$$

$$\kappa_- = -\frac{1}{\cos\theta} \tag{14b}$$

We see that they are indeed real (property of a Hermitian matrix, *cf.* § D-2-a of Chapter II). If we want to express $1/\cos\theta$ in terms of H_{ij}, all we need to do is use (9) and notice that $\cos\theta$ and $\tan\theta$ have the same sign since $0 \leq \theta < \pi$:

$$\frac{1}{\cos\theta} = \frac{\sqrt{(H_{11} - H_{22})^2 + 4|H_{12}|^2}}{H_{11} - H_{22}} \tag{15}$$

3-c. Eigenvalues of H

Using (8), (14) and (15), we immediately obtain:

$$E_+ = \frac{1}{2}(H_{11} + H_{22}) + \frac{1}{2}\sqrt{(H_{11} - H_{22})^2 + 4|H_{12}|^2} \tag{16a}$$

$$E_- = \frac{1}{2}(H_{11} + H_{22}) - \frac{1}{2}\sqrt{(H_{11} - H_{22})^2 + 4|H_{12}|^2} \tag{16b}$$

Comments:

(*i*) As we have already said, the eigenvalues (16) can easily be obtained from the characteristic equation of the matrix (H). If we need only the eigenvalues of (H), it is therefore not necessary to introduce the angles θ and φ as we have done here. On the other hand, we shall see in the following section that this method is very useful when we need to use the normalized eigenvectors of H.

(*ii*) It can be verified immediately from formulas (16) that:

$$E_+ + E_- = H_{11} + H_{22} = \mathrm{Tr}(H) \tag{17}$$

$$E_+ E_- = H_{11} H_{22} - |H_{l2}|^2 = \mathrm{Det}(H) \tag{18}$$

(*iii*) To have $E_+ = E_-$, we must have $(H_{11} - H_{22})^2 + 4|H_{12}|^2 = 0$; that is, $H_{11} = H_{22}$ and $H_{12} = H_{21} = 0$. A 2×2 Hermitian matrix with a degenerate spectrum is therefore necessarily proportional to the unit matrix.

3-d. Normalized eigenvectors of H

Let a and b be the components of $|\psi_+\rangle$ on $|\varphi_1\rangle$ and $|\varphi_2\rangle$. According to (7), (12) and (14a), they must satisfy:

$$\begin{pmatrix} 1 & \tan\theta\ \mathrm{e}^{-i\varphi} \\ \tan\theta\ \mathrm{e}^{i\varphi} & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\cos\theta} \begin{pmatrix} a \\ b \end{pmatrix} \tag{19}$$

which yields:

$$\left(1 - \frac{1}{\cos\theta}\right) a + \tan\,\theta\ \mathrm{e}^{-i\varphi} b = 0 \tag{20}$$

that is:

$$-\left(\sin\frac{\theta}{2}\ \mathrm{e}^{i\varphi/2}\right) a + \left(\cos\frac{\theta}{2}\ \mathrm{e}^{-i\varphi/2}\right) b = 0 \tag{21}$$

The normalized eigenvector $|\psi_+\rangle$ can therefore be written:

$$|\psi_+\rangle = \cos\frac{\theta}{2}\ \mathrm{e}^{-i\varphi/2}|\varphi_1\rangle + \sin\frac{\theta}{2}\ \mathrm{e}^{i\varphi/2}|\varphi_2\rangle \tag{22}$$

An analogous calculation would yield:

$$|\psi_-\rangle = -\sin\frac{\theta}{2}\ \mathrm{e}^{-i\varphi/2}|\varphi_1\rangle + \cos\frac{\theta}{2}\ \mathrm{e}^{i\varphi/2}|\varphi_2\rangle \tag{23}$$

It can be verified that $|\psi_+\rangle$ and $|\psi_-\rangle$ are orthogonal.

Comment:

While the trigonometric functions of the angle θ can be expressed rather simply in terms of the matrix elements H_{ij} [see, for example, formulas (9) and (15)], the same is not true of those of the angle $\theta/2$. Consequently, formulas (22) and (23) for the normalized eigenvectors $|\psi_+\rangle$ and $|\psi_-\rangle$ become complicated when $\cos\theta/2$ and $\sin\theta/2$ are replaced by their expressions in terms of H_{ij}; they are no longer very convenient. It is better to use expressions (22) and (23) directly, keeping the functions $\cos\theta/2$ and $\sin\theta/2$ during the entire calculation involving the normalized eigenvectors of H. Furthermore, the final result of the calculation often

involves only functions of the angle θ (see, for example, the calculation of § C-3-b of Chapter IV) and, consequently, can be expressed simply in terms of the H_{ij}. Expressions (22) and (23) thus enable us to carry out the intermediate calculations elegantly, avoiding unnecessarily complicated expressions. This is the advantage of the method presented in this complement. Another advantage concerns the physical interpretation and will be discussed in the next complement.

Complement C_{IV}

Fictitious spin 1/2 associated with a two-level system

1 Introduction 435
2 Interpretation of the Hamiltonian in terms of fictitious spin 435
3 Geometrical interpretation 437
3-a Fictitious magnetic fields associated with H_0, W and H . . . 437
3-b Effect of coupling on the eigenvalues and eigenvectors of the Hamiltonian 438
3-c Geometrical interpretation of $\mathscr{P}_{12}(t)$ 439

1. Introduction

Consider a two-level system whose Hamiltonian H is represented, in an orthonormal basis $\{|\varphi_1\rangle, |\varphi_2\rangle\}$, by the Hermitian matrix (H) [formula (1) of Complement B_{IV}][1]. If we choose $(H_{11} + H_{22})/2$ as the new energy origin, the matrix (H) becomes:

$$(H) = \begin{pmatrix} \frac{1}{2}(H_{11} - H_{22}) & H_{12} \\ H_{21} & -\frac{1}{2}(H_{11} - H_{22}) \end{pmatrix} \tag{1}$$

Although the two-level system under consideration is not necessarily a spin 1/2, we can always associate with it a spin 1/2 whose Hamiltonian H is represented by the same matrix (H) in the $|+\rangle, |-\rangle$ basis of eigenstates of the S_z component of this spin. We shall see that (H) can then be interpreted as describing the interaction of this "fictitious spin" with a static magnetic field $\mathbf{B}$, whose direction and modulus are very simply related to the parameters introduced in the preceding complement in the discussion of the diagonalization of (H). Thus it is possible to give a simple physical meaning to these parameters.

Moreover, if the Hamiltonian H is the sum $H = H_0 + W$ of two operators, we shall see that we can associate with H, H_0 and W three magnetic fields, $\mathbf{B}$, $\mathbf{B}_0$ and $\mathbf{b}$, such that $\mathbf{B} = \mathbf{B}_0 + \mathbf{b}$. Introducing the coupling W is equivalent, in terms of fictitious spin, to adding the field $\mathbf{b}$ to $\mathbf{B}_0$. We shall show that this point of view enables us to interpret very simply the different effects studied in § C of Chapter IV.

2. Interpretation of the Hamiltonian in terms of fictitious spin

We saw in Chapter IV that the Hamiltonian $\widetilde{H}$ of the coupling between a spin 1/2 and a magnetic field $\mathbf{B}$, of components B_x, B_y, B_z can be written:

$$\widetilde{H} = -\gamma \mathbf{B}.\mathbf{S} = -\gamma(B_x S_x + B_y S_y + B_z S_z) \tag{2}$$

[1] We are using the same notation as in Complement B_{IV} and Chapter IV.

To calculate the matrix associated with this operator, we substitute into this relation the matrices associated with S_x, S_y, S_z [Chap. IV, relations (A-15), (A-16), (A-17)]. This immediately yields:

$$(\widetilde{H}) = -\frac{\gamma\hbar}{2}\begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix} \tag{3}$$

Therefore, to make matrix (1) identical to $(\widetilde{H})$, we must simply choose a "fictitious field" $\mathbf{B}$ defined by:

$$\begin{cases} B_x = -\dfrac{2}{\gamma\hbar}\,\mathrm{Re}(H_{12}) \\ B_y = \dfrac{2}{\gamma\hbar}\,\mathrm{Im}(H_{12}) \\ B_z = \dfrac{1}{\gamma\hbar}(H_{22} - H_{11}) \end{cases} \tag{4}$$

Note that the modulus $B_\perp$ of the projection $\mathbf{B}_\perp$ of $\mathbf{B}$ onto the xOy plane is then equal to:

$$B_\perp = \frac{2}{\hbar}\left|\frac{H_{12}}{\gamma}\right| \tag{5}$$

According to formulas (9) and (10) of Complement B_{IV}, the angles θ and φ associated with the matrix $(H) = (\widetilde{H})$ written in (3) are given by:

$$\begin{cases} \mathrm{tg}\,\theta = \dfrac{|\gamma B_\perp|}{-\gamma B_z} & 0 \leq \theta < \pi \\ -\gamma(B_x + iB_y) = |\gamma B_\perp| e^{i\varphi} & 0 \leq \theta < 2\pi \end{cases} \tag{6}$$

The gyromagnetic ratio γ is a simple calculation tool and can have an arbitrary value. If we agree to choose γ negative, relations (6) show that the angles θ and φ associated with the matrix (H) are simply the polar angles of the direction of the field $\mathbf{B}$ (if we had chosen γ positive, they would be those of the opposite direction).

Finally, we see that we can forget the two-level system with which we started and consider the matrix (H) as representing, in the basis of the eigenstates $|+\rangle$ and $|-\rangle$ of S_z, the Hamiltonian $\widetilde{H}$ of a spin 1/2 placed in a field $\mathbf{B}$ whose components are given by (4). $\widetilde{H}$ can also be written:

$$\widetilde{H} = \omega S_u$$

where S_u is the operator $\mathbf{S}.\mathbf{u}$ which describes the spin component along the direction $\mathbf{u}$, whose polar angles are θ and φ, and ω is the Larmor angular velocity:

$$\omega = |\gamma||\mathbf{B}| \tag{7}$$

The following table summarizes the various correspondences between the two-level system and the associated fictitious spin 1/2.

Two-level system		Fictitious spin 1/2
$\|\varphi_1\rangle$	$\longleftrightarrow$	$\|+\rangle$
$\|\varphi_2\rangle$	$\longleftrightarrow$	$\|-\rangle$
$\|\psi_+\rangle$	$\longleftrightarrow$	$\|+\rangle_u$
$\|\psi_-\rangle$	$\longleftrightarrow$	$\|-\rangle_u$
$E_+ - E_-$	$\longleftrightarrow$	$\hbar\omega$
Angles θ and φ introduced in $\mathrm{B_{IV}}$	$\longleftrightarrow$	Polar angles of the fictitious field $\mathbf{B}$
$H_{11} - H_{22}$	$\longleftrightarrow$	$-\gamma\hbar B_z$
$\|H_{21}\|$	$\longleftrightarrow$	$-\gamma\hbar B_\perp/2$

3. Geometrical interpretation of the various effects discussed in § C of Chapter IV

3-a. Fictitious magnetic fields associated with H_0, W and H

Assume, as in § C of Chapter IV, that H appears as the sum of two terms:

$$H = H_0 + W \tag{8}$$

In the $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ basis, the unperturbed Hamiltonian H_0 is represented by a diagonal matrix which, with a suitable choice of the energy origin, is written:

$$H_0 = \begin{pmatrix} \dfrac{E_1 - E_2}{2} & 0 \\ 0 & -\dfrac{E_1 - E_2}{2} \end{pmatrix} \tag{9}$$

As far as the coupling W is concerned, we assume, as in § C of Chapter IV, that it is purely non-diagonal:

$$(W) = \begin{pmatrix} 0 & W_{12} \\ W_{12} & 0 \end{pmatrix} \tag{10}$$

The discussion of the preceding section then enables us to associate with (H_0) and (W) two fields $\mathbf{B}_0$ and $\mathbf{b}$ such that [*cf.* formulas (4) and (5)]:

$$\begin{cases} B_{0z} = \dfrac{E_2 - E_1}{\gamma\hbar} \\ B_{0\perp} = 0 \end{cases} \tag{11}$$

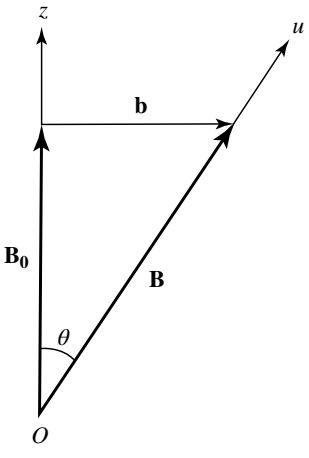


Figure 1: Relative disposition of the fictitious fields. $\mathbf{B}_0$ *is associated with* H_0, $\mathbf{b}$ *with* W, *and* $\mathbf{B} = \mathbf{B}_0 + \mathbf{b}$ *with the total Hamiltonian* $H = H_0 + W$.

$$\begin{cases} b_z = 0 \\ b_\perp = \dfrac{2}{\hbar}\left|\dfrac{W_{12}}{\gamma}\right| \end{cases} \tag{12}$$

$\mathbf{B}_0$ is therefore parallel to Oz and proportional to $(E_1 - E_2)/2$; $\mathbf{b}$ is perpendicular to Oz and proportional to $|W_{12}|$. Since $(H) = (H_0) + (W)$, the field $\mathbf{B}$ associated with the total Hamiltonian is the vector sum of $\mathbf{B}_0$ and $\mathbf{b}$:

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{b} \tag{13}$$

The three fields $\mathbf{B}_0$, $\mathbf{b}$ and $\mathbf{B}$ are shown in Figure 1; the angle θ introduced in § C-2-a of Chapter IV is the angle between $\mathbf{B}_0$ and $\mathbf{B}$, since $\mathbf{B}_0$ is parallel to Oz.

The strong coupling condition introduced in § C-2 of Chap. IV ($|W_{12}| \gg |E_1 - E_2|$) is equivalent to $|\mathbf{b}| \gg |\mathbf{B}_0|$ (Fig. 2-a). The weak coupling condition ($|W_{12}| \ll |E_1 - E_2|$) is equivalent to $|\mathbf{b}| \ll |\mathbf{B}_0|$ (Fig. 2-b).

3-b. Effect of coupling on the eigenvalues and eigenvectors of the Hamiltonian

$E_1 - E_2$ and $E_+ - E_-$ correspond respectively to the Larmor angular velocities $\omega_0 = |\gamma||\mathbf{B}_0|$ and $\omega = |\gamma||\mathbf{B}|$ in the fields $\mathbf{B}_0$ and $\mathbf{B}$. We see in Figure 1 that $\mathbf{B}_0$, $\mathbf{b}$ and $\mathbf{B}$ form a right triangle whose hypotenuse is $\mathbf{B}$; we therefore always have $|\mathbf{B}| \geq |\mathbf{B}_0|$, which again shows that $E_+ - E_-$ is always greater than $|E_1 - E_2|$.

For a weak coupling (Fig. 2-b), the difference between $|\mathbf{B}|$ and $|\mathbf{B}_0|$ is very small in relative value, being of second order in $|\mathbf{b}|/|\mathbf{B}_0|$. From this we deduce immediately that $E_+ - E_-$ and $E_1 - E_2$ differ in relative value by terms of second order in $|W_{12}|/(E_1 - E_2)$. On the other hand, for a strong coupling (Fig. 2-a), $|\mathbf{B}|$ is much larger than $|\mathbf{B}_0|$ and practically equal to $|\mathbf{b}|$; $E_+ - E_-$ is then much larger than $|E_1 - E_2|$ and practically proportional to $|W_{12}|$. We thus find again all the results of § C-2 of Chapter IV.

As far as the effect of the coupling on the eigenvectors is concerned, it can also be understood very simply from Figures 1 and 2. The eigenvectors of H and H_0 are

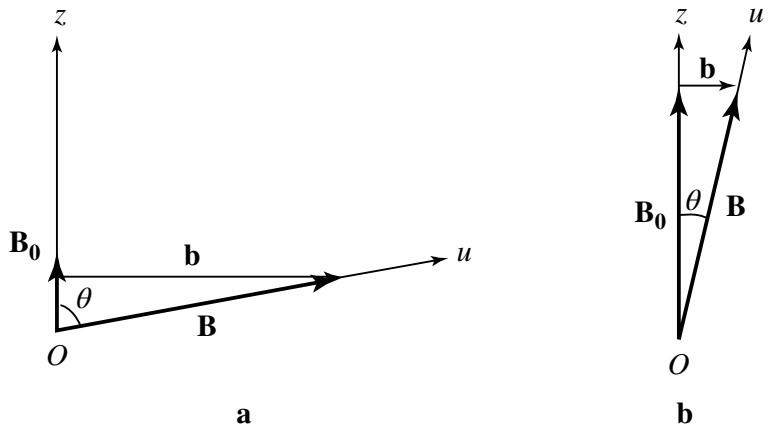


Figure 2: Relative disposition of the fictitious fields $\mathbf{B}_0$, $\mathbf{b}$ *and* $\mathbf{B}$ *in the case of strong coupling (fig. a) and weak coupling (fig. b).*

associated respectively with the eigenvectors of the components of $\mathbf{S}$ on the Ou and Oz axes. These two axes are practically parallel in the case of weak coupling (Fig. 2-b) and perpendicular in the case of strong coupling (Fig. 2-a). The eigenvectors of S_u and S_z, and, consequently, those of H and H_0, are very close in the first case and very different in the second one.

3-c. Geometrical interpretation of $\mathscr{P}_{12}(t)$

In terms of fictitious spin, the problem considered in § C-3 of Chapter IV can be put in the following way: at time $t = 0$, the fictitious spin associated with the two-level system is in the eigenstate $|+\rangle$ of S_z; $\mathbf{b}$ is added to $\mathbf{B}_0$; what is the probability $\mathscr{P}_{+-}(t)$ of finding the spin in the state $|-\rangle$ at time t? With the correspondences summarized in the table, $\mathscr{P}_{12}(t)$ must be identical to $\mathscr{P}_{+-}(t)$.

The calculation of $\mathscr{P}_{+-}(t)$ is then very simple since the time evolution of the spin reduces to a Larmor precession about $\mathbf{B}$ (Fig. 3). During this precession, the angle θ between the spin and the direction Ou of $\mathbf{B}$ remains constant. At time t, the spin points in the direction On, making an angle α with Oz; the angle formed by the (Oz, Ou) and (Ou, On) planes is equal to ωt. A classical formula of spherical trigonometry enables us to write:

$$\cos\alpha = \cos^2\theta + \sin^2\theta\cos\omega t \tag{14}$$

Now, when the spin points in a direction that makes an angle of α with Oz, the probability of finding it in the state $|-\rangle$ of S_z is equal (*cf.* § B-2-b of Chapter IV) to $\sin^2\alpha/2 = (1-\cos\alpha)/2$. From this we deduce, using (14), that:

$$\mathscr{P}_{+-}(t) = \sin^2\frac{\alpha}{2} = \frac{1}{2}\ \sin^2\theta\ (1-\cos\omega t) \tag{15}$$

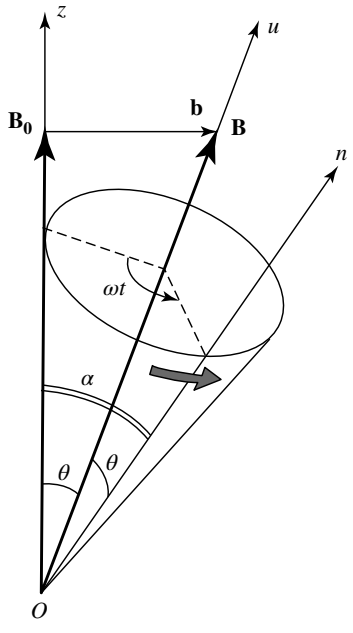


Figure 3: Geometrical interpretation of Rabi's formula in terms of fictitious spin. Under the effect of the coupling (represented by $\mathbf{b}$*), the spin, initially oriented along* Oz*, precesses about* $\mathbf{B}$*; consequently, the probability of finding* $-\hbar/2$ *in a measurement of its* S_z *component on* Oz *is an oscillating function of time.*

This result is identical, when we replace ω by $(E_+ - E_-)/\hbar$, to formula (C-31) of Chapter IV (Rabi's formula). We have thus given this formula a purely geometrical interpretation.

References and suggestions for further reading:

Abragam (14.1), Chap. II, § F; Sargent et al. (15.5), § 7-5; Allen (15.7), Chap. 2; see also the article by Feynman et al. (1.33).

Complement D_{IV}

System of two spin 1/2 particles

1 Quantum mechanical description 441

1-a State space 441

1-b Complete sets of commuting observables 442

1-c The most general state 443

2 Prediction of the measurement results 444

2-a Measurements bearing simultaneously on the two spins . . . 444

2-b Measurements bearing on one spin alone 445

In this complement, we intend to use the formalism introduced in § A-2 of Chapter IV to describe a system of two spin 1/2 particles. This case is hardly more complicated than that of a single spin 1/2 particle. Its interest, as far as the postulates are concerned, lies in the fact that none of the various spin observables alone constitutes a C.S.C.O. (while this is the case for one spin alone). Thus, we shall be able to consider measurements bearing either on one observable with a degenerate spectrum or simultaneously on two observables. In addition, this study provides a very simple illustration of the concept of a tensor product, introduced in § F of Chapter II. We shall be concerned, as in Chapter IV, only with the internal degrees of freedom (spin states), and we shall moreover assume that the two particles which constitute the system are not identical (systems of identical particles will be studied in a general way in Chapters XIV and XV).

1. Quantum mechanical description

We saw in Chapter IV how to describe quantum mechanically the spin state of a spin 1/2 particle. Thus, all we need to do is apply the results of § F of Chapter II in order to know how to describe systems of two spin 1/2 particles.

1-a. State space

We shall use the indices 1 and 2 to distinguish between the two particles. When particle (1) is alone, its spin state is defined by a ket which belongs to a two-dimensional state space $\mathcal{E}_S(1)$. Similarly, the spin states of particle (2) alone form a two-dimensional space $\mathcal{E}_S(2)$. We shall designate by $\mathbf{S}_1$ and $\mathbf{S}_2$ the spin observables of particles (1) and (2) respectively. In $\mathcal{E}_S(1)$ [or $\mathcal{E}_S(2)$], we choose as a basis the eigenkets of S_{1z} (or S_{2z}), which we shall denote by $|1:+\rangle$ and $|1:-\rangle$ (or $|2:+\rangle$ and $|2:-\rangle$). The most general ket of $\mathcal{E}_S(1)$ can be written:

$$|\varphi(1)\rangle = \alpha_1|1:+\rangle + \beta_1|1:-\rangle \tag{1}$$

and that of $\mathcal{E}_S(2)$:

$$|\chi(2)\rangle = \alpha_2|2:+\rangle + \beta_2|2:-\rangle \tag{2}$$

(α_1, β_1, α_2, β_2 are arbitrary complex numbers).

When we join the two particles to make a single system, the state space $\mathcal{E}_S$ of such a system is the tensor product of the two preceding spaces:

$$\mathcal{E}_S = \mathcal{E}_S(1) \otimes \mathcal{E}_S(2) \tag{3}$$

In the first place, this means that a basis of $\mathcal{E}_S$ can be obtained by multiplying tensorially the two bases defined above for $\mathcal{E}_S(1)$ and $\mathcal{E}_S(2)$. We shall use the following notation:

$$\begin{aligned}
|++\rangle &= |1:+\rangle|2:+\rangle \\
|+-\rangle &= |1:+\rangle|2:-\rangle \\
|-+\rangle &= |1:-\rangle|2:+\rangle \\
|--\rangle &= |1:-\rangle|2:-\rangle
\end{aligned} \tag{4}$$

In the state $|+-\rangle$, for example, the component along Oz of the spin of particle (1) is $+\hbar/2$, with absolute certainty; that of the spin of particle (2) is $-\hbar/2$, with absolute certainty. We shall agree here to denote by $\langle +-|$ the conjugate bra of the ket $|+-\rangle$; the order of the symbols is therefore the same in the ket and in the bra: the first symbol is always associated with particle (1) and the second, with particle (2).

The space $\mathcal{E}_S$ is therefore four-dimensional. Since the $\{|1:\pm\rangle\}$ and $\{|2:\pm\rangle\}$ bases are orthonormal in $\mathcal{E}_S(1)$ and $\mathcal{E}_S(2)$ respectively, the basis (4) is orthonormal in $\mathcal{E}_S$:

$$\langle\varepsilon_1\varepsilon_2|\varepsilon_1'\varepsilon_2'\rangle = \delta_{\varepsilon_1\varepsilon_1'}\ \delta_{\varepsilon_2\varepsilon_2'} \tag{5}$$

($\varepsilon_1, \varepsilon_2, \varepsilon_1', \varepsilon_2'$ are to be replaced by $+$ or $-$ depending on the case; $\delta_{\varepsilon\varepsilon'}$ is equal to 1 if ε and ε' are identical and 0 if they are different). The system of vectors (4) also satisfies a closure relation in $\mathcal{E}_S$:

$$\begin{aligned}
\sum_{\varepsilon_1\varepsilon_2} |\varepsilon_1\varepsilon_2\rangle\langle\varepsilon_1\varepsilon_2| = &|++\rangle\langle++| + |+-\rangle\langle+-| + \\
&|-+\rangle\langle-+| + |--\rangle\langle--| = \mathbb{1}
\end{aligned} \tag{6}$$

1-b. Complete sets of commuting observables

We extend into $\mathcal{E}_S$ the observables $\mathbf{S}_1$ and $\mathbf{S}_2$ which were originally defined in $\mathcal{E}_S(1)$ and $\mathcal{E}_S(2)$ (as in Chapter II, we shall continue to denote these extensions by $\mathbf{S}_1$ and $\mathbf{S}_2$). Their action on the kets of the basis (4) is simple: the components of $\mathbf{S}_1$, for example, act only on the part of the ket related to particle (1). In particular, the vectors of the basis (4) are simultaneous eigenvectors of S_{1z} and S_{2z}:

$$\begin{aligned}
S_{1z}|\varepsilon_1\varepsilon_2\rangle &= \frac{\hbar}{2}\varepsilon_1|\varepsilon_1\varepsilon_2\rangle \\
S_{2z}|\varepsilon_1\varepsilon_2\rangle &= \frac{\hbar}{2}\varepsilon_2|\varepsilon_1\varepsilon_2\rangle
\end{aligned} \tag{7}$$

For the other components of $\mathbf{S}_1$ and $\mathbf{S}_2$, we apply the formulas given in § A-2 of Chapter IV. For example, we know from relation (A-16) of Chapter IV how S_{1x} acts on the kets $|1:\pm\rangle$:

$$\begin{aligned} S_{1x}|1:+\rangle &= \frac{\hbar}{2}|1:-\rangle \\ S_{1x}|1:-\rangle &= \frac{\hbar}{2}|1:+\rangle \end{aligned} \tag{8}$$

From this we deduce the action of S_{1x} on the kets (4):

$$\begin{aligned} S_{1x}|++\rangle &= \frac{\hbar}{2}|-+\rangle \\ S_{1x}|+-\rangle &= \frac{\hbar}{2}|--\rangle \\ S_{1x}|-+\rangle &= \frac{\hbar}{2}|++\rangle \\ S_{1x}|--\rangle &= \frac{\hbar}{2}|+-\rangle \end{aligned} \tag{9}$$

It is then easy to verify that, although the three components of $\mathbf{S}_1$ (or of $\mathbf{S}_2$) do not commute with each other, *any component of* $\mathbf{S}_1$ *commutes with any component of* $\mathbf{S}_2$.

In $\mathcal{E}_S(1)$, the observable S_{1z} alone constituted a C.S.C.O., and the same was true of S_{2z} in $\mathcal{E}_S(2)$. In $\mathcal{E}_S$, the eigenvalues of S_{1z} and S_{2z} remain $\pm\hbar/2$, but each of them is two-fold degenerate. To the eigenvalue $+\hbar/2$ of S_{1z}, for example, correspond two orthogonal vectors, $|++\rangle$ and $|+-\rangle$ [formulas (7)] and all their linear combinations. Therefore, in $\mathcal{E}_S$, neither S_{1z} nor S_{2z} (taken separately) constitutes a C.S.C.O. On the other hand, the set $\{S_{1z}, S_{2z}\}$ is a C.S.C.O. in $\mathcal{E}_S$, as can be seen from formulas (7).

This is obviously not the only C.S.C.O. that can be constructed. For example, another one is $\{S_{1z}, S_{2x}\}$. These two observables commute, as we noted above, and each of them constitutes a C.S.C.O. in the space in which it was initially defined. The eigenvectors which are common to S_{1z} and S_{2x} are obtained by taking the tensor product of their respective eigenvectors in $\mathcal{E}_S(1)$ and $\mathcal{E}_S(2)$. Using relation (A-20) of Chapter IV, we find:

$$\begin{aligned} |1:+\rangle|2:+\rangle_x &= \frac{1}{\sqrt{2}}[|++\rangle + |+-\rangle] \\ |1:+\rangle|2:-\rangle_x &= \frac{1}{\sqrt{2}}[|++\rangle - |+-\rangle] \\ |1:-\rangle|2:+\rangle_x &= \frac{1}{\sqrt{2}}[|-+\rangle + |--\rangle] \\ |1:-\rangle|2:-\rangle_x &= \frac{1}{\sqrt{2}}[|-+\rangle - |--\rangle] \end{aligned} \tag{10}$$

1-c. The most general state

The vectors (4) were obtained by multiplying tensorially a ket of $\mathcal{E}_S(1)$ and a ket of $\mathcal{E}_S(2)$. More generally, using an arbitrary ket of $\mathcal{E}_S(1)$ [such as (1)] and an arbitrary

ket of $\mathcal{E}_S(2)$ [such as (2)], we can construct a ket of $\mathcal{E}_S$:

$$|\varphi(1)\rangle|\chi(2)\rangle = \alpha_1\alpha_2|++\rangle + \alpha_1\beta_2|+-\rangle + \alpha_2\beta_1|-+\rangle + \beta_1\beta_2|--\rangle \tag{11}$$

The components of such a ket in the basis (4) are the products of the components of $|\varphi(1)\rangle$ and $|\chi(2)\rangle$ in the bases of $\mathcal{E}_S(1)$ and $\mathcal{E}_S(2)$ which were used to construct (4).

But *all the kets of $\mathcal{E}_S$ are not tensor products*. The most general ket of $\mathcal{E}_S$ is an arbitrary linear combination of the basis vectors:

$$|\psi\rangle = \alpha|++\rangle + \beta|+-\rangle + \gamma|-+\rangle + \delta|--\rangle \tag{12}$$

If we want to normalize $|\psi\rangle$, we must choose:

$$|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1 \tag{13}$$

Given $|\psi\rangle$, it is not in general possible to find two kets $\varphi(1)$ and $\chi(2)$ of which it is the tensor product. For (12) to be of the form (11), we must have, in particular:

$$\frac{\alpha}{\beta} = \frac{\gamma}{\delta} \tag{14}$$

and this condition is not necessarily fulfilled.

2. Prediction of the measurement results

We are now going to envisage a certain number of measurements that can be performed on a system of two spin 1/2 particles and we shall calculate the predictions provided by the postulates for each of them. We shall assume each time that the state of the system immediately before the measurement is described by the normalized ket (12).

2-a. Measurements bearing simultaneously on the two spins

Since any component of $\mathbf{S}_1$ commutes with any component of $\mathbf{S}_2$, we can envisage measuring them simultaneously (Chap. III, § C-6-a). To calculate the predictions related to such measurements, all we need to do is use the eigenvectors common to the two observables.

α. First example

Imagine we are simultaneously measuring S_{1z} and S_{2z}. What are the probabilities of the various results that can be obtained?

Since the set S_{1z}, S_{2z} is a C.S.C.O., there exists only one state associated with each measurement result. If the system is in the state (12) before the measurement, we can therefore find:

$$\begin{array}{llll}
+\frac{\hbar}{2} \text{ for } S_{1z} \text{ and } +\frac{\hbar}{2} \text{ for } S_{2z} & \text{with the probability} & |\langle++|\psi\rangle|^2 = |\alpha|^2 \\
+\frac{\hbar}{2} \quad '' \quad -\frac{\hbar}{2} \quad '' & '' & |\langle+-|\psi\rangle|^2 = |\beta|^2 \\
-\frac{\hbar}{2} \quad '' \quad +\frac{\hbar}{2} \quad '' & '' & |\langle-+|\psi\rangle|^2 = |\gamma|^2 \\
-\frac{\hbar}{2} \quad '' \quad -\frac{\hbar}{2} \quad '' & '' & |\langle--|\psi\rangle|^2 = |\delta|^2
\end{array} \tag{15}$$

β. Second example

We now measure S_{1y} and S_{2z}. What is the probability of obtaining $+\hbar/2$ for each of the two observables?

Here again, $\{S_{1y}, S_{2z}\}$ constitutes a C.S.C.O. The eigenvector common to S_{1y} and S_{2z} that corresponds to the eigenvalues $+\hbar/2$ and $+\hbar/2$ is the tensor product of the vector $|1:+\rangle_y$ and the vector $|2:+\rangle$:

$$|1:+\rangle_y|2:+\rangle = \frac{1}{\sqrt{2}}[|++\rangle + i|-+\rangle] \tag{16}$$

Applying the fourth postulate of Chapter III, we find that the probability we are looking for is:

$$\begin{aligned}\mathscr{P} &= \left|\frac{1}{\sqrt{2}}\left[\langle++| - i\langle-+|\right]|\psi\rangle\right|^2 \\ &= \frac{1}{2}|\alpha - i\gamma|^2\end{aligned} \tag{17}$$

The result therefore appears in the form of a "square of a sum"[1].

After the measurement, if we have actually found $+\hbar/2$ for S_{1y} and $+\hbar/2$ for S_{2z}, the system is in the state (16).

2-b. Measurements bearing on one spin alone

It is obviously possible to measure only one component of one of the two spins. In this case, since none of these components constitutes by itself a C.S.C.O., there exist several eigenvectors corresponding to the same measurement result, and the corresponding probability will be a "sum of squares".

α. First example

We measure only S_{1z}. What results can be found, and with what probabilities?

The possible results are the eigenvalues $\pm\hbar/2$ of S_{1z}. Each of them is doubly degenerate. In the associated eigensubspace, we choose an orthonormal basis: we can, for example, take $\{|++\rangle, |+-\rangle\}$ for $+\hbar/2$ and $\{|-+\rangle, |--\rangle\}$ for $-\hbar/2$. We then obtain:

$$\begin{aligned}\mathscr{P}\left(+\frac{\hbar}{2}\right) &= |\langle++|\psi\rangle|^2 + |\langle+-|\psi\rangle|^2 \\ &= |\alpha|^2 + |\beta|^2 \\ \mathscr{P}\left(-\frac{\hbar}{2}\right) &= |\langle-+|\psi\rangle|^2 + |\langle--|\psi\rangle|^2 \\ &= |\gamma|^2 + |\delta|^2\end{aligned} \tag{18}$$

[1] It must be remembered that the sign of i changes when we go from (16) to the conjugate bra. If this were to be forgotten, the result obtained would be incorrect ($|\alpha + iy|^2 \neq |\alpha - i\gamma|^2$ since α/γ is not in general real).

Comment:

Since we are not performing any measurement on the spin (2), the choice of the basis in $\mathcal{E}_S(2)$ is arbitrary. We can, for example, choose as a basis of the eigensubspace of S_{1z} associated with the eigenvalue $+\hbar/2$ the vectors:

$$|1:+\rangle|2:\pm\rangle_x = \frac{1}{\sqrt{2}}[|++\rangle \pm |+-\rangle] \tag{19}$$

which again yields:

$$\begin{aligned}\mathscr{P}\left(+\frac{\hbar}{2}\right) &= \frac{1}{2}|\alpha+\beta|^2 + \frac{1}{2}|\alpha-\beta|^2 \\ &= |\alpha|^2 + |\beta|^2\end{aligned} \tag{20}$$

The general proof of the fact that the probability obtained is independent (in the case of a degenerate eigenvalue) of the choice of the basis in the corresponding eigensubspace was given in § B-3-b-α of Chapter III.

β. Second example

We now choose to measure S_{2x}. What is the probability of obtaining $-\hbar/2$? The eigensubspace associated with the eigenvalue $-\hbar/2$ of S_{2x} is two-dimensional. We can choose as a basis in that subspace:

$$\begin{aligned}|1:+\rangle|2:-\rangle_x &= \frac{1}{\sqrt{2}}[|++\rangle - |+-\rangle] \\ |1:-\rangle|2:-\rangle_x &= \frac{1}{\sqrt{2}}[|-+\rangle - |--\rangle]\end{aligned} \tag{21}$$

We then find:

$$\begin{aligned}\mathscr{P} &= \left|\frac{1}{\sqrt{2}}\left[\langle++|-\langle+-|\right]|\psi\rangle\right|^2 + \left|\frac{1}{\sqrt{2}}\left[\langle-+|-\langle--|\right]|\psi\rangle\right|^2 \\ &= \frac{1}{2}|\alpha-\beta|^2 + \frac{1}{2}|\gamma-\delta|^2\end{aligned} \tag{22}$$

In this result, each of the terms of the "sum of squares" is itself the "square of a sum".

If the measurement actually yields $-\hbar/2$, the state $|\psi'\rangle$ of the system immediately after this measurement is the (normalized) projection of $|\psi\rangle$ onto the corresponding eigensubspace. We have just calculated the components of $|\psi\rangle$ on the basis vectors (21) of this subspace: they are equal, respectively, to $\frac{1}{\sqrt{2}}(\alpha-\beta)$ and $\frac{1}{\sqrt{2}}(\gamma-\delta)$. Consequently:

$$\begin{aligned}|\psi'\rangle = &\frac{1}{\sqrt{\frac{1}{2}|\alpha-\beta|^2 + \frac{1}{2}|\gamma-\delta|^2}} \\ &\times\left[\frac{1}{2}(\alpha-\beta)(|++\rangle - |+-\rangle) + \frac{1}{2}(\gamma-\delta)(|-+\rangle - |--\rangle)\right]\end{aligned} \tag{23}$$

Comment:

We have considered, in this complement, only the components of $\mathbf{S}_1$ and $\mathbf{S}_2$ on the coordinate axes. It is obviously possible to measure their components $\mathbf{S}_1 \cdot \mathbf{u}$ and $\mathbf{S}_2 \cdot \mathbf{v}$ on arbitrary unit vectors $\mathbf{u}$ and $\mathbf{v}$. The reasoning is the same as above.

Complement E_{IV}

Spin $1/2$ density matrix

1	**Introduction**	**449**
2	**Density matrix of a perfectly polarized spin (pure case)**	**449**
3	**Example of a statistical mixture: unpolarized spin**	**450**
4	**Spin 1/2 at thermodynamic equilibrium in a static field**	**452**
5	**Expansion of the density matrix in terms of the Pauli matrices**	**453**

1. Introduction

The aim of this complement is to illustrate the general considerations developed in Complement E_{III}, using a very simple physical system, that of a spin 1/2. We are going to study the density matrices which describe a spin 1/2 in a certain number of cases: perfectly polarized spin (pure case), unpolarized or partially polarized spin (statistical mixture). This will lead us to verify and interpret the general properties stated in Complement E_{III}. In addition, we shall see that the expansion of the density matrix in terms of the Pauli matrices can be expressed very simply as a function of the mean values of the various spin components.

2. Density matrix of a perfectly polarized spin (pure case)

Consider a spin 1/2, coming out of an "atomic polarizer" of the type described in § B of Chapter IV. We assume that it is in the eigenstate $|+\rangle_{\mathbf{u}}$ (eigenvalue $+\hbar/2$) of the $\mathbf{S}\cdot\mathbf{u}$ component of the spin (recall that the polar angles of the unit vector $\mathbf{u}$ are θ and φ). The spin state is then perfectly well-known and is written [*cf.* formula (A-22a) of Chapter IV]:

$$|\psi\rangle = \cos\frac{\theta}{2}\,\mathrm{e}^{-i\varphi/2}|+\rangle + \sin\frac{\theta}{2}\,\mathrm{e}^{i\varphi/2}|-\rangle \tag{1}$$

We saw in Complement E_{III} that, by definition, such a situation corresponds to a pure case. We shall say that the beam which leaves the "polarizer" is perfectly polarized. Recall also that, for each spin, the mean value $\langle\mathbf{S}\rangle$ is equal to $\frac{\hbar}{2}\mathbf{u}$ [Chap. IV, relations (B-14)].

It is simple to write, in the $\{|+\rangle, |-\rangle\}$ basis, the density matrix $\rho(\theta,\varphi)$ corresponding to the state (1). We write the matrix of the projector onto this state:

$$\rho(\theta,\varphi) = \begin{pmatrix} \cos^2\dfrac{\theta}{2} & \sin\dfrac{\theta}{2}\cos\dfrac{\theta}{2}\,\mathrm{e}^{-i\varphi} \\ \sin\dfrac{\theta}{2}\cos\dfrac{\theta}{2}\,\mathrm{e}^{i\varphi} & \sin^2\dfrac{\theta}{2} \end{pmatrix} \tag{2}$$

This matrix is generally non-diagonal. The "populations" ρ_{++} and ρ_{--} have a very simple physical significance. Their difference is equal to $\cos\theta = 2\langle S_z\rangle/\hbar$ [*cf.* equations (B-14) of Chapter IV], and their sum is, of course, equal to 1. The populations are therefore related to the longitudinal polarization $\langle S_z\rangle$. Similarly, the modulus of the "coherences" ρ_{+-} and ρ_{-+} is $|\rho_{+-}| = |\rho_{-+}| = \frac{1}{2}\sin\theta = \frac{1}{\hbar}|\langle \mathbf{S}_\perp\rangle|$ (where $\langle \mathbf{S}_\perp\rangle$ is the projection of $\langle \mathbf{S}\rangle$ onto the xOy plane). The argument of ρ_{-+} is φ, which is the angle between $\langle \mathbf{S}_\perp\rangle$ and Ox: the coherences are therefore related to the transverse polarization $\langle \mathbf{S}_\perp\rangle$.

It can also be verified that:

$$[\rho(\theta,\varphi)]^2 = \rho(\theta,\varphi) \tag{3}$$

a relation characteristic of a pure state.

3. Example of a statistical mixture: unpolarized spin

Now let us consider the spin of a silver atom leaving a furnace, such as the one in Figure 1 of Chapter IV, and which has not passed through an "atomic polarizer" (the spin has not been prepared in a particular state). The only information we then possess about this spin is the following: it can point in any direction of space, and all directions are equally probable. With the notation of Complement E_{III}, such a situation corresponds to a statistical mixture of the states $|+\rangle_u$ with equal weights. Formula (28) of Complement E_{III} defines the density matrix ρ that corresponds to this case. Nevertheless, the discrete sum Σ_k must here be replaced by an integral over all possible directions:

$$\rho = \frac{1}{4\pi}\int \mathrm{d}\Omega\ \rho(\theta,\varphi) = \frac{1}{4\pi}\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi}\sin\theta\ \mathrm{d}\theta\ \rho(\theta,\varphi) \tag{4}$$

(the factor $1/4\pi$ insures the normalization of the probabilities associated with the various directions). The integrals which give the matrix elements of ρ are simple to calculate and lead to the following result:

$$\rho = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \tag{5}$$

It is easy to deduce from (5) that $\rho^2 = \rho/2$, which shows that, in the case of a statistical mixture of states, ρ^2 is different from ρ.

In addition, if we calculate from (5) the mean values of S_x, S_y, S_z, we obtain:

$$\langle S_i\rangle = \mathrm{Tr}\{\rho S_i\} = \frac{1}{2}\mathrm{Tr}\{S_i\} = 0 \qquad \text{with: } i = x, y, z \tag{6}$$

We again find the fact that the spin is unpolarized: since all the directions are equivalent, the mean value of the spin is zero.

Comments:

(*i*) It is clear from this example how the non-diagonal elements (coherences) of ρ can disappear from the summation over the various states of the statistical mixture. As we saw in § 2, the coherences ρ_{+-} and ρ_{-+} are related to the transverse polarization $\langle \mathbf{S}_\perp \rangle$ of the spin. Upon summing the vectors $\langle \mathbf{S}_\perp \rangle$ corresponding to all (equiprobable) directions of the xOy plane, we obviously find a null result.

(*ii*) The case of an unpolarized spin is also very instructive, since it helps us understand the impossibility of describing a statistical mixture by an "average state vector". Assume that we are trying to choose α and β so that the vector:

$$|\psi\rangle = \alpha|+\rangle + \beta|-\rangle \tag{7}$$

with:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{8}$$

represents an unpolarized spin, for which $\langle S_x \rangle$, $\langle S_y \rangle$ and $\langle S_z \rangle$ are zero. A simple calculation gives:

$$\begin{aligned} \langle S_x \rangle &= \frac{\hbar}{2}(\alpha^*\beta + \alpha\beta^*) \\ \langle S_y \rangle &= \frac{\hbar}{2i}(\alpha^*\beta - \alpha\beta^*) \\ \langle S_z \rangle &= \frac{\hbar}{2}(\alpha^*\alpha - \beta^*\beta) \end{aligned} \tag{9}$$

If we want to make $\langle S_x \rangle$ zero, we must choose α and β so as to make $\alpha^*\beta$ a pure imaginary; similarly, $\alpha^*\beta$ must be real for $\langle S_y \rangle$ to be zero. We must therefore have $\alpha^*\beta = 0$, that is:

either $\alpha = 0$, which implies $|\beta| = 1$ and $\langle S_z \rangle = -\hbar/2$
or $\beta = 0$, which implies $|\alpha| = 1$ and $\langle S_z \rangle = \hbar/2$

Therefore, $\langle S_z \rangle$, $\langle S_x \rangle$ and $\langle S_y \rangle$ cannot all be zero at the same time; consequently, the state (7) cannot represent an unpolarized spin.

Furthermore, the discussion of § B-1-c of Chapter IV shows that for any α and β that satisfy (8), one can always associate with them two angles θ and φ fixing a direction $\mathbf{u}$ such that $|\psi\rangle$ is an eigenvector of $\mathbf{S}.\mathbf{u}$ with the eigenvalue $+\hbar/2$. Thus we see directly that a state such as (7) always describes a spin which is perfectly polarized in a certain direction of space.

(*iii*) The density matrix (5) represents a statistical mixture of the various states $|+\rangle_\mathbf{u}$, all the directions $\mathbf{u}$ being equiprobable (this is how we obtained it). We could, however, imagine other statistical mixtures which would lead to the same density matrix: for example, a statistical mixture of equal proportions of the states $|+\rangle$ and $|-\rangle$, or a statistical mixture of equal proportions of three states $|+\rangle_\mathbf{u}$, such that the tips of the three corresponding vectors $\mathbf{u}$ are the vertices of an equilateral triangle centered at O. Thus we see that the same

density matrix can be obtained in several different ways. In fact, since all the physical predictions depend only on the density matrix, it is impossible to distinguish physically between the various types of statistical mixtures that lead to the same density matrix. They must be considered to be different expressions of the same incomplete information we possess about the system.

4. Spin 1/2 at thermodynamic equilibrium in a static field

Consider a spin 1/2 placed in a static field $\mathbf{B}_0$ parallel to Oz. We saw in § B-3-a of Chapter IV that the stationary states of this spin are the states $|+\rangle$ and $|-\rangle$, of energies $+\hbar\omega_0/2$ and $-\hbar\omega_0/2$ (with $\omega_0 = -\gamma B_0$, where γ is the gyromagnetic ratio of the spin). If we know only that the system is in thermodynamic equilibrium at the temperature T, we can assert that it has a probability $Z^{-1}\mathrm{e}^{-\hbar\omega_0/2kT}$ of being in the state $|+\rangle$ and Z^{-1} $\mathrm{e}^{+\hbar\omega_0/2kT}$ of being in the state $|-\rangle$, where $Z = \mathrm{e}^{-\hbar\omega_0/2kT} + \mathrm{e}^{+\hbar\omega_0/2kT}$ is a normalization factor (Z is called the "partition function"). We have here another example of a statistical mixture, described by the density matrix:

$$\rho = Z^{-1} \begin{pmatrix} \mathrm{e}^{-\hbar\omega_0/2kT} & 0 \\ 0 & \mathrm{e}^{+\hbar\omega_0/2kT} \end{pmatrix} \tag{10}$$

Once more, it is easy to verify that $\rho^2 \neq \rho$. The non-diagonal elements are zero since all directions perpendicular to $\mathbf{B}_0$ (that is, to Oz) and fixed by the angle φ are equivalent.

From (10), it is easy to calculate:

$$\begin{aligned} \langle S_x \rangle &= \mathrm{Tr}(\rho S_x) = 0 \\ \langle S_y \rangle &= \mathrm{Tr}(\rho S_y) = 0 \\ \langle S_z \rangle &= \mathrm{Tr}(\rho S_z) = -\frac{\hbar}{2} \tanh\left(\frac{\hbar\omega_0}{2kT}\right) \end{aligned} \tag{11}$$

We see that the spin acquires a polarization parallel to the field in which it is placed. The larger ω_0 (that is, B_0) and the lower the temperature T, the greater the polarization. Since $|\tanh x| < 1$, this polarization is less than the value $\hbar/2$ corresponding to a spin that is perfectly polarized along Oz. The density matrix (10) can therefore be said to describe a spin which is "partially polarized" along Oz.

Comment:

The magnetization $\langle M_z \rangle$ is equal to $\gamma\langle S_z \rangle$. It is possible to calculate from (11) the paramagnetic susceptibility χ of the spin, defined by:

$$\langle M_z \rangle = \gamma \langle S_z \rangle = \chi B_0 \tag{12}$$

We find (Brillouin's formula):

$$\chi = \frac{\hbar\gamma}{2B_0} \tanh\left(\frac{\hbar\gamma B_0}{2kT}\right) \tag{13}$$

5. Expansion of the density matrix in terms of the Pauli matrices

We saw in Complement A_{IV} that the unit matrix I and the Pauli matrices σ_x, σ_y and σ_z form a convenient basis for expanding a 2×2 matrix. We therefore set, for the density matrix ρ of a spin 1/2:

$$\rho = a_0\, I + \mathbf{a} \cdot \boldsymbol{\sigma} \tag{14}$$

where the coefficients a_i are given by [*cf.* Complement A_{IV}, relations (23)]:

$$\begin{aligned} a_0 &= \frac{1}{2}\ \mathrm{Tr}\{\rho\} \\ a_x &= \frac{1}{2}\ \mathrm{Tr}\{\rho\sigma_x\} = \frac{1}{\hbar}\ \mathrm{Tr}\{\rho S_x\} \\ a_y &= \frac{1}{2}\ \mathrm{Tr}\{\rho\sigma_y\} = \frac{1}{\hbar}\ \mathrm{Tr}\{\rho S_y\} \\ a_z &= \frac{1}{2}\ \mathrm{Tr}\{\rho\sigma_z\} = \frac{1}{\hbar}\ \mathrm{Tr}\{\rho S_z\} \end{aligned} \tag{15}$$

Thus we have:

$$\begin{aligned} a_0 &= \frac{1}{2} \\ \mathbf{a} &= \frac{1}{\hbar}\langle \mathbf{S} \rangle \end{aligned} \tag{16}$$

and ρ can be written:

$$\rho = \frac{1}{2}\ I + \frac{1}{\hbar}\langle \mathbf{S} \rangle \cdot \boldsymbol{\sigma} \tag{17}$$

Therefore, the density matrix ρ of a spin 1/2 can be expressed very simply in terms of the mean value $\langle \mathbf{S} \rangle$ of the spin.

Comment:

Let us square expression (17). We obtain, using identity (17) of Complement A_{IV}:

$$\rho^2 = \frac{1}{4} I + \frac{1}{\hbar^2}\langle \mathbf{S} \rangle^2\ I + \frac{1}{\hbar}\langle \mathbf{S} \rangle \cdot \boldsymbol{\sigma} \tag{18}$$

The condition $\rho^2 = \rho$, characteristic of the pure case, is therefore equivalent, for a spin 1/2, to the condition:

$$\langle \mathbf{S} \rangle^2 = \frac{\hbar^2}{4} \tag{19}$$

This condition is obviously not satisfied for an unpolarized spin ($\langle \mathbf{S} \rangle$ is then zero) or for a spin in thermodynamic equilibrium (we saw in § 4 that in this case $|\langle \mathbf{S} \rangle| < \hbar/2$). On the other hand, it can be verified, using formulas (B-14) of Chapter IV, that, for a spin in the state $|\psi\rangle$ given in (1), $\langle \mathbf{S} \rangle^2$ is indeed equal to $\hbar^2/4$.

References and suggestions for further reading:

Abragam (14.1), Chap. II, § C.

Complement F_{IV}

Spin 1/2 particle in a static and a rotating magnetic fields: magnetic resonance

1 Classical treatment; rotating reference frame 455
1-a Motion in a static field; Larmor precession 455
1-b Influence of a rotating field; resonance 456
2 Quantum mechanical treatment 458
2-a The Schrödinger equation 458
2-b Changing to the rotating frame 459
2-c Transition probability; Rabi's formula 460
2-d Case where the two levels are unstable 461
3 Relation between the classical treatment and the quantum mechanical treatment: evolution of $\langle \mathbf{M} \rangle$ 463
4 Bloch equations 463
4-a A concrete example 464
4-b Solution in the case of a rotating field 466

In Chapter IV, we used quantum mechanics to study the evolution of a spin 1/2 in a static magnetic field. In this complement, we shall consider the case of a spin 1/2 simultaneously subjected to several magnetic fields, some of which can be time dependent, as is the case in magnetic resonance experiments. Before attacking this problem quantum mechanically, we shall briefly review several results obtained using classical mechanics.

1. Classical treatment; rotating reference frame

1-a. Motion in a static field; Larmor precession

Consider a system with angular momentum $\mathbf{j}$ that possesses a magnetic moment $\mathbf{m} = \gamma \mathbf{j}$ collinear with $\mathbf{j}$ (the constant γ is the gyromagnetic ratio). The system is placed in a static magnetic field $\mathbf{B}_0$, which exerts a torque $\mathbf{m} \times \mathbf{B}_0$ on the system. The classical equation of motion of $\mathbf{j}$ is:

$$\frac{\mathrm{d}\mathbf{j}}{\mathrm{d}t} = \mathbf{m} \times \mathbf{B}_0 \tag{1}$$

or:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{m}(t) = \gamma\ \mathbf{m}(t) \times \mathbf{B}_0 \tag{2}$$

Performing a scalar multiplication of both sides of this equation by either $\mathbf{m}(t)$ or $\mathbf{B}_0$, we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{m}(t)]^2 = 0 \tag{3}$$

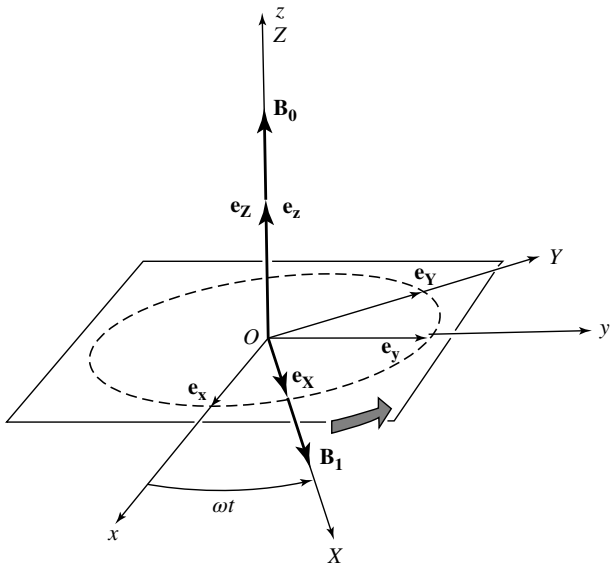


Figure 1: Oxyz is a fixed coordinate system. The static magnetic field $\mathbf{B}_0$ *is directed along the Oz axis. The OXYZ system [the OX axis is in the direction of the field* $\mathbf{B}_1(t)$*] rotates about Oz with the angular velocity* ω*.*

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{m}(t)\cdot\mathbf{B}_0]=0 \tag{4}$$

$\mathbf{m}(t)$ therefore evolves with a constant modulus, maintaining a constant angle with $\mathbf{B}_0$. If we project equation (2) onto the plane perpendicular to $\mathbf{B}_0$, we see that $\mathbf{m}(t)$ rotates about $\mathbf{B}_0$ (Larmor precession) with an angular velocity of $\omega_0=-\gamma B_0$ (the rotation is clockwise if γ is positive).

1-b. Influence of a rotating field; resonance

Now assume that we add to the static field $\mathbf{B}_0$ a field $\mathbf{B}_1(t)$, perpendicular to $\mathbf{B}_0$, and which is of constant modulus and rotates about $\mathbf{B}_0$ with an angular velocity ω (*cf.* Fig. 1). We set:

$$\begin{aligned}\omega_0&=-\gamma B_0\\ \omega_1&=-\gamma B_1\end{aligned} \tag{5}$$

We shall designate by $Oxyz$ (unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$) a fixed coordinate system, whose Oz axis is the direction of the field $\mathbf{B}_0$, and by $OXYZ$ (unit vectors $\mathbf{e}_X$, $\mathbf{e}_Y$, $\mathbf{e}_Z$), the axes obtained from $Oxyz$ by rotation through an angle ωt about Oz [OX is the direction of the rotating field $\mathbf{B}_1(t)$]. The equation of motion of $\mathbf{m}(t)$ in the presence of the total field $\mathbf{B}(t)=\mathbf{B}_0+\mathbf{B}_1(t)$ then becomes:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{m}(t)=\gamma\ \mathbf{m}(t)\times[\mathbf{B}_0+\mathbf{B}_1(t)] \tag{6}$$

To solve this equation, it is convenient to place ourselves, not in the laboratory reference frame $Oxyz$, but in the rotating reference frame $OXYZ$, with respect to which the relative velocity of the vector $\mathbf{m}(t)$ is:

$$\left(\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t}\right)_{\mathrm{rel}} = \frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t} - \omega\ \mathbf{e}_z \times \mathbf{m}(t) \tag{7}$$

Let us set:

$$\Delta\omega = \omega - \omega_0 \tag{8}$$

Substituting (6) into (7), we obtain:

$$\left(\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t}\right)_{\mathrm{rel}} = \mathbf{m}(t) \times [\Delta\omega\ \mathbf{e}_Z - \omega_1\ \mathbf{e}_X] \tag{9}$$

This equation is much simpler to solve than equation (6), since the coefficients of the right-hand side are now time-independent. Moreover, its form is analogous to that of (2): the relative motion of the vector $\mathbf{m}(t)$ is therefore a rotation about the "effective field" $\mathbf{B}_{\mathrm{eff}}$ (which is static with respect to the rotating reference frame), given by (*cf.* Fig. 2):

$$\mathbf{B}_{\mathrm{eff}} = \frac{1}{\gamma}[\Delta\omega\ \mathbf{e}_Z - \omega_1 \mathbf{e}_X] \tag{10}$$

To obtain the absolute motion of $\mathbf{m}(t)$, we must combine this precession about $\mathbf{B}_{\mathrm{eff}}$ with a rotation about Oz of angular velocity ω.

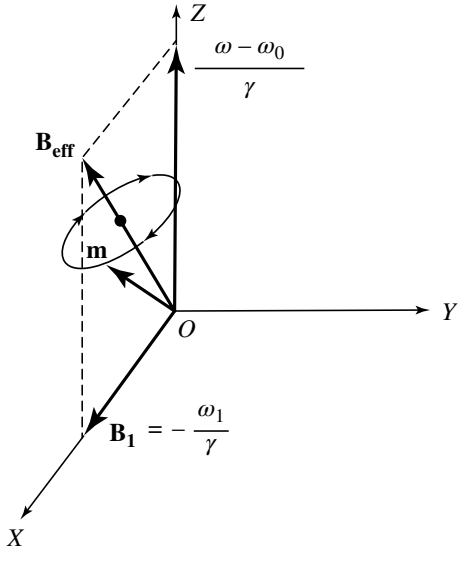


Figure 2: In the rotating reference frame $OXYZ$, the effective field $\mathbf{B}_{\mathrm{eff}}$ has a fixed direction, about which the magnetic moment $\mathbf{m}(t)$ rotates with a constant angular velocity (precession in the rotating reference frame).

These first results already enable us to understand the essence of the magnetic resonance phenomenon. Let us consider a magnetic moment which, at time $t = 0$, is

parallel to the field $\mathbf{B}_0$ (the case, for example, of a magnetic moment in thermodynamic equilibrium at very low temperatures: it is in the lowest energy state possible in the presence of the field $\mathbf{B}_0$). What happens when we apply a weak rotating field $\mathbf{B}_1(t)$? If the rotation frequency $\omega/2\pi$ of this field is very different from the natural frequency $\omega_0/2\pi$ (more precisely, if $\Delta\omega = \omega - \omega_0$ is much larger than ω_1), the effective field is directed practically along Oz. The precession of $\mathbf{m}(t)$ about $\mathbf{B}_{\text{eff}}$ then has a very small amplitude and hardly modifies the direction of the magnetic moment. On the other hand, if the resonance condition $\omega \simeq \omega_0$ is satisfied ($\Delta\omega \ll \omega_1$), the angle between the field $\mathbf{B}_{\text{eff}}$ and Oz is large. The precession of the magnetic moment then has a large amplitude and, at resonance ($\Delta\omega = 0$), the magnetic moment can even be completely flipped.

2. Quantum mechanical treatment

2-a. The Schrödinger equation

Let $|+\rangle$ and $|-\rangle$ be two eigenvectors of the projection S_z of the spin onto Oz, with respective eigenvalues $+\hbar/2$ and $-\hbar/2$. The state vector of the system can be written:

$$|\psi(t)\rangle = a_+(t)|+\rangle + a_-(t)|-\rangle \tag{11}$$

The Hamiltonian operator $H(t)$ of the system is[1]:

$$H(t) = -\mathbf{M}\cdot\mathbf{B}(t) = -\gamma\ \mathbf{S}\cdot[\mathbf{B}_0 + \mathbf{B}_1(t)] \tag{12}$$

that is, expanding the scalar product:

$$H(t) = \omega_0 S_z + \omega_1[\cos\omega t\ S_x + \sin\omega t\ S_y] \tag{13}$$

Using formulas (A-16) and (A-17) of Chapter IV, we obtain the matrix that represents H in the $|+\rangle, |-\rangle$ basis:

$$H = \frac{1}{2}\begin{pmatrix} \omega_0 & \omega_1\,\mathrm{e}^{-i\omega t} \\ \omega_1\,\mathrm{e}^{i\omega t} & -\omega_0 \end{pmatrix} \tag{14}$$

Using (11) and (14), we can write the Schrödinger equation in the form:

$$\begin{cases} i\dfrac{\mathrm{d}}{\mathrm{d}t}a_+(t) = \dfrac{\omega_0}{2}a_+(t) + \dfrac{\omega_1}{2}\mathrm{e}^{-i\omega t}a_-(t) \\ i\dfrac{\mathrm{d}}{\mathrm{d}t}a_-(t) = \dfrac{\omega_1}{2}\mathrm{e}^{i\omega t}a_+(t) - \dfrac{\omega_0}{2}a_-(t) \end{cases} \tag{15}$$

[1] In expression (12), $\mathbf{M}.\mathbf{B}(t)$ symbolizes the scalar product $M_xB_x(t) + M_yB_y(t) + M_zB_z(t)$, where M_x, M_y and M_z are operators (observables of the system under study), while $B_x(t), B_y(t)$ and $B_z(t)$ are numbers (since we consider the magnetic field to be a classical quantity whose value is imposed by an external device, independent of the system under study).

2-b. Changing to the rotating frame

Equations (15) constitute a linear homogeneous system with time-dependent coefficients. It is convenient to define new functions by setting:

$$b_+(t) = \mathrm{e}^{i\omega t/2} a_+(t)$$
$$b_-(t) = \mathrm{e}^{-i\omega t/2} a_-(t) \tag{16}$$

Substituting (16) into (15), we obtain a system which has constant coefficients:

$$\begin{cases} i\dfrac{\mathrm{d}}{\mathrm{d}t} b_+(t) = -\dfrac{\Delta\omega}{2} b_+(t) + \dfrac{\omega_1}{2} b_-(t) \\ i\dfrac{\mathrm{d}}{\mathrm{d}t} b_-(t) = \dfrac{\omega_1}{2} b_+(t) + \dfrac{\Delta\omega}{2} b_-(t) \end{cases} \tag{17}$$

This system can also be written:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\widetilde{\psi}(t)\rangle = \widetilde{H} |\widetilde{\psi}(t)\rangle \tag{18}$$

if we introduce the ket $|\widetilde{\psi}(t)\rangle$ and the operator $\widetilde{H}$ defined by:

$$|\widetilde{\psi}(t)\rangle = b_+(t)|+\rangle + b_-(t)|-\rangle \tag{19}$$

$$\widetilde{H} = \frac{\hbar}{2} \begin{pmatrix} -\Delta\omega & \omega_1 \\ \omega_1 & \Delta\omega \end{pmatrix} \tag{20}$$

Transformation (16) has led to equation (18), which is analogous to a Schrödinger equation in which the operator $\widetilde{H}$, given in (20), plays the role of a time-independent Hamiltonian. $\widetilde{H}$ describes the interaction of the spin with a *fixed* field, whose components are none other than those of the effective field introduced above in the $OXYZ$ frame [formula (10)]. We can therefore consider that the transformation (16) is the quantum mechanical equivalent of the change from the fixed $Oxyz$ frame to the rotating $OXYZ$ frame.

This result can be proved rigorously. According to (16), we can write:

$$|\widetilde{\psi}(t)\rangle = R(t)|\psi(t)\rangle \tag{21}$$

where $R(t)$ is the unitary operator defined by:

$$R(t) = \mathrm{e}^{i\omega t S_z/\hbar} \tag{22}$$

We shall see later (*cf.* Complement $\mathrm{B_{VI}}$) that $R(t)$ describes a rotation of the coordinate system through an angle ωt about Oz. (18) is therefore indeed the transformed Schrödinger equation in the rotating $OXYZ$ frame.

Equation (18) is very simple to solve. To determine $|\widetilde{\psi}(t)\rangle$, given $|\widetilde{\psi}(0)\rangle$, all we need to do is expand $|\widetilde{\psi}(0)\rangle$ on the eigenvectors of $\widetilde{H}$ (which can be calculated exactly) and then apply rule (D-54) of Chapter III (which is possible since $\widetilde{H}$ is not explicitly time-dependent). We then go from $|\widetilde{\psi}(t)\rangle$ to $|\psi(t)\rangle$ by using formulas (16).

2-c. Transition probability; Rabi's formula

Consider a spin which, at time $t = 0$, is in the state $|+\rangle$:

$$|\psi(0)\rangle = |+\rangle \tag{23}$$

According to (16), this corresponds to:

$$|\widetilde{\psi}(0)\rangle = |+\rangle \tag{24}$$

What is the probability $\mathscr{P}_{+-}(t)$ of finding this spin in the state $|-\rangle$ at time t? Since $a_-(t)$ and $b_-(t)$ have the same modulus, we can write:

$$\mathscr{P}_{+-}(t) = |\langle -|\psi(t)\rangle|^2 = |a_-(t)|^2 = |b_-(t)|^2 = |\langle -|\widetilde{\psi}\rangle|^2 \tag{25}$$

We must therefore calculate $|\langle -|\widetilde{\psi}(t)\rangle|^2$, where $|\widetilde{\psi}(t)\rangle$ is the solution of (18) that corresponds to the initial condition (24).

The problem we have just posed has already been solved, in § C-3-b of Chapter IV. To use the calculations of that section, all we need to do is apply the following correspondences:

$$\begin{aligned}
&|\varphi_1\rangle \longrightarrow |+\rangle \\
&|\varphi_2\rangle \longrightarrow |-\rangle \\
&E_1 \longrightarrow -\frac{\hbar}{2}\Delta\omega \\
&E_2 \longrightarrow \frac{\hbar}{2}\Delta\omega \\
&W_{12} \longrightarrow \frac{\hbar}{2}\omega_1
\end{aligned} \tag{26}$$

Rabi's formula [equation (C-32) of Chapter IV] then becomes:

$$\mathscr{P}_{+-}(t) = \frac{\omega_1^2}{\omega_1^2 + (\Delta\omega)^2} \sin^2\left[\sqrt{\omega_1^2 + (\Delta\omega)^2}\ \frac{t}{2}\right] \tag{27}$$

The probability $\mathscr{P}_{+-}(t)$ is zero at time $t = 0$, and then varies sinusoidally with time between the values 0 and $\frac{\omega_1^2}{\omega_1^2+(\Delta\omega)^2}$. Again, we have a resonance phenomenon. For $|\Delta\omega| \gg |\omega_1|$, $\mathscr{P}_{+-}(t)$ remains almost zero (*cf.* Fig. 3-a); near resonance, the oscillation amplitude of $\mathscr{P}_{+-}(t)$ becomes large and, when the condition $\Delta\omega = 0$ is exactly satisfied, we have $\mathscr{P}_{+-}(t) = 1$ at times $t = \frac{(2n+1)\pi}{\omega_1}$ (*cf.* Fig. 3-b).

Thus we again find the result which we have already obtained classically: at resonance, a very weak rotating field is able to reverse the direction of the spin. Note, moreover, that the angular frequency of the oscillation of $\mathscr{P}_{+-}(t)$ is $\sqrt{\omega_1^2 + (\Delta\omega)^2} = |\gamma \mathbf{B}_{\text{eff}}|$. This oscillation corresponds, in the rotating frame, to the projection onto OZ of the precession of the magnetic moment about the effective field, sometimes called "Rabi precession" [see also the calculation of $\mathscr{P}_{+-}(t)$ in Complement C_{IV}, § 3-c].

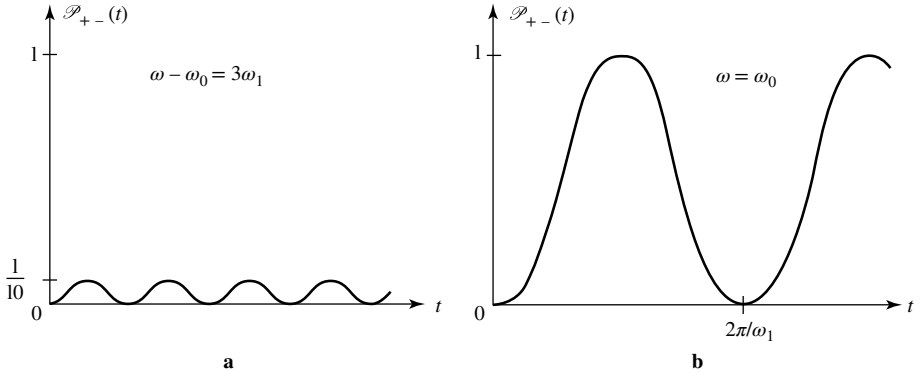


*Figure 3: Variation with respect to time of the transition probability between the states $|+\rangle$ and $|-\rangle$, under the effect of a rotating magnetic field $B_1(t)$. Outside resonance (fig. **a**), this probability remains small; at resonance (fig. **b**), however small the field B_1, there exist times when the transition probability is equal to 1.*

2-d. Case where the two levels are unstable

We are now going to assume that the two states $|\pm\rangle$ correspond to two sublevels of an excited atomic level (whose angular momentum is assumed equal to 1/2). n atoms are excited[2] per unit time, all being raised to the state $|+\rangle$. Each atom decays, by spontaneous emission of radiation, with a probability per unit time of $1/\tau$, which is the same for the two sublevels $|\pm\rangle$. We know that, under these conditions, an atom which was excited at time $-t$ has a probability $\mathrm{e}^{-t/\tau}$ of still being excited at time $t = 0$ (*cf.* Complement $\mathrm{K_{III}}$).

We assume that the experiment is performed in the steady state: in the presence of the fields $\mathbf{B}_0$ and $\mathbf{B}_1(t)$, the atoms are excited at a constant rate n into the state $|+\rangle$. After a time much longer than the lifetime τ, what is the number N of atoms which decay per unit time from the state $|-\rangle$? If an atom is excited at time $-t$, the probability of finding it in the state $|-\rangle$ at $t = 0$ is $\mathrm{e}^{-t/\tau}\mathscr{P}_{+-}(t)$, where $\mathscr{P}_{+-}(t)$ is given by relation (27). The total number of atoms in the state $|-\rangle$ is obtained by taking the sum of atoms excited at all previous times $-t$, that is, by calculating the integral:

$$\int_0^\infty \mathrm{e}^{-t/\tau}\mathscr{P}_{+-}(t)\; n\; \mathrm{d}t \tag{28}$$

This calculation presents no difficulties. Multiplying the number of atoms thus obtained by their probability $1/\tau$ of decay per unit time, we obtain:

$$N = \frac{n}{2}\frac{\omega_1^2}{(\Delta\omega)^2 + \omega_1^2 + (1/\tau)^2} \tag{29}$$

[2]In practice, this excitation can be produced, for example, by placing the atoms in a light beam. When the incident photons are polarized, conservation of angular momentum, in certain cases, requires that the atoms which absorb them can attain only the state $|+\rangle$ (and not the state $|-\rangle$). Similarly, by detecting the polarization of the photons re-emitted by the atoms, one can know whether the atoms fall back into the ground state from the state $|+\rangle$ or the state $|-\rangle$.

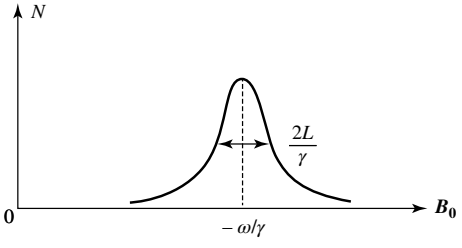


Figure 4: Resonance curve. To observe a resonance phenomenon, we perform an experiment in which n atoms are excited per unit time into the state $|+\rangle$. Under the effect of a field $\mathbf{B}_1(t)$, rotating at the frequency $\omega/2\pi$, the atoms undergo transitions towards the state $|-\rangle$. In the steady state, if we measure the number N of atoms which decay per unit time from the state $|-\rangle$, we obtain a resonant variation when we scan the static field B_0 about the value $-\omega/\gamma$.

The variation of N with respect to $\Delta\omega$ corresponds to a Lorentz curve whose half-width is:

$$L = \sqrt{\omega_1^2 + (1/\tau)^2} \tag{30}$$

In the experiment described above, let us measure, for various values of the magnetic field B_0 (that is, with ω assumed to be fixed, for various values of $\Delta\omega$), the number of atoms which decay from the level $|-\rangle$. According to (29), we obtain a resonance curve which has the shape shown in Figure 4.

It is very interesting to obtain such a curve experimentally, since one can use it to determine several parameters:

- if we know ω and measure the value B_0^m of the field B_0 corresponding to the peak of the curve, we can deduce the value of the gyromagnetic ratio γ through the relation $\gamma = -\omega/B_0^m$.
- if we know γ, we can, by measuring the frequency $\omega/2\pi$ corresponding to resonance, measure the static magnetic field B_0. Various magnetometers, often of very great precision, operate on this principle. In certain cases, one can derive interesting information from such a measurement of the field. If, for example, the spin being considered is that of a nucleus which belongs to a molecule or to a crystal lattice, one can find the local field seen by the nucleus, its variation with the site occupied, etc.
- if we trace the square L^2 of the half-width as a function of ω_1^2, we obtain a straight line which, extrapolated to $\omega_1 = 0$, gives the lifetime τ of the excited level (*cf.* Fig. 5).

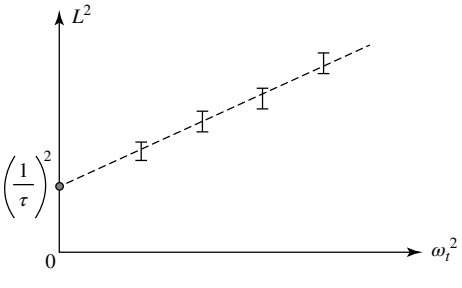


Figure 5: The extrapolation to $\omega_1 = 0$ of the squared half-width L of the resonance curve of Figure 4 gives the lifetime of the level being studied.

3. Relation between the classical treatment and the quantum mechanical treatment: evolution of $\langle \mathbf{M} \rangle$

The results obtained in §§ 1 and 2 are very similar, although we used classical mechanics in one case and quantum mechanics in the other. We are now going to show that this similarity is not accidental. It arises from the fact that the quantum mechanical evolution equations of the mean value of a magnetic moment placed in an arbitrary magnetic field are identical to the corresponding classical equations.

The mean value of the magnetic moment associated with a spin 1/2 is:

$$\langle \mathbf{M} \rangle(t) = \gamma \langle \mathbf{S} \rangle(t) \tag{31}$$

To calculate the evolution of $\langle \mathbf{M} \rangle(t)$, we use theorem (D-27) of Chapter III:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbf{M} \rangle(t) = \langle [\mathbf{M} \ , \ H(t)] \rangle \tag{32}$$

where $H(t)$ is the operator:

$$H(t) = -\mathbf{M} \cdot \mathbf{B}(t) \tag{33}$$

Let us calculate for example the commutator $[M_x, H(t)]$. Using the fact that the field components $B_y(t)$ and $B_z(t)$ are numbers (*cf.* note 1), we find:

$$\begin{aligned} [M_x, \ H(t)] &= -\gamma^2 [S_x, S_x B_x(t) + S_y B_y(t) + S_z B_z(t)] \\ &= -\gamma^2 B_y(t) [S_x, S_y] - \gamma^2 B_z(t) [S_x, S_z] \end{aligned} \tag{34}$$

Using relations (14) of Complement $\mathrm{A_{IV}}$, we obtain:

$$[M_x, H(t)] = i\hbar\gamma^2 [B_z(t) S_y - B_y(t) S_z] \tag{35}$$

Substituting (35) into (32):

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle M_x \rangle(t) = \gamma [B_z(t) \langle M_y \rangle(t) - B_y(t) \langle M_z \rangle(t)] \tag{36}$$

By cyclic permutation, we can calculate analogous expressions for the components on Oy and Oz; the three equations obtained can be condensed into:

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbf{M} \rangle(t) = \gamma \ \langle \mathbf{M} \rangle(t) \times \mathbf{B}(t) \tag{37}$$

Let us compare (37) with (6): the evolution of the mean value $\langle \mathbf{M} \rangle(t)$ obeys the classical equations exactly, whatever the time-dependence of the magnetic field $\mathbf{B}(t)$.

4. Bloch equations

In practice, in a magnetic resonance experiment, it is not the magnetic moment of a single spin that is observed, but rather that of a great number of identical spins (as in the experiment described in § 2-d above, where the number of atoms which decay from the state $|-\rangle$ is detected). Moreover, one is not concerned solely with the quantity $\mathscr{P}_{+-}(t)$, calculated above. One can also measure the global magnetization $\mathscr{M}$ of the sample under study: the sum of the mean values

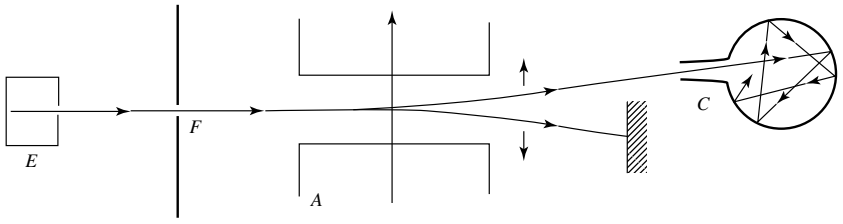


Figure 6: Schematic drawing of an experimental device which supplies cell C with atoms in the state $|+\rangle$.

of the observable $\langle \mathbf{M} \rangle$ corresponding to each spin of the sample[3]. It is interesting, therefore, to obtain the equations of motion of $\mathscr{M}$, called the *Bloch equations*.

In order to understand the physical significance of the various terms appearing in these equations, we are going to derive them for a simple concrete case. The results obtained can be generalized to other more complicated situations.

4-a. A concrete example

Consider a beam of atoms coming from an atomic polarizer of the type studied in § B-1-a of Chapter IV. All the atoms of the beam[4] are in the spin state $|+\rangle$ and therefore have their magnetic moments parallel to Oz. They enter a cell C through a small opening (Fig. 6), rebound a certain number of times from the inside walls of the cell and, after a certain time, escape through the same opening.

We shall denote by n the number of polarized atoms entering the cell per unit time; n is generally small and the atomic density inside the cell is low enough to allow atomic interactions to be neglected. Moreover, if the inside walls of the cell are suitably coated, collisions with the walls have little effect on the spin state of the atoms[5]. We shall assume that there is a probability per unit time $1/T_R$ for the elementary magnetization introduced into the cell by a polarized atom to disappear, either because of a depolarizing collision with the walls or simply because the atom has left the cell. T_R is called the "relaxation time". The cell is placed in a magnetic field $\mathbf{B}(t)$ which may have a static component and a rotating component. The problem consists of finding the equation of motion of the global magnetization $\mathscr{M}(t)$ of the atoms which are inside the cell at time t.

First, let us write the exact expression for $\mathscr{M}(t)$:

$$\mathscr{M}(t) = \sum_{i=1}^{\mathscr{N}} \langle \psi^{(i)}(t) | \mathbf{M} | \psi^{(i)}(t) \rangle = \sum_{i=1}^{\mathscr{N}} \mathscr{M}^{(i)}(t) \tag{38}$$

In (38), the sum is taken over the $\mathscr{N}$ spins which are already in the cell and which, at time t, have neither left nor undergone a depolarizing collision. $|\psi^{(i)}(t)\rangle$ is the state vector of such a spin (i) at time t [we are not counting, in (38), the spins which have undergone a depolarizing

[3] It is possible to detect, for example, the electromotive force emf induced in a coil by the variation of $\mathscr{M}$ with respect to time.

[4] For example, silver or hydrogen atoms in the ground state. For the sake of simplicity, all effects related to nuclear spin are neglected.

[5] For example, for hydrogen atoms bouncing off teflon walls, tens of thousands of collisions are required for the magnetic moment of the hydrogen atom to become disoriented.

collision and have not yet left the cell, since their global contribution is zero: their spins point randomly in all directions].

Between times t and $t + \mathrm{d}t$, $\boldsymbol{\mathscr{M}}(t)$ varies for three different reasons:

(i) A certain proportion, $\mathrm{d}t/T_R$, of the $\mathscr{N}$ spins undergo a depolarizing collision or leave the compartment; these spins disappear from the sum (38) and $\boldsymbol{\mathscr{M}}(t)$ therefore decreases by:

$$\mathrm{d}\boldsymbol{\mathscr{M}}(t) = -\frac{\mathrm{d}t}{T_R}\,\boldsymbol{\mathscr{M}}(t) \tag{39}$$

(ii) The other spins evolve freely in the field $\mathbf{B}(t)$. We saw in § 3 above that, for each of them, the evolution of the mean value of $\mathbf{M}$:

$$\boldsymbol{\mathscr{M}}^{(i)}(t) = \langle\psi^{(i)}(t)|\mathbf{M}|\psi^{(i)}(t)\rangle$$

obeys the classical law:

$$\mathrm{d}\boldsymbol{\mathscr{M}}^{(i)}(t) = \gamma\,\boldsymbol{\mathscr{M}}^{(i)}(t) \times \mathbf{B}(t)\mathrm{d}t \tag{40}$$

Since the right-hand side of (40) is linear with respect to $\boldsymbol{\mathscr{M}}^{(i)}(t)$, the contribution of these spins to the variation of $\boldsymbol{\mathscr{M}}(t)$ is given by:

$$\mathrm{d}\boldsymbol{\mathscr{M}}(t) = \gamma\,\boldsymbol{\mathscr{M}}(t) \times \mathbf{B}(t)\;\mathrm{d}t \tag{41}$$

(iii) Finally, a certain number, $n\mathrm{d}t$, of new spins have entered the cell. Each of them adds to the global magnetization a contribution $\boldsymbol{\mu}_0$, equal to the mean value of $\mathbf{M}$ in the state $|+\rangle$ ($\boldsymbol{\mu}_0$ is parallel to Oz and $|\boldsymbol{\mu}_0| = |\gamma|\,\hbar/2$). $\boldsymbol{\mathscr{M}}$ therefore increases by:

$$\mathrm{d}\boldsymbol{\mathscr{M}}(t) = n\boldsymbol{\mu}_0\mathrm{d}t \tag{42}$$

The global variation of $\boldsymbol{\mathscr{M}}$ is obtained by adding (39), (41) and (42). Dividing by $\mathrm{d}t$, we obtain the equation of motion of $\boldsymbol{\mathscr{M}}(t)$ (Bloch equation):

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathscr{M}}(t) = n\boldsymbol{\mu}_0 - \frac{1}{T_R}\boldsymbol{\mathscr{M}}(t) + \gamma\boldsymbol{\mathscr{M}}(t) \times \mathbf{B}(t) \tag{43}$$

We have derived (43) in a specific case, making certain hypotheses. However, the main features of this equation remain valid for a great number of other experiments where the rate of variation of $\boldsymbol{\mathscr{M}}(t)$ appears in the form of a sum of three terms:

- a source term (here $n\boldsymbol{\mu}_0$) which describes the preparation of the system. It would, in fact, be impossible to observe magnetic resonance without a preliminary polarization of the spins, which can be achieved through selection using a magnetic field gradient (as in the example studied here), a polarized optical excitation (as in the example studied in § 2-d above), cooling of the sample in a strong static field, etc.
- a damping term $\left(\text{here} - \frac{1}{T_R}\boldsymbol{\mathscr{M}}(t)\right)$ which describes the disappearance or "relaxation" of the global magnetization under the effect of various processes: collisions, disappearance of atoms, change in atomic levels through spontaneous emission (as in the example studied in § 2-d), etc.
- a term which describes the precession of $\boldsymbol{\mathscr{M}}(t)$ in the field $\mathbf{B}(t)$ [last term of (43)].

4-b. Solution in the case of a rotating field

When the field $\mathbf{B}(t)$ is the sum of a static field $\mathbf{B}_0$ and a rotating field $\mathbf{B}_1(t)$, such as those considered above, equations (43) can be solved exactly. As in §§ 1 and 2, one changes to the rotating frame $OXYZ$, with respect to which the relative variation of $\boldsymbol{\mathscr{M}}(t)$ is:

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathscr{M}}\right)_{\mathrm{rel}} = n\boldsymbol{\mu}_0 - \frac{1}{T_R}\boldsymbol{\mathscr{M}} + \gamma\boldsymbol{\mathscr{M}} \times \mathbf{B}_{\mathrm{eff}} \tag{44}$$

where $\mathbf{B}_{\mathrm{eff}}$ is defined by equation (10).

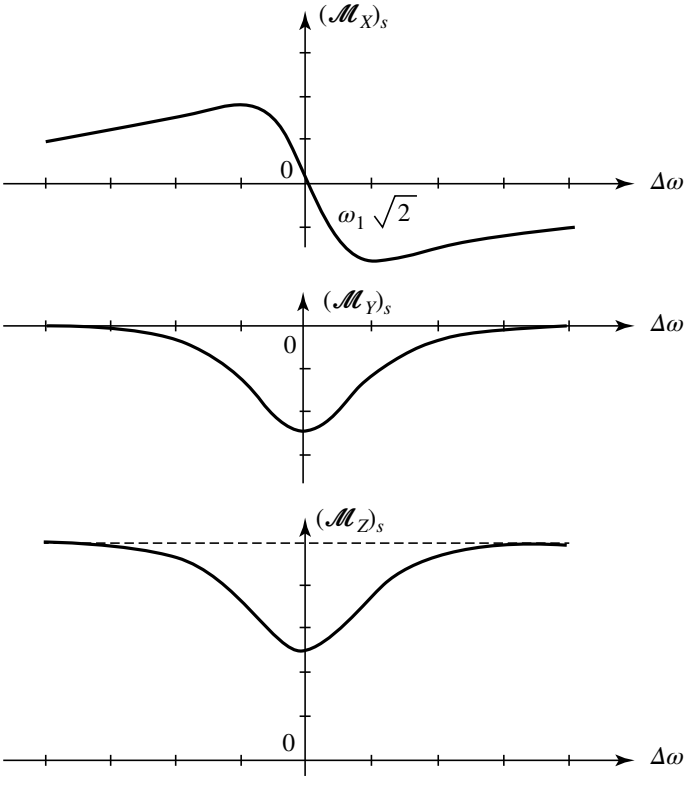


Figure 7: Variation with respect to $\Delta\omega = \omega - \omega_0$ of the stationary values of the components of $\boldsymbol{\mathscr{M}}$ in the rotating frame. One obtains a dispersion curve for $(\mathscr{M}_X)_S$ and absorption curves for $(\mathscr{M}_Y)_S$ and $(\mathscr{M}_Z)_S$. The three curves have the same width, $2\sqrt{\omega_1^2 + (1/T_R)^2}$, which increases with ω_1. They have been drawn assuming that $\omega_1 = 1/T_R$ ("half-saturation").

Projecting this equation onto OX, OY and OZ, we obtain a system of three linear differential equations with constant coefficients, whose stationary solution (valid after a time

much greater than T_R) is:

$$
\begin{aligned}
(\mathscr{M}_X)_S &= -n\boldsymbol{\mu}_0 T_R \, \frac{\omega_1 \Delta\omega}{(\Delta\omega)^2 + \omega_1^2 + (1/T_R)^2} \\
(\mathscr{M}_Y)_S &= -n\mu_0 T_R \, \frac{\omega_1}{(\Delta\omega)^2 + \omega_1^2 + (1/T_R)^2} \\
(\mathscr{M}_Z)_S &= n\mu_0 T_R \left[1 - \frac{\omega_1^2}{(\Delta\omega)^2 + \omega_1^2 + (1/T_R)^2} \right]
\end{aligned}
\tag{45}
$$

The three components of the stationary magnetization $(\boldsymbol{\mathscr{M}})_S$, when the field B_0 varies, have resonant variations about the value $B_0 = -\omega/\gamma$ (*cf.* Fig. 7). $(\mathscr{M}_Y)_S$ and $(\mathscr{M}_Z)_S$ follow absorption curves (Lorentz curves of width $2\sqrt{\omega_1^2 + (1/T_R)^2/\gamma}$). $(\boldsymbol{\mathscr{M}}_X)_S$ follows a dispersion curve (of the same width).

References and suggestions for further reading:

Feynman II (7.2), Chap. 35; Cagnac and Pebay-Peyroula (11.2), Chaps. IX § 5, X § 5, XI §§ 2 to 5, XIX § 3; Kuhn (11.1), § VI, D.

See the references of section 14 of the bibliography, particularly Abragam (14.1) and Slichter (14.2).

Complement G_{IV}

A simple model of the ammonia molecule

1 Description of the model 469
2 Eigenfunctions and eigenvalues of the Hamiltonian 471
2-a Infinite potential barrier 471
2-b Finite potential barrier 473
2-c Evolution of the molecule. Inversion frequency 476
3 The ammonia molecule considered as a two-level system . . 477
3-a The state space 478
3-b Energy levels. Removal of the degeneracy due to the transparency of the potential barrier 479
3-c Influence of a static electric field 481

1. Description of the model

In the ammonia molecule NH_3, the three hydrogen atoms form the base of a pyramid whose apex is the nitrogen atom (*cf.* Fig. 1). We shall study this molecule by using a simplified model with the following features: the nitrogen atom (of mass m_N), much heavier than its partners (of mass m_H), is motionless; the hydrogen atoms form a rigid equilateral triangle whose axis always passes through the nitrogen atom. The potential energy of the system is thus a function of only one parameter, the (algebraic) distance x between the nitrogen atom and the plane defined by the three hydrogen atoms[1]. The

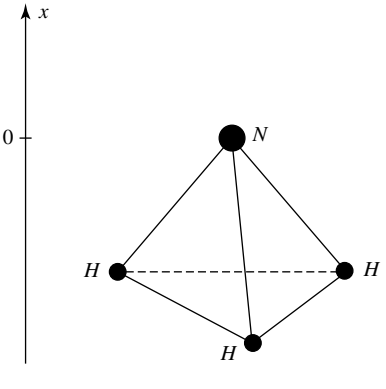


Figure 1: Schematic drawing of the ammonia molecule; x is the algebraic distance between the plane of the hydrogen atoms and the nitrogen atom, which is assumed to be motionless.

shape of this potential energy $V(x)$ is given by the solid-line curve in Figure 2. The

[1] In this one-dimensional model, effects linked to the rotation of the molecule are obviously not taken into account.

symmetry of the problem with respect to the $x = 0$ plane requires $V(x)$ to be an even function of x. The two minima of $V(x)$ correspond to two symmetrical configurations of the molecule in which, classically, it is stable; we shall choose the energy origin such that its energy is then zero. The potential barrier at $x = 0$, of height V_1, expresses the fact that, if the nitrogen atom is in the plane of the hydrogen atoms, they repel it. Finally, the increase in $V(x)$ when $|x|$ is greater than b corresponds to the chemical bonding force which insures the cohesion of the molecule.

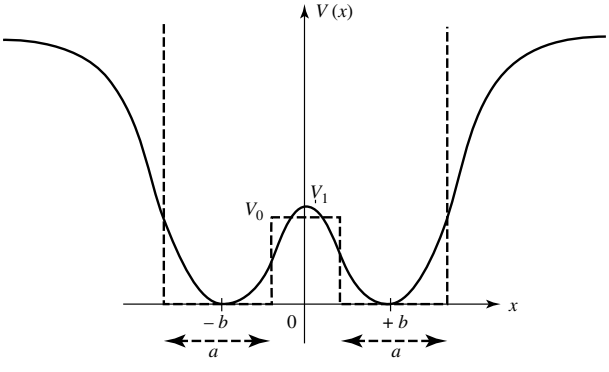


Figure 2: Variation with respect to x of the potential energy $V(x)$ of the molecule. $V(x)$ has two minima (classical equilibrium positions), separated by a potential barrier due to the repulsion for small $|x|$ between the nitrogen atom and the three hydrogen atoms. The "square potential" used to approximate $V(x)$ is shown in dashed lines.

This model therefore reduces the problem to a one-dimensional one in which a fictitious particle of mass m (it can be shown that the "reduced mass" m of the system is equal to $\frac{3m_H m_N}{3m_H+m_N}$) is under the influence of the potential $V(x)$. Under these conditions, what are the energy levels predicted by quantum mechanics? With respect to classical predictions, two major differences appear:

(*i*) The Heisenberg uncertainty relation forbids the molecule to have an energy equal to the minimum of $V(x)$ ($V_{min} = 0$ in our case). We have already seen, in Complements C_I and M_{III} why this energy must be greater than $V_{\min}$.

(*ii*) Classically, the potential barrier at $x = 0$ cannot be cleared by a particle whose energy is less than V_1: the nitrogen atom thus always remains on the same side of the plane of the hydrogen atoms, and the molecule cannot invert itself. Quantum mechanically, such a particle can cross this barrier by the tunnel effect (*cf.* Chap. I, § D-2-c): the inversion of the molecule is therefore always possible. We are going to discuss the consequences of this effect.

We are concerned here only with a qualitative discussion of the physical phenomena and not with an exact quantitative calculation which would not have much significance in this approximate model. For example, we shall try to demonstrate the existence of an inversion frequency of the ammonia molecule, without giving an exact or even an approximate value of this frequency. We shall therefore simplify the problem further by replacing the function $V(x)$ by the square potential drawn in dashed lines in Figure 2 [two infinite potential steps at $x = \pm(b + a/2)$ and a potential barrier of height V_0 centered at $x = 0$ and of width $(2b - a)$].

2. Eigenfunctions and eigenvalues of the Hamiltonian

2-a. Infinite potential barrier

Before calculating the eigenfunctions and eigenvalues of the Hamiltonian corresponding to the "square" potential of Figure 2, we are going to assume, in this first stage, that the potential barrier V_0 is infinite (in which case, no tunnel effect is possible). This will lead us to a better understanding of the consequences of the tunnel effect across the finite potential barrier of Figure 2. We shall therefore consider, to begin with, a particle in a potential $\widetilde{V}(x)$ composed of two infinite wells of width a centered at $x = \pm b$ (Fig. 3). If the particle is in one of these two wells, it obviously cannot go into the other one.

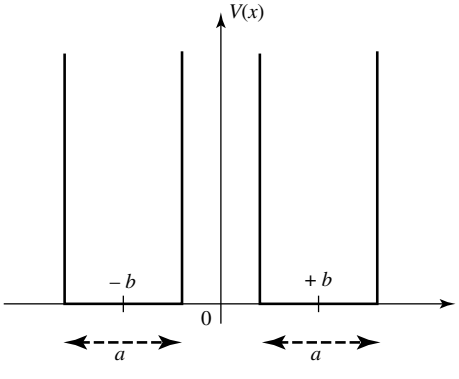


Figure 3: When the height V_0 of the potential barrier of Figure 2 is large, we have two practically infinite potential wells of width a whose centers are separated by a distance of $2b$.

Each of the two wells of Figure 3 is similar to the one studied in Complement H_I, in § 2-c-β. We can therefore use the results obtained in that complement. The possible energies of the particle are:

$$E_n = \frac{\hbar^2 k_n^2}{2m} \tag{1}$$

with:

$$k_n = \frac{n\pi}{a} \tag{2}$$

(where n is a positive integer). Each of the energy values is twofold degenerate, since two wave functions correspond to it:

$$\varphi_1^n(x) = \begin{cases} \sqrt{\dfrac{2}{a}} \sin\left[k_n\left(b + \dfrac{a}{2} - x\right)\right] & \text{if} \quad b - \dfrac{a}{2} \leq x \leq b + \dfrac{a}{2} \\ 0 & \text{everywhere else.} \end{cases}$$

$$\varphi_2^n(x) = \begin{cases} \sqrt{\dfrac{2}{a}} \sin\left[k_n\left(b + \dfrac{a}{2} + x\right)\right] & \text{if} \quad b - \dfrac{a}{2} \leq -x \leq b + \dfrac{a}{2} \\ 0 & \text{everywhere else.} \end{cases} \tag{3}$$

In the state $|\varphi_1^n\rangle$, the particle is in the infinite well on the right; in the state $|\varphi_2^n\rangle$, it is in the one on the left.

Figure 4 shows the first two energy levels of the molecule, which are two-fold degenerate. The Bohr frequency $(E_2-E_1)/h$ associated with these two levels corresponds, as we saw in Complement A_{III} (§ 2-b), to the to-and-fro motion of the particle between the two sides of the well on the right (or on the left) when its state is a linear superposition of $|\varphi_1^1\rangle$ and $|\varphi_1^2\rangle$ (or of $|\varphi_2^1\rangle$ and $|\varphi_2^2\rangle$). Physically, such an oscillation represents a molecular vibration of the plane of the three hydrogen atoms about its stable equilibrium position, which corresponds to $x = +b$ (or $x = -b$). The frequency of this oscillation falls in the infrared part of the spectrum.

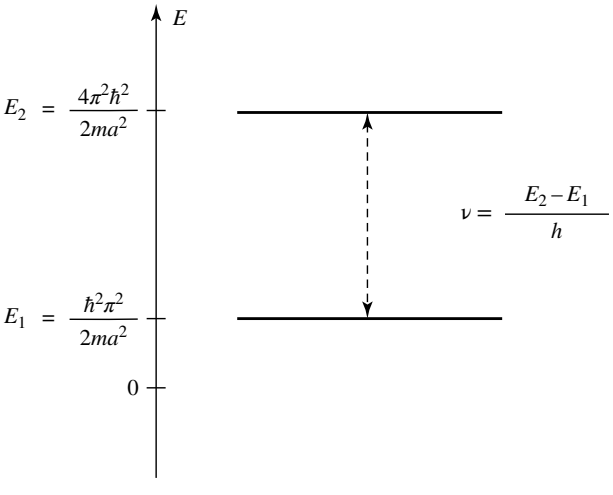


Figure 4: First energy levels obtained in the potential wells of Figure 3. The oscillation of the system in one of the two wells at the Bohr frequency $\nu = (E_2 - E_1)/h$ represents the vibration of the molecule about one of its two classical equilibrium positions.

In the rest of the calculation, it is convenient to change bases, in each of the eigensubspaces of the Hamiltonian of the particle. Since the function $V(x)$ is even, this Hamiltonian H commutes with the parity operator Π (*cf.* Complement F_{II}, § 4). In this case, we can find a basis of eigenvectors of H that are even or odd; the wave functions of these vectors are the symmetrical and antisymmetrical linear combinations:

$$\begin{aligned} \varphi_s^n(x) &= \frac{1}{\sqrt{2}}[\varphi_1^n(x) + \varphi_2^n(x)] \\ \varphi_a^n(x) &= \frac{1}{\sqrt{2}}[\varphi_1^n(x) - \varphi_2^n(x)] \end{aligned} \tag{4}$$

In the states $|\varphi_s^n\rangle$ and $|\varphi_a^n\rangle$, the particle can be found in one or the other of the two potential wells.

In what follows, we shall confine ourselves to the study of the ground state, for which the wave functions $\varphi_1^1(x)$, $\varphi_2^1(x)$, $\varphi_s^1(x)$ and $\varphi_a^1(x)$ are shown in Figure 5.

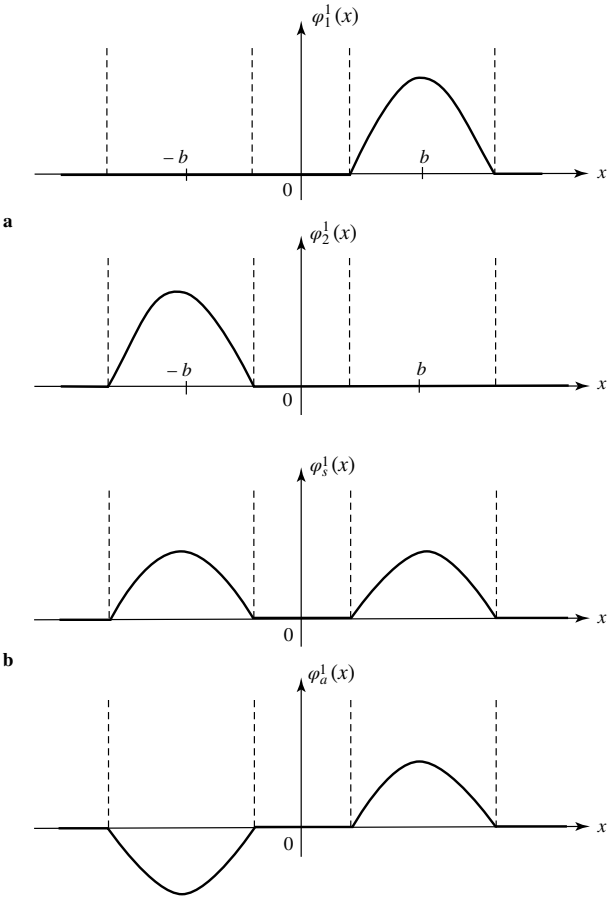


*Figure 5: The states $\varphi_1^1(x)$ and $\varphi_2^1(x)$, shown in figure **a**, are stationary states with the same energy, respectively localized in the right-hand well and the left-hand well of Figure 3. To use the symmetry of the problem, it is more convenient to choose as stationary states the symmetrical state $\varphi_s^1(x)$ and the antisymmetrical state $\varphi_a^1(x)$, linear combinations of $\varphi_1^1(x)$ and $\varphi_2^1(x)$ (figure **b**)*

2-b. Finite potential barrier

Let us try to find the shape of the eigenfunctions of the first energy levels when V_0 has a finite value (assumed, nevertheless, to be greater than the energy of these levels).

Inside the two "square" potential wells (dashed lines in Figure 2), $V(x) = 0$. The wave function is therefore of the form:

$$\begin{aligned} \chi(x) &= A \sin\left[k\left(b + \frac{a}{2} - x\right)\right] \quad \text{if} \quad b - \frac{a}{2} \leq x \leq b + \frac{a}{2} \\ \chi(x) &= A' \sin\left[k\left(b + \frac{a}{2} + x\right)\right] \quad \text{if} \quad b - \frac{a}{2} \leq -x \leq b + \frac{a}{2} \end{aligned} \tag{5}$$

where k is related to the energy E of the level by the relation:

$$E = \frac{\hbar^2 k^2}{2m} \tag{6}$$

As in the preceding paragraph, $\chi(x)$ always goes to zero at $x = \pm(b + a/2)$, since $V(x)$ becomes infinite at these two points. On the other hand, since V_0 is finite, $\chi(x)$ no longer goes to zero at $x = \pm(b - a/2)$; consequently, k no longer satisfies relation (2).

Once again, since $V(x)$ is even, we can look for eigenfunctions of the Hamiltonian, $\chi_s(x)$ and $\chi_a(x)$, which are respectively even and odd. Let us denote by A_s and A'_s, A_a and A'_a the values of the coefficients A and A', introduced in (5), which correspond to $\chi_s(x)$ and $\chi_a(x)$. We have, obviously:

$$\begin{aligned} A'_s &= A_s \\ A'_a &= -A_a \end{aligned} \tag{7}$$

The eigenvalues associated with χ_s and χ_a will be denoted by E_s and E_a, which enables us, using (6), to define the corresponding values k_s and k_a of the parameter k.

In the interval $-(b-a/2) \leq x \leq +(b-a/2)$, the wave function is no longer zero, as it was before, since V_0 is finite. It must be a linear combination, even or odd depending on whether we are considering χ_s or χ_a, of exponentials $e^{q_{s,a}x}$ and $e^{-q_{s,a}x}$; q_s and q_a are defined in terms of $E_{s,a}$ and V_0 by:

$$q_{s,a} = \sqrt{\frac{2m}{\hbar^2}(V_0 - E_{s,a})} = \sqrt{\alpha^2 - k_{s,a}^2} \tag{8}$$

with:

$$V_0 = \frac{\hbar^2 \alpha^2}{2m} \tag{9}$$

Therefore, for $-(b - a/2) \leq x \leq (b - a/2)$, the functions χ_s and χ_a are written:

$$\begin{aligned} \chi_s(x) &= B_s \cosh(q_s x) \\ \chi_a(x) &= B_a \sinh(q_a x) \end{aligned} \tag{10}$$

Finally, we must match the eigenfunctions and their derivatives at $x = \pm(b-a/2)$. The even solution $\chi_s(x)$ must therefore satisfy the conditions:

$$\begin{aligned} A_s \sin(k_s a) &= B_s \cosh\left[q_s\left(b - \frac{a}{2}\right)\right] \\ -A_s k_s \cos(k_s a) &= B_s q_s \sinh\left[q_s\left(b - \frac{a}{2}\right)\right] \end{aligned} \tag{11}$$

Since A_s and B_s cannot be zero simultaneously, we can take the ratio of equations (11) to obtain:

$$\tan(k_s a) = -\frac{k_s}{q_s} \coth\left[q_s\left(b - \frac{a}{2}\right)\right] \tag{12}$$

For the odd solution $\chi_a(x)$, we obtain in the same way:

$$\tan(k_a a) = -\frac{k_a}{q_s} \tanh\left[q_a\left(b - \frac{a}{2}\right)\right] \tag{13}$$

If q_s and q_a are replaced by their values in terms of k_s and k_a, relations (12) and (13) can be written:

$$\tan(k_s a) = -\frac{k_s}{\sqrt{\alpha^2 - k_s^2}} \coth\left[\sqrt{\alpha^2 - k_s^2}\left(b - \frac{a}{2}\right)\right] \tag{14}$$

and:

$$\tan(k_a a) = -\frac{k_a}{\sqrt{\alpha^2 - k_a^2}} \tanh\left[\sqrt{\alpha^2 - k_a^2}\left(b - \frac{a}{2}\right)\right] \tag{15}$$

In theory, therefore, the problem is solved. Relations (14) and (15) express the energy quantization since they give the possible values of k_s and k_a and therefore, thanks to relation (6), the energies E_s and E_a (with the condition that they be less than V_0). The transcendental equations (14) and (15) can be solved graphically. A certain number of roots are found: k_s^1, k_s^2,..., k_a^1, k_a^2... The root k_s^n is different from k_a^n, since equations (14) and (15) are not the same: *the energies E_s^n and E_a^n are therefore different.* Of course, when V_0 becomes very large, k_s^n and k_a^n both approach the value $n\pi/a$ found in the preceding section; this can be seen by letting α approach infinity in equations (14) and (15), which yields $\tan(k_{s,a}a) = 0$, an equation equivalent to (2). The energies E_s^n and E_a^n therefore approach the value $E_n = \hbar^2 n^2 \pi^2 / 2ma^2$ calculated in the preceding section for V_0 approaching infinity. Finally, it is easy to see that, the more V_0 exceeds E_n, the closer together the two energies E_s^n and E_a^n will be.

The exact values of E_s^n and E_a^n are of little importance to us here. We shall content ourselves with sketching the shape of the energy spectrum in Figure 6, which shows what happens to the energies of levels E_1 and E_2 of Figure 4 when the finite height V_0 of the potential barrier is taken into account. We see that the tunnel effect across this barrier removes the degeneracy of E_1 and E_2, giving rise to doublets, (E_s^1, E_a^1) and (E_s^2, E_a^2) (assuming, of course, that all these energies are less than V_0). Since the (E_s^1, E_a^1) doublet is the deeper one, it is clear that $|E_s^1 - E_a^1| < |E_s^2 - E_a^2|$. Finally, the distance between the doublets is much greater than the spacing within each doublet (experimentally, their ratio is of the order of a thousand). These spacings enable us, moreover, to define new Bohr (angular) frequencies:

$$\Omega_1 = \frac{E_a^1 - E_s^1}{\hbar} \quad , \quad \Omega_2 = \frac{E_a^2 - E_s^2}{\hbar}, \ldots$$

whose physical significance we shall study in the next paragraph (the corresponding transitions are represented by arrows in Figure 6).

Figure 7 shows the shape of the eigenfunctions $\chi_s^1(x)$ and $\chi_a^1(x)$, which are given by equations (5), (7) and (10) once k_s^1 and k_a^1 have been determined from (14) and (15). We see that they greatly resemble the functions $\varphi_s^1(x)$ and $\varphi_a^1(x)$ of Figure 5, the essential difference being that the wave function is no longer zero in the interval $-(b - a/2) \leq x \leq (b - a/2)$. The reason for introducing the φ_s^1 and φ_a^1 basis in the preceding paragraph can now be understood: the eigenfunctions χ_s^1 and χ_a^1, in the presence of the tunnel effect, resemble φ_s^1 and φ_a^1 much more than φ_1^1 and φ_2^1.

2-c. Evolution of the molecule. Inversion frequency

Assume that at time $t = 0$, the molecule is in the state:

$$|\psi(t=0)\rangle = \frac{1}{\sqrt{2}}\left[|\chi_s^1\rangle + |\chi_a^1\rangle\right] \tag{16}$$

The state vector $|\psi(t)\rangle$ at time t can be obtained by using the general formula (D-54) of Chapter III; we obtain:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} e^{-i\frac{E_s^1+E_a^1}{2\hbar}t}\left[e^{+i\Omega_1 t/2}|\chi_s^1\rangle + e^{-i\Omega_1 t/2}|\chi_a^1\rangle\right] \tag{17}$$

From this we deduce the probability density:

$$|\psi(x,t)|^2 = \frac{1}{2}\left[\chi_s^1(x)\right]^2 + \frac{1}{2}\left[\chi_a^1(x)\right]^2 + \cos(\Omega_1 t)\chi_s^1(x)\chi_a^1(x) \tag{18}$$

The variation with respect to time of this probability density is simple to obtain graphically from the curves of Figure 7. They are shown in Figure 8. For $t = 0$ (Fig. 8-a), we see that the initial state chosen in (16) corresponds to a probability density which is concentrated in the right-hand well (in the left-hand well, the functions χ_s^1 and χ_a^1 are of opposite sign and very close in absolute value, so their sum is practically zero). It can therefore be said that the particle, initially, is practically in the right-hand well. At time $t = \pi/2\Omega_1$ (Fig. 8-b), it has moved appreciably, through the tunnel effect, into the left-hand well, is practically there at time $t = \pi/\Omega_1$ (Fig. 8-c), and then performs the process in reverse (Figures 8-d and 8-e).

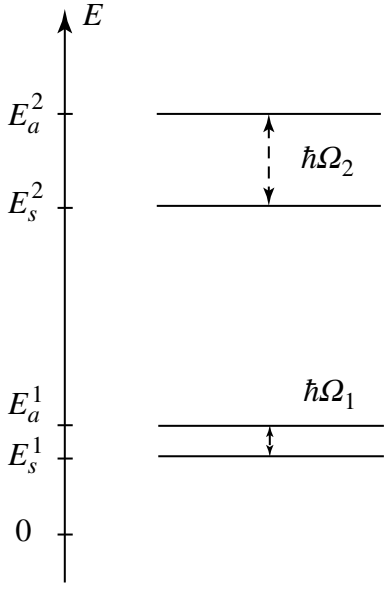


Figure 6: When one takes the finite height V_0 of the barrier into account, one finds that the energy spectrum of Figure 4 is modified: each level splits into two distinct ones. The Bohr frequencies $\Omega_1/2\pi$ and $\Omega_2/2\pi$ corresponding to tunnelling from one well to the other are the inversion frequencies of the ammonia molecule for the first two vibration levels. The tunnel effect is more important in the higher vibration level, so $\Omega_2 > \Omega_1$.

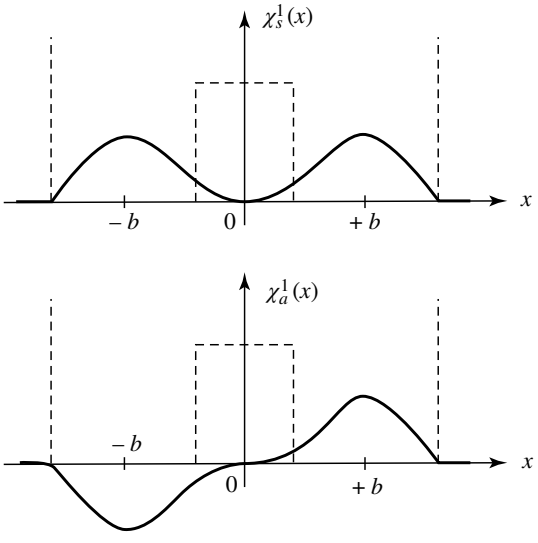


Figure 7: Wave functions associated with the levels E_s^1 and E_a^1 in Figure 6. Note the analogy with the functions in Figure 5-b; however, these new functions do not vanish on the interval $-b+a/2 \leq x \leq b-a/2$.

The fictitious particle therefore moves from one side of the potential barrier to the other with the frequency $\Omega_1/2\pi$, which means that the plane of the hydrogen atoms continually passes from one side of the nitrogen atom to the other. This is why the frequency $\Omega_1/2\pi$ is called the *inversion frequency* of the molecule. Note that this inversion frequency has no classical analogue, since its existence is related to the tunnel effect of the fictitious particle across the potential barrier.

Since the nitrogen atom tends to attract the electrons of the three hydrogen atoms, the ammonia molecule possesses an electric dipole moment which is proportional to the mean value $\langle X \rangle$ of the position of the fictitious particle we have studied; we see in Figure 8 that this dipole moment is an oscillating function of time. Under these conditions, the ammonia molecule is capable of emitting or absorbing electromagnetic radiation of frequency $\Omega_1/2\pi$.

Experimentally, this is indeed observed; the value of Ω_1 falls in the domain of centimeter waves. In radioastronomy, ammonia molecules in interstellar space have been shown to emit and absorb electromagnetic waves of this frequency. Let us also point out that the principle of the ammonia maser is based on the stimulated emission of these waves by the NH_3 molecule.

3. The ammonia molecule considered as a two-level system

We see in Figure 6 that we have a situation which is analogous to the one mentioned in the introduction of § C of Chapter IV. The system under study possesses two levels, E_s^1 and E_a^1, which are very close to each other and very far from all other levels E_s^2, E_a^2, ... If we are interested only in the two levels E_s^1 and E_a^1, we can "forget" all the others (the exact justification for such an approximation will be given in the framework of perturbation theory in Chapter XI).

We are going to return to the preceding discussion with a slightly different point of view and show that the general considerations of Chapter IV concerning two-level

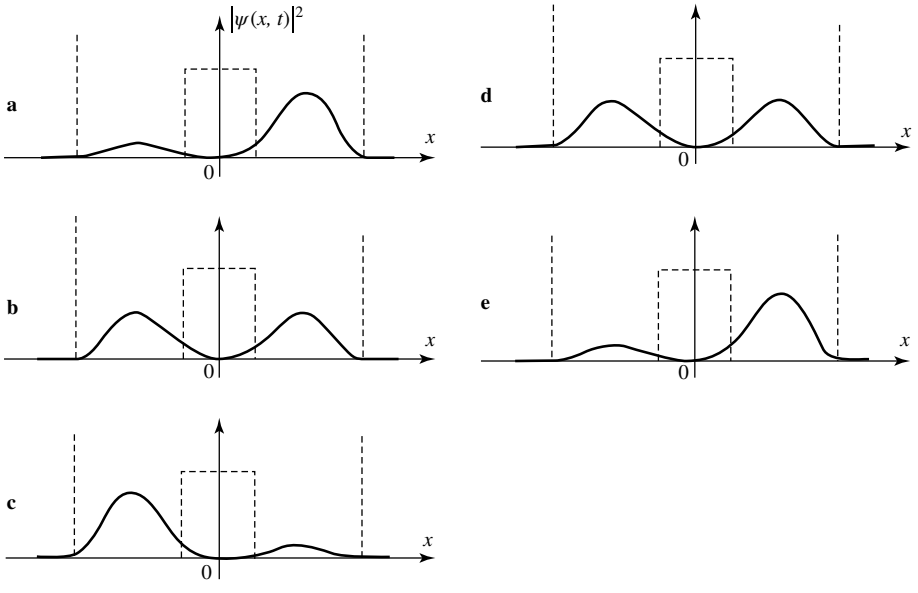


Figure 8: Evolution of a wave packet obtained by superposing the two stationary wave functions of Figure 7. The particle, initially, is in the right-hand well (fig. ***a****), tunnels into the left-hand well (fig.* ***b****) and, after a certain time, becomes localized there (fig.* ***c****); then it returns to the right-hand well (fig.* ***d****) and the initial state (fig.* ***e****), and so on.*

systems can be applied to the ammonia molecule. This point of view will also enable us to study very simply the effect of a static external electric field on this molecule.

3-a. The state space

The state space we are going to consider is spanned by the two orthogonal vectors $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$, whose wave functions are given by (3). As we explained above, we shall ignore the other states $|\varphi_1^n\rangle$ and $|\varphi_2^n\rangle$ for which $n > 1$. In the states $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$, the nitrogen atom is either above or below the plane of the hydrogen atoms. We introduced in (4) a second orthonormal basis of the state space, composed of the even and odd vectors:

$$\begin{aligned} |\varphi_s^1\rangle &= \frac{1}{\sqrt{2}}\left[|\varphi_1^1\rangle + |\varphi_2^1\rangle\right] \\ |\varphi_a^1\rangle &= \frac{1}{\sqrt{2}}\left[|\varphi_1^1\rangle - |\varphi_2^1\rangle\right] \end{aligned} \tag{19}$$

There is the same probability in these two states of finding the nitrogen atom above or below the plane of the hydrogen atoms.

3-b. Energy levels. Removal of the degeneracy due to the transparency of the potential barrier

When the height V_0 of the potential barrier is infinite, the states $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$ have the same energy (as do the states $|\varphi_s^1\rangle$ and $|\varphi_a^1\rangle$), so that H_0, the Hamiltonian of the system, is written:

$$H_0 = E_1 \times \mathbb{1} \tag{20}$$

(where $\mathbb{1}$ is the identity operator in the two-dimensional state space).

To take into account phenomenologically the fact that the barrier is not infinite, we add to H_0 a perturbation W which is non-diagonal in the $|\varphi_1^1\rangle, |\varphi_2^1\rangle$ basis and is represented by the matrix:

$$W = -A \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{21}$$

where A is a real positive coefficient[2].

If we want to find the stationary states of the molecule, we must diagonalize the total Hamiltonian operator $H = H_0 + W$, whose matrix is written:

$$H = \begin{pmatrix} E_1 & -A \\ -A & E_1 \end{pmatrix} \tag{22}$$

A simple calculation gives the eigenvalues and eigenvectors of H:

$$\begin{array}{lll} E_1 + A & \text{corresponding to the eigenket} & |\varphi_a^1\rangle \\ E_1 - A & \quad '' \quad '' \quad '' & |\varphi_s^1\rangle \end{array} \tag{23}$$

We see that, under the effect of the perturbation W, the two levels, which were degenerate when A was zero, now split; an energy difference, equal to $2A$, appears, and the new eigenstates are the states $|\varphi_s^1\rangle$ and $|\varphi_a^1\rangle$. We again find the results of § 2.

If, at time $t = 0$, the molecule is in the state $|\varphi_1^1\rangle$:

$$|\psi(t=0)\rangle = |\varphi_1^1\rangle = \frac{1}{\sqrt{2}} \left[|\varphi_s^1\rangle + |\varphi_a^1\rangle \right] \tag{24}$$

the state vector at time t will be:

$$\begin{aligned} |\psi(t)\rangle &= \frac{1}{\sqrt{2}} \mathrm{e}^{-iE_1 t/\hbar} \left[\mathrm{e}^{iAt/\hbar} |\varphi_s^1\rangle + \mathrm{e}^{-iAt/\hbar} |\varphi_a^1\rangle \right] \\ &= \mathrm{e}^{-iE_1 t/\hbar} \left[\cos\left(\frac{At}{\hbar}\right) |\varphi_1^1\rangle + i \sin\left(\frac{At}{\hbar}\right) |\varphi_2^1\rangle \right] \end{aligned} \tag{25}$$

[2] We are forced to assume $A > 0$ in order to obtain the relative disposition of the E_s^1 and E_a^1 levels of Figure 6 [see eigenvalues (23)].

In a measurement performed at time t, we therefore have a probability $\cos^2(At/\hbar)$ of finding the molecule in the state $|\varphi_1^1\rangle$ (the nitrogen atom above the plane of the hydrogen atoms) and $\sin^2(At/\hbar)$ of finding it in the state $|\varphi_2^1\rangle$ (the nitrogen atom below). Thus we again find that, under the effect of the coupling W, the ammonia molecule inverts periodically.

Comment:

The perturbation W [given in (21)] describes (phenomenologically) the fact that the potential barrier is finite. This approach is less precise than the discussion above, since we obtain here eigenfunctions $\varphi_s^1(x)$ and $\varphi_a^1(x)$ which, unlike χ_s^1 and χ_a^1, go to zero in the region $(-b + a/2) \leq x \leq (b - a/2)$. This much more simple description nevertheless explains two fundamental physical effects: the removal of the degeneracy of E_1 and the periodic oscillation of the molecule between the states $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$ (inversion).

3-c. Influence of a static electric field

We saw above that, in the states $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$, the electric dipole moment of the molecule takes on two opposite values, which we shall denote by $+\eta$ and $-\eta$. If we call D the observable associated with this physical quantity, we can therefore assume that D is represented in the $\{|\varphi_1^1\rangle, |\varphi_2^1\rangle\}$ basis by a diagonal matrix whose eigenvalues are $+\eta$ and $-\eta$:

$$D = \begin{pmatrix} \eta & 0 \\ 0 & -\eta \end{pmatrix} \tag{26}$$

When the molecule is placed in a static electric field[3] $\mathscr{E}$, the interaction energy with this field is:

$$W'(\mathscr{E}) = -\mathscr{E} D \tag{27}$$

This term of the Hamiltonian[4] is represented in the $\{|\varphi_1^1\rangle, |\varphi_2^1\rangle\}$ basis by the matrix:

$$W'(\mathscr{E}) = -\eta\mathscr{E} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{28}$$

Let us then write the matrix which represents, in the $\{|\varphi_1^1\rangle, |\varphi_2^1\rangle\}$ basis, the total Hamiltonian operator of the molecule, $H_0 + W + W'(\mathscr{E})$:

$$H_0 + W + W'(\mathscr{E}) = \begin{pmatrix} E_1 - \eta\mathscr{E} & -A \\ -A & E_1 + \eta\mathscr{E} \end{pmatrix} \tag{29}$$

This matrix can easily be diagonalized; its eigenvalues E_+ and E_- and its eigenvectors $|\psi_+\rangle$ and $|\psi_-\rangle$ are given by:

$$\begin{aligned} E_+ &= E_1 + \sqrt{A^2 + \eta^2\mathscr{E}^2} \\ E_- &= E_1 - \sqrt{A^2 + \eta^2\mathscr{E}^2} \end{aligned} \tag{30}$$

[3] For the sake of simplicity, we assume here that this field is parallel to the Ox axis of Figure 1 (one-dimensional model).

[4] In $W'(\mathscr{E})$, D is an observable, while $\mathscr{E}$ is a classical quantity which is externally imposed (*cf.* note, page 458).

and:

$$\begin{aligned} |\psi_+\rangle &= \cos\frac{\theta}{2}\ |\varphi_1^1\rangle - \sin\frac{\theta}{2}\ |\varphi_2^1\rangle \\ |\psi_-\rangle &= \sin\frac{\theta}{2}\ |\varphi_1^1\rangle + \cos\frac{\theta}{2}\ |\varphi_2^1\rangle \end{aligned} \tag{31}$$

where we have set:

$$\tan\ \theta = -\frac{A}{\eta\mathscr{E}} \qquad (0 \leq \theta < \pi) \tag{32}$$

[*cf.* Complement B_{IV}, relations (9), (10), (22) and (23); since A is real and negative, the angle φ introduced in that complement is here equal to π].

When $\mathscr{E}$ is zero, $\theta = \pi/2$, and we again obtain the results of § 3-b, since:

$$\begin{aligned} E_+(\mathscr{E}=0) &= E_1 + A \\ E_-(\mathscr{E}=0) &= E_1 - A \end{aligned} \tag{33}$$

with:

$$\begin{aligned} |\psi_+(\mathscr{E}=0)\rangle &= |\varphi_a^1\rangle \\ |\psi_-(\mathscr{E}=0)\rangle &= |\varphi_s^1\rangle \end{aligned} \tag{34}$$

When, for an arbitrary $\mathscr{E}$, A is zero (a perfectly opaque potential barrier), we obtain:

$$\begin{aligned} E_+(A=0) &= E_1 + \eta|\mathscr{E}| \\ E_-(A=0) &= E_1 - \eta|\mathscr{E}| \end{aligned} \tag{35}$$

with, if $\mathscr{E}$ is positive[5]:

$$\begin{aligned} |\psi_-(A=0)\rangle &= |\varphi_1^1\rangle \\ |\psi_+(A=0)\rangle &= -|\varphi_2^1\rangle \end{aligned} \tag{36}$$

In this case, the energies therefore vary linearly with $\mathscr{E}$ (dashed straight lines in Figure 9). Physically, results (35) and (36) are easy to understand: when the electric field alone acts on the molecule, it "pulls" the positively charged hydrogen atoms above or below the nitrogen atom; this is why the stationary states are $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$.

When the electric field $\mathscr{E}$ and the coupling constant A are both arbitrary, the states $|\psi_+\rangle$ and $|\psi_-\rangle$ are linear superpositions of the states $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$ (and of the states $|\varphi_s^1\rangle$ and $|\varphi_a^1\rangle$ as well), and result from a compromise between the action of the electric field, which tends to pull the hydrogen atoms to one side of the nitrogen atom, and that of the coupling W, which tends to draw the nitrogen atom accross the potential barrier. The variation of the energies E_+ and E_- is shown graphically in Figure 9, in which we see the phenomenon of anti-crossing (*cf.* Chap. IV, § C-2-b) due to the coupling W. E_+ and E_- correspond to the two branches of a hyperbola whose asymptotes are the dashed lines associated with the energies in the absence of coupling. Finally, we can calculate

[5] If $\mathscr{E}$ is negative, the roles of $|\varphi_1^1\rangle$ and $|\varphi_2^1\rangle$ are inverted in (36).

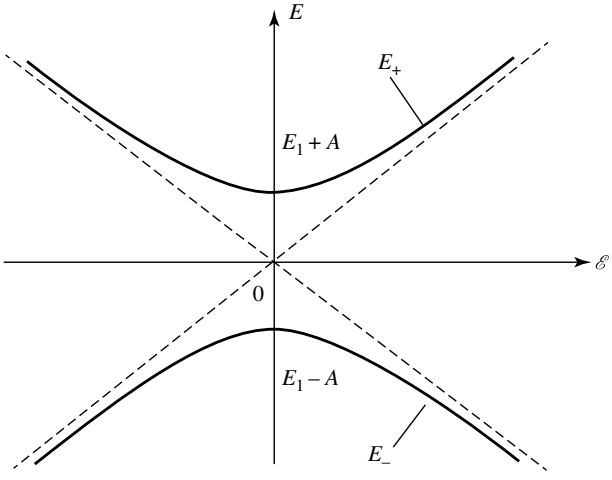


Figure 9: Influence of an electric field $\mathscr{E}$ on the first two levels of the ammonia molecule (their spacing $2A$ in a zero field is due to the tunnel effect coupling). For weak $\mathscr{E}$, the molecule acquires a dipole moment proportional to $\mathscr{E}$, and the corresponding energy varies with $\mathscr{E}^2$. For large $\mathscr{E}$, the dipole moment approaches a limit (corresponding to the nitrogen atom either above or below the plane of the hydrogen atoms), and the energy becomes a linear function of $\mathscr{E}$.

the mean values of the electric dipole moment D in each of the two stationary states $|\psi_+\rangle$ and $|\psi_-\rangle$. Using (26) and (31), we find:

$$\langle\psi_+|D|\psi_+\rangle = -\langle\psi_-|D|\psi_-\rangle = \eta\cos\theta \tag{37}$$

which, according to (32), yields:

$$\langle\psi_+|D|\psi_+\rangle = -\langle\psi_-|D|\psi_-\rangle = -\frac{\eta^2\mathscr{E}}{\sqrt{A^2+\eta^2\mathscr{E}^2}} \tag{38}$$

For $\mathscr{E}=0$, these two mean values are zero. This corresponds to the fact that, in the two states $|\varphi^1_{s,a}\rangle$, the particle has an equal probability of being in one or the other of the two wells. On the other hand, when $\eta\mathscr{E} \gg A$, we again find the dipole moment $+\eta$ (or $-\eta$) corresponding to the state $|\varphi^1_1\rangle$ (or $|\varphi^1_2\rangle$).

When the electric field is weak ($\eta\mathscr{E} \ll A$), formulas (38) can be written in the form:

$$\langle\psi_+|D|\psi_+\rangle = -\langle\psi_-|D|\psi_-\rangle = -\frac{\eta^2}{A}\mathscr{E} \tag{39}$$

We see that the molecule in the stationary state $|\psi_+\rangle$ (or $|\psi_-\rangle$) acquires an electric dipole moment proportional to the external field $\mathscr{E}$. If we define an electric susceptibility of the molecule in the state $|\psi_-\rangle$ by the relation:

$$\langle\psi_-|D|\psi_-\rangle = \varepsilon_-\mathscr{E} \tag{40}$$

we find, according to (39), that:

$$\varepsilon_- = \frac{\eta^2}{A} \tag{41}$$

(the same calculations are valid for $|\psi_+\rangle$ and yield $\varepsilon_+ = -\varepsilon_-$).

Comment:

In a weak field, formulas (30) can be expanded as a power series of $\eta\mathscr{E}/A$:

$$E_- = E_1 - A - \frac{1}{2}\frac{\eta^2\mathscr{E}^2}{A} + \ldots \tag{42a}$$

$$E_+ = E_1 + A + \frac{1}{2}\frac{\eta^2\mathscr{E}^2}{A} + \ldots \tag{42b}$$

Let us now consider ammonia molecules moving in a region where $\mathscr{E}$ is weak but where $\mathscr{E}^2$ has a strong gradient in the Ox direction (i.e. along the axis of the molecules):

$$\frac{\mathrm{d}}{\mathrm{d}x}(\mathscr{E}^2) = \lambda \tag{43}$$

According to (42a), the molecules in the state $|\psi_-\rangle$ are subjected to a force parallel to Ox which is equal to:

$$F_- = -\frac{\mathrm{d}E_-}{\mathrm{d}x} = \frac{1}{2}\lambda\frac{\eta^2}{A} \tag{44}$$

Relation (42b) indicates that the molecules in the state $|\psi_+\rangle$ are subjected to an opposite force:

$$F_+ = -\frac{\mathrm{d}E_+}{\mathrm{d}x} = -F_- \tag{45}$$

This result is the basis of the method used in the ammonia maser to sort the molecules and select those in the higher energy state. The device is analogous to the Stern-Gerlach apparatus: a beam of ammonia molecules crosses a region where there is a strong electric field gradient, the molecules follow different trajectories depending on whether they are in one state or the other; one can, using a suitable diaphragm, isolate either one of the two states.

References and suggestions for further reading:

Feynman III (1.2), § 8-6 and Chap. 9; Alonso and Finn III (1.4), § 2-8; article by Vuylsteke (1.34); Townes and Schawlow (12.10), Chap. 12; see (15.11) for references to original articles on masers; articles by Lyons (15.14), Gordon (15.15), and Turner (12.14).

See also Encrenaz (12.11), Chap. VI.

Complement H_{IV}

Effects of a coupling between a stable state and an unstable state

1 Introduction. Notation 485
2 Influence of a weak coupling on states of different energies 486
3 Influence of an arbitrary coupling on states of the same energy 487

1. Introduction. Notation

The effects of a coupling W between two states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ of energies E_1 and E_2 were discussed in detail in § C of Chapter IV. What modifications appear when one of the two states ($|\varphi_1\rangle$, for example) is unstable?

The concepts of an unstable state and a lifetime were introduced in Complement K_{III}. We shall assume, for example, that $|\varphi_1\rangle$ is an excited atomic state. When the atom is in this state, it can fall back to a lower energy level through spontaneous emission of one or several photons, with a probability $1/\tau_1$ per unit time: τ_1 is the lifetime of the unstable state $|\varphi_1\rangle$. On the other hand, we assume that in the absence of the coupling W, the state $|\varphi_2\rangle$ is stable (τ_2 is infinite).

We saw in Complement K_{III} that a simple way of taking the instability of a state into account consists of adding an imaginary term to the corresponding energy. We shall therefore replace the energy E_1 of the state $|\varphi_1\rangle$ by:

$$E_1' = E_1 - i\frac{\hbar}{2}\gamma_1 \tag{1}$$

with:

$$\gamma_1 = \frac{1}{\tau_1} \tag{2}$$

(since τ_2 is infinite, γ_2 is zero and $E_2' = E_2$). In the absence of coupling, the matrix representing the "Hamiltonian" H_0 of the system is now written in the $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ basis[1]:

$$H_0 = \begin{pmatrix} E_1' & 0 \\ 0 & E_2' \end{pmatrix} = \begin{pmatrix} E_1 - i\frac{\hbar}{2}\gamma_1 & 0 \\ 0 & E_2 \end{pmatrix} \tag{3}$$

[1] The operator H_0 is not Hermitian and is therefore not really a Hamiltonian (see the comment at the end of Complement K_{III}).

2. Influence of a weak coupling on states of different energies

Let us assume, as in § C of Chapter IV, that we add to H_0 a perturbation W, whose matrix in the $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ basis is:

$$W = \begin{pmatrix} 0 & W_{12} \\ W_{21} & 0 \end{pmatrix} \tag{4}$$

What now happens to the energies and lifetimes of the levels?

Let us calculate the eigenvalues ε'_1 and ε'_2 of the matrix:

$$H = H_0 + W = \begin{pmatrix} E_1 - i\frac{\hbar}{2}\gamma_1 & W_{12} \\ W_{21} & E_2 \end{pmatrix} \tag{5}$$

ε'_1 and ε'_2 are the solutions of the equation in ε:

$$\varepsilon^2 - \varepsilon\left(E_1 + E_2 - i\frac{\hbar}{2}\gamma_1\right) + E_1E_2 - i\frac{\hbar}{2}\gamma_1 E_2 - |W_{12}|^2 = 0 \tag{6}$$

To simplify the calculation, we shall confine ourselves to the case where the coupling is weak, i.e. $|W_{12}| \ll \sqrt{(E_1+E_2)^2 + \frac{\hbar^2}{4}\gamma_1^2}$; we then find:

$$\begin{aligned} \varepsilon'_1 &\simeq E_1 - i\frac{\hbar}{2}\gamma_1 + \frac{|W_{12}|^2}{E_2 - E_1 - i\frac{\hbar}{2}\gamma_1} \\ \varepsilon'_2 &\simeq E_2 + \frac{|W_{12}|^2}{E_2 - E_1 + i\frac{\hbar}{2}\gamma_1} \end{aligned} \tag{7}$$

The energies of the eigenstates in the presence of the coupling are the real parts of ε'_1 and ε'_2; the lifetimes are inversely proportional to their imaginary parts. We see from (7) that the coupling changes, to second order in $|W_{12}|$, both the energies and the lifetimes. In particular, we observe that ε'_1 and ε'_2 are both complex when $|W_{12}|$ is not zero: in the presence of the coupling, there is no longer any stable state. We can write ε'_2 in the form:

$$\varepsilon'_2 = \Delta_2 - i\frac{\hbar}{2}\Gamma_2 \tag{8}$$

with:

$$\Delta_2 = E_2 + \frac{(E_2 - E_1)|W_{12}|^2}{(E_2 - E_1)^2 + \frac{\hbar^2}{4}\gamma_1^2} \tag{9a}$$

$$\Gamma_2 = \gamma_1 \frac{|W_{12}|^2}{(E_2 - E_1) + \frac{\hbar^2}{4}\gamma_1^2} \tag{9b}$$

The state $|\varphi_2\rangle$ therefore acquires, under the effect of the coupling, a finite lifetime whose inverse is given in (9b) (Bethe's formula). This result is easy to understand physically:

if the system at $t = 0$ is in the stable state $|\varphi_2\rangle$, there is a non-zero probability at a subsequent time t of finding it in the state $|\varphi_1\rangle$, in which the system has a finite lifetime. It is sometimes said figuratively that "the coupling brings into the stable state part of the instability of the other state". Moreover, it can be seen from expressions (7) that, as in the case studied in § C of Chapter IV, the smaller the difference between the unperturbed energies E_1 and E_2, the more effectively the perturbation acts on the energies and lifetimes. We shall therefore study in the next section the case where this difference is zero.

3. Influence of an arbitrary coupling on states of the same energy

When the energies E_1 and E_2 are equal, the operator H is written, if we make its trace appear explicitly, as in § 2 of Complement B_{IV}:

$$H = \left(E_1 - i\frac{\hbar}{4}\gamma_1\right)\mathbb{1} + K \tag{10}$$

where $\mathbb{1}$ is the identity operator and K is the operator which, in the $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ basis, has for its matrix:

$$(K) = \begin{pmatrix} -i\frac{\hbar}{4}\gamma_1 & W_{12} \\ (W_{12})^* & i\frac{\hbar}{4}\gamma_1 \end{pmatrix} \tag{11}$$

The eigenvalues k_1 and k_2 of K are the two solutions of the characteristic equation:

$$k^2 = |W_{12}|^2 - \frac{\hbar^2}{16}\gamma_1^2 \tag{12}$$

They therefore have opposite values:

$$k_1 = -k_2 \tag{13}$$

which yields for the eigenvalues of H:

$$\begin{aligned} \varepsilon_1' &= E_1 - i\frac{\hbar}{4}\gamma_1 + k_1 \\ \varepsilon_2' &= E_1 - i\frac{\hbar}{4}\gamma_1 - k_1 \end{aligned} \tag{14}$$

The eigenvectors of H and K are the same; a simple calculation[2] enables us to obtain these vectors $|\psi_1'\rangle$ and $|\psi_2'\rangle$:

$$\begin{aligned} |\psi_1'\rangle &= W_{12}|\varphi_1\rangle + \left(k_1 + i\frac{\hbar}{4}\gamma_1\right)|\varphi_2\rangle \\ |\psi_2'\rangle &= W_{12}|\varphi_1\rangle + \left(-k_1 + i\frac{\hbar}{4}\gamma_1\right)|\varphi_2\rangle \end{aligned} \tag{15}$$

[2]For the calculation performed here, it is not indispensable to normalize $|\psi_1'\rangle$ and $|\psi_2'\rangle$. Note also that, since H is not Hermitian, $|\psi_1'\rangle$ and $|\psi_2'\rangle$ are not orthogonal.

Assume that the system at time $t = 0$ is in the state $|\varphi_2\rangle$ (which would be stable in the absence of the coupling):

$$|\psi(t=0)\rangle = |\varphi_2\rangle = \frac{1}{2k_1}\,[|\psi_1'\rangle - |\psi_2'\rangle] \tag{16}$$

Using (14), we see that, at time t, the state vector is:

$$|\psi(t)\rangle = \frac{1}{2k_1}\,\mathrm{e}^{-iE_1t/\hbar}\,\mathrm{e}^{-\frac{1}{4}\gamma_1}\,[\mathrm{e}^{-ik_1t/\hbar}|\psi_1'\rangle - \mathrm{e}^{ik_1t/\hbar}|\psi_2'\rangle] \tag{17}$$

The probability $\mathscr{P}_{21}(t)$ of finding the system at time t in the state $|\varphi_1\rangle$ is:

$$\begin{aligned}\mathscr{P}_{21}(t) &= |\langle\varphi_1|\psi(t)\rangle|^2 \\ &= \frac{1}{4|k_1|^2}\mathrm{e}^{-\gamma_1 t/2}\,\left|\mathrm{e}^{-ik_1t/\hbar}\langle\varphi_1|\psi_1'\rangle - \mathrm{e}^{ik_1t/\hbar}\langle\varphi_1|\psi_2'\rangle\right|^2 \\ &= \frac{1}{4|k_1|^2}\mathrm{e}^{-\gamma_1 t/2}\,|W_{12}|^2\,\left|\mathrm{e}^{-ik_1t/\hbar} - \mathrm{e}^{ik_1t/\hbar}\right|^2\end{aligned} \tag{18}$$

We shall distinguish between several cases:

(i) When the condition:

$$|W_{12}| > \frac{\hbar}{4}\gamma_1 \tag{19}$$

is satisfied, we obtain directly, using (12):

$$k_1 = -k_2 = \sqrt{|W_{12}|^2 - \left(\frac{\hbar}{4}\gamma_1\right)^2} \tag{20}$$

and the eigenvalues ε_1' and ε_2' are given by:

$$\begin{aligned}\varepsilon_1' &= E_1 + \sqrt{|W_{12}|^2 - \left(\frac{\hbar}{4}\gamma_1\right)^2} - i\frac{\hbar}{4}\gamma_1 \\ \varepsilon_2' &= E_2 - \sqrt{|W_{12}|^2 - \left(\frac{\hbar}{4}\gamma_1\right)^2} - i\frac{\hbar}{4}\gamma_1\end{aligned} \tag{21}$$

ε_1' and ε_2' have the same imaginary part, but different real parts. The states $|\psi_1'\rangle$ and $|\psi_2'\rangle$ therefore have the same lifetime, $2\tau_1$, but different energies.

Substituting (20) into (18), we obtain:

$$\mathscr{P}_{21}(t) = \frac{|W_{12}|^2}{|W_{12}|^2 - \left(\frac{\hbar}{4}\gamma_1\right)^2}\,\mathrm{e}^{-\gamma_1 t/2}\sin^2\left(\sqrt{|W_{12}|^2 - \left(\frac{\hbar}{4}\gamma_1\right)^2}\;\frac{t}{\hbar}\right) \tag{22}$$

The form of this result recalls Rabi's formula [*cf.* Chap. IV, equation (C-32)]. The function $\mathscr{P}_{21}(t)$ is represented by a damped sinusoid with time constant $2\tau_1$ (Fig. 1). Condition (19) thus expresses the fact that the coupling is sufficiently strong to make

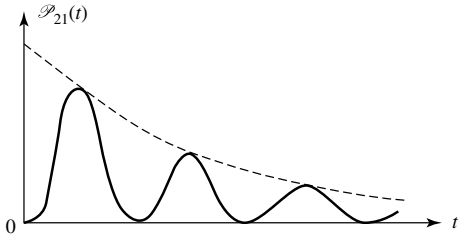


Figure 1: Effect of a strong coupling between a stable state $|\varphi_2\rangle$ and an unstable state $|\varphi_1\rangle$. If the system is initially in the state $|\varphi_2\rangle$, the probability $\mathscr{P}_{21}(t)$ of finding it in the state $|\varphi_1\rangle$ at time t presents damped oscillations.

the system oscillate between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ before the instability of the state $|\varphi_1\rangle$ can have a real effect.

(*ii*) If, on the other hand, the condition:

$$|W_{12}| < \frac{\hbar}{4}\gamma_1 \tag{23}$$

is satisfied, we then have:

$$k_1 = -k_2 = i\sqrt{\left(\frac{\hbar}{4}\gamma_1\right)^2 - |W_{12}|^2} \tag{24}$$

and:

$$\begin{aligned}
\varepsilon_1' &= E_1 - i\left[\frac{\hbar}{4}\gamma_1 - \sqrt{\left(\frac{\hbar}{4}\gamma_1\right)^2 - |W_{12}|^2}\right] \\
\varepsilon_2' &= E_1 - i\left[\frac{\hbar}{4}\gamma_1 + \sqrt{\left(\frac{\hbar}{4}\gamma_1\right)^2 - |W_{12}|^2}\right]
\end{aligned} \tag{25}$$

The states $|\psi_1'\rangle$ and $|\psi_2'\rangle$ then have the same energy and different lifetimes. Formula (18) becomes:

$$\mathscr{P}_{21}(t) = \frac{|W_{12}|^2}{\left(\frac{\hbar}{4}\gamma_1\right)^2 - |W_{12}|^2}\ \mathrm{e}^{-\gamma_1 t/2}\ \sinh^2\left(\sqrt{\left(\frac{\hbar}{4}\gamma_1\right)^2 - |W_{12}|^2}\ \frac{t}{\hbar}\right) \tag{26}$$

This time, $\mathscr{P}_{21}(t)$ is a sum of damped exponentials (Fig. 2).

This result has a simple physical interpretation: condition (23) expresses the fact that the lifetime τ_1 is so short that the system is completely damped before the coupling W has had the time to make it oscillate between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$.

(*iii*) Finally, let us examine the case where we have exactly:

$$|W_{12}| = \frac{\hbar}{4}\gamma_1 \tag{27}$$

We see then from (14) that the states $|\psi_1'\rangle$ and $|\psi_2'\rangle$ both have the same energy E_1 and the same lifetime $2\tau_1$.

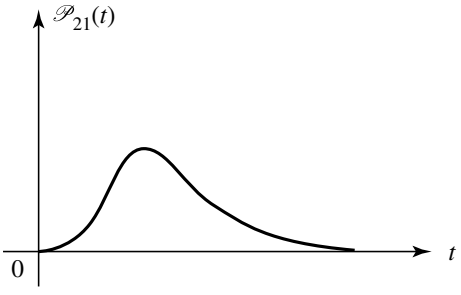


Figure 2: When the coupling is weak, oscillations between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ do not have time to occur.

Equations (22) and (26), in this case, take on indeterminate forms, which can be resolved and both yield:

$$\mathscr{P}_{21}(t) = \frac{|W_{12}|^2}{\hbar^2} t^2 \mathrm{e}^{-\gamma_1 t/2} \tag{28}$$

Comment:

The preceding discussion is very similar to that of the classical motion of a damped harmonic oscillator. Conditions (19), (23) and (27) correspond respectively to weak, strong and critical damping.

References and suggestions for further reading:

An important application of the phenomenon discussed in this complement is the shortening of the lifetime of a metastable state due to an electric field. See: Lamb and Retherford (3.11), App. II; Sobel'man (11.12), Chap. 8, § 28-5.

Complement J_{IV}

Exercises

1. Consider a spin 1/2 particle of magnetic moment $\mathbf{M} = \gamma\mathbf{S}$. The spin state space is spanned by the basis of the $|+\rangle$ and $|-\rangle$ vectors, eigenvectors of S_z with eigenvalues $+\hbar/2$ and $-\hbar/2$. At time $t = 0$, the state of the system is:

$$|\psi(t=0)\rangle = |+\rangle$$

a. If the observable S_x is measured at time $t = 0$, what results can be found, and with what probabilities?

b. Instead of performing the preceding measurement, we let the system evolve under the influence of a magnetic field parallel to Oy, of modulus B_0. Calculate, in the $\{|+\rangle, |-\rangle\}$ basis, the state of the system at time t.

c. At this time t, we measure the observables S_x, S_y, S_z. What values can we find, and with what probabilities? What relation must exist between B_0 and t for the result of one of the measurements to be certain? Give a physical interpretation of this condition.

2. Consider a spin 1/2 particle, as in the previous exercise (using the same notation).

a. At time $t = 0$, we measure S_y and find $+\hbar/2$. What is the state vector $|\psi(0)\rangle$ immediately after the measurement?

b. Immediately after this measurement, we apply a uniform time-dependent field parallel to Oz. The Hamiltonian operator of the spin $H(t)$ is then written:

$$H(t) = \omega_0(t)\ S_z$$

Assume that $\omega_0(t)$ is zero for $t < 0$ and $t > T$ and increases linearly from 0 to ω_0 when $0 \leq t \leq T$ (T is a given parameter having the dimension of a time). Show that at time t the state vector can be written:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}[\mathrm{e}^{i\theta(t)}|+\rangle + i\mathrm{e}^{-i\theta(t)}|-\rangle]$$

where $\theta(t)$ is a real function of t (to be calculated by the student).

c. At a time $t = \tau > T$, we measure S_y. What results can we find, and with what probabilities? Determine the relation that must exist between ω_0 and T in order for us to be sure of the result. Give the physical interpretation.

3. Consider a spin 1/2 particle placed in a magnetic field $\mathbf{B}_0$ with components:

$$\begin{cases} B_x = \dfrac{1}{\sqrt{2}} B_0 \\ B_y = 0 \\ B_z = \dfrac{1}{\sqrt{2}} B_0 \end{cases}$$

The notation is the same as that of exercise (1).

a. Calculate the matrix representing, in the $\{|+\rangle, |-\rangle\}$ basis, the operator H, the Hamiltonian of the system.

b. Calculate the eigenvalues and the eigenvectors of H.

c. The system at time $t = 0$ is in the state $|-\rangle$. What values can be found if the energy is measured, and with what probabilities?

d. Calculate the state vector $|\psi(t)\rangle$ at time t. At this instant, S_x is measured; what is the mean value of the results that can be obtained? Give a geometrical interpretation.

4. Consider the experimental device described in § B-2-b of Chapter IV (*cf.* Fig. 8): a beam of atoms of spin 1/2 passes through one apparatus, which serves as a "polarizer" in a direction which makes an angle θ with Oz in the xOz plane, and then through another apparatus, the "analyzer", which measures the S_z component of the spin. We assume in this exercise that between the polarizer and the analyzer, over a length L of the atomic beam, a magnetic field $\mathbf{B}_0$ is applied which is uniform and parallel to Ox. We call v the speed of the atoms and $T = L/v$ the time during which they are submitted to the field $\mathbf{B}_0$. We set $\omega_0 = -\gamma B_0$.

a. What is the state vector $|\psi_1\rangle$ of a spin at the moment it enters the analyzer?

b. Show that when the measurement is performed in the analyzer, there is a probability equal to $\frac{1}{2}(1 + \cos\theta \cos\omega_0 T)$ of finding $+\hbar/2$ and $\frac{1}{2}(1 - \cos\theta \cos\omega_0 T)$ of finding $-\hbar/2$. Give a physical interpretation.

c. (This question and the following one involve the concept of a density operator, defined in Complement E_{III}. The reader is also advised to refer to Complement E_{IV}). Show that the density matrix ρ_1 of a particle which enters the analyzer is written, in the $\{|+\rangle, |-\rangle\}$ basis:

$$\rho_1 = \frac{1}{2} \begin{pmatrix} 1 + \cos\theta \ \cos\omega_0 T & \sin\theta + i\cos\theta \ \sin\omega_0 T \\ \sin\theta - i\cos\theta \ \sin\omega_0 T & 1 - \cos\theta \ \cos\omega_0 T \end{pmatrix}$$

Calculate $\mathrm{Tr}\{\rho_1 S_x\}$, $\mathrm{Tr}\{\rho_1 S_y\}$ and $\mathrm{Tr}\{\rho_1 S_z\}$. Give an interpretation. Does the density operator ρ_1 describe a pure state?

d. Now assume that the speed of an atom is a random variable, and hence the time T is known only to within a certain uncertainty ΔT. In addition, the field B_0 is assumed to be sufficiently strong that $\omega_0 \Delta T \gg 1$. The possible values of the product $\omega_0 T$ are then (modulus 2π) all values included between 0 and 2π, all of which are equally probable.

In this case, what is the density operator ρ_2 of an atom at the moment it enters the analyzer? Does ρ_2 correspond to a pure case? Calculate the quantities $\mathrm{Tr}\{\rho_2 S_x\}$, $\mathrm{Tr}\{\rho_2 S_y\}$ and $\mathrm{Tr}\{\rho_2 S_z\}$. What is your interpretation? In which case does the density operator describe a completely polarized spin? A completely unpolarized spin?

Describe qualitatively the phenomena observed at the analyzer exit when ω_0 varies from zero to a value where the condition $\omega_0 \Delta T \gg 1$ is satisfied.

5. Evolution operator of a spin 1/2 (*cf.* Complement $\mathrm{F_{III}}$)

Consider a spin 1/2, of magnetic moment $\mathbf{M} = \gamma \mathbf{S}$, placed in a magnetic field $\mathbf{B}_0$ of components $B_x = -\omega_x/\gamma$, $B_y = -\omega_y/\gamma$, $B_z = -\omega_z/\gamma$.

We set:

$$\omega_0 = -\gamma |\mathbf{B}_0|$$

a. Show that the evolution operator of this spin is:

$$U(t,0) = \mathrm{e}^{iMt}$$

where M is the operator:

$$M = \frac{1}{\hbar}[\omega_x S_x + \omega_y S_y + \omega_z S_z] = \frac{1}{2}[\omega_x \sigma_x + \omega_y \sigma_y + \omega_z \sigma_z]$$

expressed as a function of the three Pauli matrices σ_x, σ_y and σ_z (*cf.* Complement $\mathrm{A_{IV}}$).

Calculate the matrix representing M in the $\{|+\rangle, |-\rangle\}$ basis of eigenvectors of S_z. Show that:

$$M^2 = \frac{1}{4}[\omega_x^2 + \omega_y^2 + \omega_z^2] = \left(\frac{\omega_0}{2}\right)^2$$

b. Put the evolution operator into the form:

$$U(t,0) = \cos\left(\frac{\omega_0 t}{2}\right) - \frac{2i}{\omega_0} M \sin\left(\frac{\omega_0 t}{2}\right)$$

c. Consider a spin which at time $t = 0$ is in the state $|\psi(0)\rangle = |+\rangle$. Show that the probability $\mathscr{P}_{++}(t)$ of finding it in the state $|+\rangle$ at time t is:

$$\mathscr{P}_{++}(t) = |\langle + | U(t,0) | + \rangle|^2$$

and derive the relation:

$$\mathscr{P}_{++}(t) = 1 - \frac{\omega_x^2 + \omega_y^2}{\omega_0^2} \sin^2\left(\frac{\omega_0 t}{2}\right)$$

Give a geometrical interpretation.

6. Consider the system composed of two spin 1/2's, $\mathbf{S}_1$ and $\mathbf{S}_2$, and the basis of four vectors $|\pm,\pm\rangle$ defined in Complement D_{IV}. The system at time $t = 0$ is in the state:

$$|\psi(0)\rangle = \frac{1}{2}|+\ +\rangle + \frac{1}{2}|+\ -\rangle + \frac{1}{\sqrt{2}}|-\ -\rangle$$

a. At time $t = 0$, $\mathbf{S}_{1z}$ is measured; what is the probability of finding $-\hbar/2$? What is the state vector after this measurement? If we then measure S_{1x}, what results can be found, and with what probabilities? Answer the same questions for the case where the measurement of S_{1z} yielded $+\hbar/2$.

b. When the system is in the state $|\psi(0)\rangle$ written above, S_{1z} and S_{2z} are measured simultaneously. What is the probability of finding opposite results? Identical results?

c. Instead of performing the preceding measurements, we let the system evolve under the influence of the Hamiltonian:

$$H = \omega_1 S_{1z} + \omega_2 S_{2z}$$

What is the state vector $|\psi(t)\rangle$ at time t? Calculate at time t the mean values $\langle \mathbf{S}_1 \rangle$ and $\langle \mathbf{S}_2 \rangle$. Give a physical interpretation.

d. Show that the lengths of the vectors $\langle \mathbf{S}_1 \rangle$ and $\langle \mathbf{S}_2 \rangle$ are less than $\hbar/2$. What must be the form of $|\psi(0)\rangle$ for each of these lengths to be equal to $+\hbar/2$?

7. Consider the same system of two spin 1/2's as in the preceding exercise; the state space is spanned by the basis of four states $|\pm,\pm\rangle$.

a. Write the 4×4 matrix representing, in this basis, the S_{1y} operator. What are the eigenvalues and eigenvectors of this matrix?

b. The normalized state of the system is:

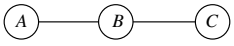
$$|\psi\rangle = \alpha|+\ +\rangle + \beta|+\ -\rangle + \gamma|-\ +\rangle + \delta|-\ -\rangle$$

where α, β, γ and δ are given complex coefficients. S_{1x} and S_{2y} are measured simultaneously; what results can be found, and with what probabilities? What happens to these probabilities if $|\psi\rangle$ is a tensor product of a vector of the state space of the first spin and a vector of the state space of the second spin?

c. Same questions for a measurement of S_{1y} and S_{2y}.

d. Instead of performing the preceding measurements, we measure only S_{2y}. Calculate, first from the results of b and then from those of c, the probability of finding $-\hbar/2$.

8. Consider an electron of a linear triatomic molecule formed by three equidistant atoms.



We use $|\varphi_A\rangle$, $|\varphi_B\rangle$, $|\varphi_C\rangle$ to denote three orthonormal states of this electron, corresponding respectively to three wave functions localized about the nuclei of atoms A, B, C. We shall confine ourselves to the subspace of the state space spanned by $|\varphi_A\rangle$, $|\varphi_B\rangle$ and $|\varphi_C\rangle$.

When we neglect the possibility of the electron jumping from one nucleus to another, its energy is described by the Hamiltonian H_0 whose eigenstates are the three states $|\varphi_A\rangle$, $|\varphi_B\rangle$, $|\varphi_C\rangle$ with the same eigenvalue E_0. The coupling between the states $|\varphi_A\rangle$, $|\varphi_B\rangle$, $|\varphi_C\rangle$ is described by an additional Hamiltonian W defined by:

$$W|\varphi_A\rangle = -a|\varphi_B\rangle$$
$$W|\varphi_B\rangle = -a|\varphi_A\rangle - a|\varphi_C\rangle$$
$$W|\varphi_C\rangle = -a|\varphi_B\rangle$$

where a is a real positive constant.

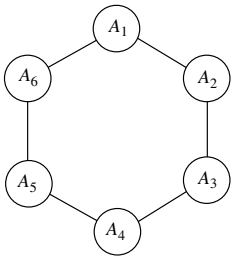
a. Calculate the energies and stationary states of the Hamiltonian $H = H_0 + W$.

b. The electron at time $t = 0$ is in the state $|\varphi_A\rangle$. Discuss qualitatively the localization of the electron at subsequent times t. Are there any values of t for which it is perfectly localized about atom A, B or C?

c. Let D be the observable whose eigenstates are $|\varphi_A\rangle$, $|\varphi_B\rangle$, $|\varphi_C\rangle$ with respective eigenvalues $-d$, 0, d. D is measured at time t; what values can be found, and with what probabilities?

d. When the initial state of the electron is arbitrary, what are the Bohr frequencies that can appear in the evolution of $\langle D\rangle$? Give a physical interpretation of D. What are the frequencies of the electromagnetic waves that can be absorbed or emitted by the molecule?

9. A molecule is composed of six identical atoms A_1, A_2,... A_6 which form a regular hexagon. Consider an electron which can be localized on each of the atoms. Call $|\varphi_n\rangle$ the state in which it is localized on the nth atom ($n = 1, 2, ..., 6$). The electron states will be confined to the space spanned by the $|\varphi_n\rangle$, assumed to be orthonormal.



a. Define an operator R by the following relations:

$$R|\varphi_1\rangle = |\varphi_2\rangle \ ; \quad R|\varphi_2\rangle = |\varphi_3\rangle \ ; \ldots \ ; \quad R|\varphi_6\rangle = |\varphi_1\rangle$$

Find the eigenvalues and eigenstates of R. Show that the eigenvectors of R form a basis of the state space.

b. When the possibility of the electron passing from one site to another is neglected, its energy is described by a Hamiltonian H_0 whose eigenstates are the six states $|\varphi_n\rangle$, with the same eigenvalue E_0. As in the previous exercise, we describe the possibility of the electron jumping from one atom to another by adding a perturbation W to the Hamiltonian H_0; W is defined by:

$$W|\varphi_1\rangle = -a|\varphi_6\rangle - a|\varphi_2\rangle \ ; \qquad W|\varphi_2\rangle = -a|\varphi_1\rangle - a|\varphi_3\rangle \ ;$$
$$\ldots\ldots \ ; \ W|\varphi_6\rangle = -a|\varphi_5\rangle - a|\varphi_1\rangle$$

Show that R commutes with the total Hamiltonian $H = H_0 + W$. From this deduce the eigenstates and eigenvalues of H. In these eigenstates, is the electron localized? Apply these considerations to the benzene molecule.

Exercise 9.
Reference: Feynman III (1.2), § 15-4.

Chapter V

The one-dimensional harmonic oscillator

A Introduction 497
A-1 Importance of the harmonic oscillator in physics 497
A-2 The harmonic oscillator in classical mechanics 499
A-3 General properties of the quantum mechanical Hamiltonian . 502
B Eigenvalues of the Hamiltonian 503
B-1 Notation 503
B-2 Determination of the spectrum 506
B-3 Degeneracy of the eigenvalues 509
C Eigenstates of the Hamiltonian 510
C-1 The $\{|\varphi_n\rangle\}$ representation 510
C-2 Wave functions associated with the stationary states 515
D Discussion 518
D-1 Mean values and root mean square deviations of X and P in a state $|\varphi_n\rangle$ 518
D-2 Properties of the ground state 520
D-3 Time evolution of the mean values 522

A. Introduction

A-1. Importance of the harmonic oscillator in physics

This chapter is devoted to the study of a particularly important physical system: the one-dimensional harmonic oscillator.

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

The simplest example of such a system is that of a particle of mass m moving in a potential which depends only on x and has the form:

$$V(x) = \frac{1}{2}kx^2 \tag{A-1}$$

(k is a real positive constant). The particle is attracted towards the $x = 0$ plane [the minimum of $V(x)$, corresponding to positions of stable equilibrium] by a restoring force:

$$F_x = -\frac{\mathrm{d}V}{\mathrm{d}x} = -kx \tag{A-2}$$

which is proportional to the distance x between the particle and the $x = 0$ plane (x is an algebraic variable: $x \lessgtr 0$). We know that in classical mechanics, the projection onto Ox of the particle's motion is a sinusoidal oscillation about $x = 0$, of angular frequency:

$$\omega = \sqrt{\frac{k}{m}} \tag{A-3}$$

Actually, a large number of systems are governed (at least approximately) by the harmonic oscillator equations. Whenever one studies the behavior of a physical system in the neighborhood of a stable equilibrium position, one arrives at equations which, in the limit of small oscillations, are those of a harmonic oscillator (see § A-2). The results we shall derive in this chapter are applicable, therefore, to a whole series of important physical phenomena – for example, the vibrations of the atoms of a molecule about their equilibrium position, the oscillations of atoms or ions of a crystalline lattice (phonons)[1].

The harmonic oscillator is also involved in the study of the electromagnetic field. We know that in a cavity, there exist an infinite number of possible stationary waves (normal modes of the cavity). The electromagnetic field can be expanded in terms of these modes and it can be shown, using Maxwell's equations, that each of the coefficients of this expansion (which describe the state of the field at each instant) obeys a differential equation, which is identical to that of a harmonic oscillator whose angular frequency ω is that of the associated normal mode. In other words, the electromagnetic field is formally equivalent to a set of independent harmonic oscillators (*cf.* Complement $\mathrm{K_V}$). The quantization of the field is obtained by quantizing these oscillators associated with the various normal modes of the cavity (*cf.* Chapter XIX). Recall, moreover, that it was the study of the behavior of these oscillators at thermal equilibrium (blackbody radiation) which, historically, led Planck to introduce, for the first time in physics, the constant h which bears his name. We shall see (*cf.* Complement $\mathrm{L_V}$) that the mean energy of a harmonic oscillator in thermodynamic equilibrium at the temperature T is different for classical and quantum mechanical oscillators.

The harmonic oscillator also plays an important role in the description of a set of identical particles which are all in the same quantum mechanical state (they must obviously be bosons, *cf.* Chap. XIV). As we shall see later, this is because the energy levels of a harmonic oscillator are equidistant, the spacing between two adjacent levels being equal to $\hbar\omega$. With the energy level labelled by the integer n (situated at a distance $n\hbar\omega$ above the ground state) can then be associated a set of n identical particles (or

[1] Complement $\mathrm{A_V}$ is devoted to a qualitative study of some physical examples of harmonic oscillators.

quanta), each possessing an energy $\hbar\omega$ (*cf.* Chapter XV). The transition of the oscillator from level n to level $n+1$ or $n-1$ corresponds to the creation or annihilation of a quantum of energy $\hbar\omega$. In this chapter, we shall introduce the operators $a^\dagger$ and a, which enable us to describe this transition from level n to level $n+1$ or $n-1$. These operators, respectively called "creation" and "annihilation" operators[2], are used throughout quantum statistical mechanics and quantum field theory[3].

The detailed study of the harmonic oscillator in quantum mechanics is therefore extremely important from a physical point of view. Moreover, we are dealing with a quantum mechanical system for which the Schrödinger equation can be solved rigorously. Having studied spin 1/2 and two-level systems in Chapter IV, we shall therefore now consider another simple example which illustrates the general formalism of quantum mechanics. We shall show in particular how to solve an eigenvalue equation by dealing only with the operators and the commutation relations (this technique will also be applied to angular momentum). We shall also study in a detailed way the motion of wave packets, particularly at the classical limit (*cf.* Complement G_V on quasi-classical states).

In § A-2, we shall review some results related to the classical oscillator before stating (§ A-3) certain general properties of the eigenvalues of the Hamiltonian H. Then, in §§ B and C, we shall determine these eigenvalues and eigenvectors by introducing creation and annihilation operators and using only the consequences of the canonical commutation relation $[X, P] = i\hbar$, as well as the particular form of H. § D is devoted to a physical study of the stationary states of the oscillator and wave packets formed by linear superpositions of these stationary states.

A-2. The harmonic oscillator in classical mechanics

The potential energy $V(x)$ [formula (A-1)] is shown in Figure 1. The motion of the particle is governed by the dynamical equation:

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\frac{\mathrm{d}V}{\mathrm{d}x} = -kx \tag{A-4}$$

The general solution of this equation is of the form:

$$x = x_M \cos(\omega t - \varphi) \tag{A-5}$$

where ω is defined by (A-3), and the constants of integration x_M and φ are determined by the initial conditions of the motion. The particle therefore *oscillates sinusoidally* about the point O, with an amplitude x_M and an angular frequency ω.

The kinetic energy of the particle is:

$$T = \frac{1}{2}m\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 = \frac{p^2}{2m} \tag{A-6}$$

[2] Annihilation operators are also often called "destruction operators".

[3] The aim of quantum field theory is to describe interactions between particles in the relativistic domain, especially the interactions between electrons, positrons and photons. It is clear that creation and annihilation operators should play an important role, since such processes are indeed observed experimentally (absorption or emission of photons, pair creation...). The quantum theory of electromagnetism is introduced in Chapter XIX.

where $p = m\dfrac{\mathrm{d}x}{\mathrm{d}t}$ is the momentum of the particle. The total energy is:

$$E = T + V = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{A-7}$$

Substituting solution (A-5) into this equation, we find:

$$E = \frac{1}{2}m\omega^2 x_M^2 \tag{A-8}$$

The energy of the particle is therefore time-independent (this is a general property of conservative systems) and can take on any positive (or zero) value, since x_M is *a priori* arbitrary.

If we fix the total energy E, the limits of the classical motion $x = \pm x_M$ can be determined from Figure 1 by taking the intersection of the parabola with the line parallel to Ox of ordinate E. At these points $x = \pm x_M$, the potential energy is at a maximum and equal to E, and the kinetic energy is zero. On the other hand, at $x = 0$, the potential energy is zero and the kinetic energy is maximum.

Comment:

Consider an arbitrary potential $V(x)$ which has a minimum at $x = x_0$ (Fig. 2). Expanding the function $V(x)$ in a Taylor's series in the neighborhood of x_0, we obtain:

$$V(x) = a + b(x - x_0)^2 + c(x - x_0)^3 + ... \tag{A-9}$$

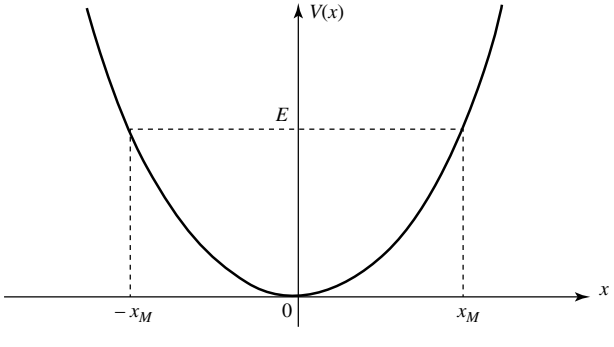


Figure 1: The potential energy $V(x)$ of a one-dimensional harmonic oscillator. The amplitude of the classical motion of energy E is x_M.

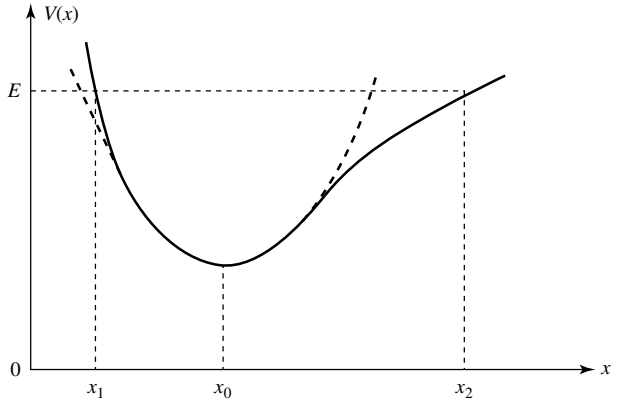


Figure 2: In the neighborhood of a minimum, any potential $V(x)$ can be approximated by a parabolic potential (dashed line). In the potential $V(x)$, a classical particle of energy E oscillates between x_1 and x_2.

The coefficients of this expansion are given by:

$$\begin{aligned} a &= V(x_0) \\ b &= \frac{1}{2}\left(\frac{\mathrm{d}^2 V}{\mathrm{d}x^2}\right)_{x=x_0} \\ c &= \frac{1}{3!}\left(\frac{\mathrm{d}^3 V}{\mathrm{d}x^3}\right)_{x=x_0} \end{aligned} \tag{A-10}$$

and the linear term in $(x - x_0)$ is zero since x_0 corresponds to a minimum of $V(x)$. The force derived from the potential $V(x)$ is, in the neighborhood of x_0:

$$F_x = -\frac{\mathrm{d}V}{\mathrm{d}x} = -2b(x - x_0) - 3c(x - x_0)^2 + ... \tag{A-11}$$

Since $x = x_0$ represents a minimum, the coefficient b is positive.

The point $x = x_0$ corresponds to a stable equilibrium position for the particle: F_x is zero for $x = x_0$; moreover, for $(x - x_0)$ sufficiently small, F_x and $(x - x_0)$ have opposite signs since b is positive.

If the amplitude of the motion of the particle about x_0 is sufficiently small for the term in $(x - x_0)^3$ of (A-9) [and therefore, the corresponding term in $(x - x_0)^2$ of (A-11)] to be negligible compared to the preceding ones, we have a harmonic oscillator since the dynamical equation can then be approximated by:

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} \simeq -2b(x - x_0) \tag{A-12}$$

The corresponding angular frequency ω is related to the second derivative of $V(x)$

at $x = x_0$ by the formula:

$$\omega = \sqrt{\frac{2b}{m}} = \sqrt{\frac{1}{m}\left(\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right)_{x=x_0}} \tag{A-13}$$

Since the amplitude of the motion must remain small, the energy of the harmonic oscillator will be low.

For higher energies E, the particle will be in *periodic but not sinusoidal motion* between the limits x_1 and x_2 (Fig. 2). If we expand the function $x(t)$ in a Fourier series giving the position of the particle, we find, not one, but several, sinusoidal terms; their frequencies are integral multiples of the lowest frequency. We then say that we are dealing with an *anharmonic* oscillator. Note also that, in this case, the period of the motion is not generally $2\pi/\omega$, where ω is given by formula (A-13).

A-3. General properties of the quantum mechanical Hamiltonian

In quantum mechanics, the classical quantities x and p are replaced respectively by the observables X and P, which satisfy:

$$[X, P] = i\hbar \tag{A-14}$$

It is then easy to obtain the Hamiltonian operator of the system from (A-7):

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2 \tag{A-15}$$

Since H is time-independent (conservative system), the quantum mechanical study of the harmonic oscillator reduces to the solution of the eigenvalue equation:

$$H|\varphi\rangle = E|\varphi\rangle \tag{A-16}$$

which is written, in the $\{|x\rangle\}$ representation:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2\right]\varphi(x) = E\ \varphi(x) \tag{A-17}$$

Before undertaking the detailed study of equation (A-16), let us indicate some important properties that can be deduced from the form (A-1) of the potential function:

(*i*) *The eigenvalues of the Hamiltonian are positive.* It can be shown that, in general (Complement $\mathrm{M_{III}}$), if the potential function $V(x)$ has a lower bound, the eigenvalues E of the Hamiltonian $H = \dfrac{P^2}{2m} + V(X)$ are greater than the minimum of $V(x)$:

$$V(x) \geq V_m \quad \text{requires} \quad E > V_m \tag{A-18}$$

For the harmonic oscillator we are studying here, we have chosen the energy origin such that V_m is zero.

(*ii*) *The eigenfunctions of H have a definite parity.* This is due to the fact that the potential $V(x)$ is an even function:

$$V(-x) = V(x) \tag{A-19}$$

We can then (*cf.* Complements F_{II} and C_V) look for eigenfunctions of H, in the $\{|x\rangle\}$ representation, amongst the functions which have a definite parity (in fact, we shall see that the eigenvalues of H are not degenerate; consequently, the wave functions associated with the stationary states are necessarily either even or odd).

(*iii*) *The energy spectrum is discrete.* Whatever the value of the total energy, the classical motion is limited to a bounded region of the Ox axis (Fig. 1), and it can be shown (Complement M_{III}) that in this case, the eigenvalues of the Hamiltonian form a discrete set.

We shall derive these properties (in a more precise form) in the following sections. However, it is interesting to note that they can be obtained simply by applying to the harmonic oscillator some general theorems concerning one-dimensional problems.

B. Eigenvalues of the Hamiltonian

We are now going to study the eigenvalue equation (A-16). First of all, using only the canonical commutation relation (A-14), we shall find the spectrum of the Hamiltonian H written in (A-15).

B-1. Notation

We shall begin by introducing some useful notations.

B-1-a. The $\hat{X}$ and $\hat{P}$ operators

The observables X and P obviously have dimensions (those of a length and a momentum, respectively). Since ω has the dimension of the inverse of a time and $\hbar$, of an action (product of an energy and a time), it is easy to see that the observables $\hat{X}$ and $\hat{P}$ defined by:

$$\begin{aligned} \hat{X} &= \sqrt{\frac{m\omega}{\hbar}} X \\ \hat{P} &= \frac{1}{\sqrt{m\hbar\omega}} P \end{aligned} \tag{B-1}$$

are dimensionless.

If we use these new operators, the canonical commutation relation will be written:

$$[\hat{X}, \hat{P}] = i \tag{B-2}$$

and the Hamiltonian can be put in the form:

$$H = \hbar\omega \, \hat{H} \tag{B-3}$$

with:

$$\hat{H} = \frac{1}{2}\left(\hat{X}^2 + \hat{P}^2\right) \tag{B-4}$$

We shall therefore seek the solutions of the eigenvalue equation:

$$\hat{H}|\varphi_\nu^i\rangle = \varepsilon_\nu \ |\varphi_\nu^i\rangle \tag{B-5}$$

where the operator $\hat{H}$ and the eigenvalues ε_ν are dimensionless. The index ν can belong to either a discrete or a continuous set, and the additional index i enables us to distinguish between the various possible orthogonal eigenvectors associated with the same eigenvalue ε_ν.

B-1-b. The a, $a^\dagger$ and N operators

If $\hat{X}$ and $\hat{P}$ were numbers and not operators, we could write the sum $\hat{X}^2 + \hat{P}^2$ appearing in expression (B-4) for $\hat{H}$ in the form of a product of linear terms, and obtain $(\hat{X} - i\hat{P})(\hat{X} + i\hat{P})$. In fact, since $\hat{X}$ and $\hat{P}$ are non-commuting operators, $\hat{X}^2 + \hat{P}^2$ is not equal to $(\hat{X} - i\hat{P})(\hat{X} + i\hat{P})$. We shall show, however, that the introduction of operators proportional to $\hat{X} + i\hat{P}$ and $\hat{X} - i\hat{P}$ enables us to simplify considerably our search for eigenvalues and eigenvectors of $\hat{H}$.

We therefore set[4]:

$$a = \frac{1}{\sqrt{2}}(\hat{X} + i\hat{P}) \tag{B-6a}$$

$$a^\dagger = \frac{1}{\sqrt{2}}(\hat{X} - i\hat{P}) \tag{B-6b}$$

These formulas can immediately be inverted to yield:

$$\hat{X} = \frac{1}{\sqrt{2}}(a^\dagger + a) \tag{B-7a}$$

$$\hat{P} = \frac{i}{\sqrt{2}}(a^\dagger - a) \tag{B-7b}$$

Since $\hat{X}$ and $\hat{P}$ are Hermitian, a and $a^\dagger$ are not (because of the factor i), but are adjoints of each other.

The commutator of a and $a^\dagger$ is easy to calculate from (B-6) and (B-2):

$$\begin{aligned}[a, a^\dagger] &= \frac{1}{2}\left[\hat{X} + i\hat{P}, \hat{X} - i\hat{P}\right] \\ &= \frac{i}{2}\left[\hat{P}, \hat{X}\right] - \frac{i}{2}\left[\hat{X}, \hat{P}\right]\end{aligned} \tag{B-8}$$

that is:

$$[a, a^\dagger] = 1 \tag{B-9}$$

This relation is completely equivalent to the canonical commutation relation (A-14).

[4] Until now, we have designated operators by capital letters. However, to conform to standard usage, we shall use the small letters a and $a^\dagger$ for the operators (B-6).

Finally, we derive some simple formulas which will be useful in the rest of this chapter. We first calculate $a^\dagger a$:

$$\begin{aligned} a^\dagger a &= \frac{1}{2}(\hat{X} - i\hat{P})(\hat{X} + i\hat{P}) \\ &= \frac{1}{2}(\hat{X}^2 + \hat{P}^2 + i\hat{X}\hat{P} - i\hat{P}\hat{X}) \\ &= \frac{1}{2}(\hat{X}^2 + \hat{P}^2 - 1) \end{aligned} \tag{B-10}$$

Comparing this with expression (B-4), we see that:

$$\hat{H} = a^\dagger a + \frac{1}{2} = \frac{1}{2}(\hat{X} - i\hat{P})(\hat{X} + i\hat{P}) + \frac{1}{2} \tag{B-11}$$

Unlike the situation in the classical case, $\hat{H}$ cannot be put in the form of a product of linear terms. The non-commutativity of $\hat{X}$ and $\hat{P}$ is at the origin of the additional term $1/2$ that appears on the right-hand side of (B-11). Similarly, it can be shown that:

$$\hat{H} = aa^\dagger - \frac{1}{2} \tag{B-12}$$

Let us now introduce the operator N defined by:

$$N = a^\dagger a \tag{B-13}$$

This operator is Hermitian since:

$$N^\dagger = a^\dagger (a^\dagger)^\dagger = a^\dagger a = N \tag{B-14}$$

Moreover, according to (B-11):

$$\hat{H} = N + \frac{1}{2} \tag{B-15}$$

so that *the eigenvectors of $\hat{H}$ are eigenvectors of N, and vice versa.*

Finally, let us calculate the commutators of N with a and $a^\dagger$:

$$\begin{aligned} [N, a] &= [a^\dagger a, a] = a^\dagger [a, a] + [a^\dagger, a]a = -a \\ [N, a^\dagger] &= [a^\dagger a, a^\dagger] = a^\dagger [a, a^\dagger] + [a^\dagger, a^\dagger]a = a^\dagger \end{aligned} \tag{B-16}$$

that is:

$$[N, a] = -a \tag{B-17a}$$

$$[N, a^\dagger] = a^\dagger \tag{B-17b}$$

Our study of the harmonic oscillator will be based on the use of the a, $a^\dagger$ and N operators. We have replaced the eigenvalue equation of H, which we first wrote in the form (B-5), by that of N:

$$N|\varphi_\nu^i\rangle = \nu|\varphi_\nu^i\rangle \tag{B-18}$$

When this equation is solved, we shall know that the eigenvector $|\varphi^i_\nu\rangle$ of N is also an eigenvector of H with the eigenvalue $E_\nu = (\nu + 1/2)\hbar\omega$ [formulas (B-3) and (B-15)]:

$$H|\varphi^i_\nu\rangle = (\nu + 1/2)\hbar\omega \ |\varphi^i_\nu\rangle \tag{B-19}$$

The solution of equation (B-18) will be based on the commutation relation (B-9), which is equivalent to the initial relation (A-14), and on formulas (B-17), which are consequences of it.

B-2. Determination of the spectrum

B-2-a. Lemmas

α. *Lemma I (property of the eigenvalues of N)*

The eigenvalues ν of the operator N are positive or zero.

Consider an arbitrary eigenvector $|\varphi^i_\nu\rangle$ of N. The square of the norm of the vector $a\ |\varphi^i_\nu\rangle$ is positive or zero:

$$\| \ a|\varphi^i_\nu\rangle \ \|^2 = \langle\varphi^i_\nu|a^\dagger a|\varphi^i_\nu\rangle \geq 0 \tag{B-20}$$

Let us then use definition (B-13) of N:

$$\langle\varphi^i_\nu|a^\dagger a|\varphi^i_\nu\rangle = \langle\varphi^i_\nu|N|\varphi^i_\nu\rangle = \nu\langle\varphi^i_\nu|\varphi^i_\nu\rangle \tag{B-21}$$

Since $\langle\varphi^i_\nu|\varphi^i_\nu\rangle$ is positive, comparison of (B-20) and (B-21) shows that:

$$\nu \geq 0 \tag{B-22}$$

β. *Lemma II (properties of the vector $a\,|\varphi^i_\nu\rangle$)*

Let $|\varphi^i_\nu\rangle$ be a (non-zero) eigenvector of N with the eigenvalue ν.

We shall prove the following:

(i) If $\nu = 0$, the ket $a|\varphi^i_{\nu=0}\rangle$ is zero.

(ii) If $\nu > 0$, the ket $a|\varphi^i_\nu\rangle$ is a non-zero eigenvector of N with the eigenvalue $\nu - 1$.

(i) According to (B-21), the square of the norm of $a|\varphi^i_\nu\rangle$ is zero if $\nu = 0$; now, the norm of a vector is zero if and only if this vector is zero. Consequently, if $\nu = 0$ is an eigenvalue of N, all eigenvectors $|\varphi^i_0\rangle$ associated with this eigenvalue satisfy the relation:

$$a\,|\varphi^i_0\rangle = 0 \tag{B-23}$$

Let us now show that relation (B-23) is characteristic of these eigenvectors. Consider a vector $|\varphi\rangle$ which satisfies:

$$a\,|\varphi\rangle = 0 \tag{B-24}$$

Multiply both sides of this equation from the left by $a^\dagger$:

$$a^\dagger a|\varphi\rangle = N|\varphi\rangle = 0 \tag{B-25}$$

Any vector which satisfies (B-24) is therefore an eigenvector of N with the eigenvalue $\nu = 0$.

(ii) Now let us assume that ν is strictly positive. According to (B-21), the vector $a|\varphi_\nu^i\rangle$ is then non-zero, since the square of its norm is not equal to zero.

Let us show that $a|\varphi_\nu^i\rangle$ is an eigenvector of N. To do this, let us apply the operator relation (B-17a) to the vector $|\varphi_\nu^i\rangle$:

$$\begin{aligned} [N,a]|\varphi_\nu^i\rangle &= -a|\varphi_\nu^i\rangle \\ Na|\varphi_\nu^i\rangle &= aN|\varphi_\nu^i\rangle - a|\varphi_\nu^i\rangle \\ &= a\nu|\varphi_\nu^i\rangle - a|\varphi_\nu^i\rangle \end{aligned} \tag{B-26}$$

Therefore:

$$N\big[a|\varphi_\nu^i\rangle\big] = (\nu-1)[a|\varphi_\nu^i\rangle] \tag{B-27}$$

which shows that $a|\varphi_\nu^i\rangle$ is an eigenvector of N with the eigenvalue $\nu - 1$.

γ. *Lemma III (properties of the vector $a^\dagger|\varphi_\nu^i\rangle$)*

Let $|\varphi_\nu^i\rangle$ be a (non-zero) eigenvector of N of eigenvalue ν.
We shall prove the following:

(i) $a^\dagger|\varphi_\nu^i\rangle$ is always non-zero.

(ii) $a^\dagger|\varphi_\nu^i\rangle$ is an eigenvector of N with the eigenvalue $\nu + 1$.

(i) It is easy to calculate the norm of the vector $a^\dagger|\varphi_\nu^i\rangle$, using formulas (B-9) and (B-13):

$$\begin{aligned} \| a^\dagger|\varphi_\nu^i\rangle \|^2 &= \langle\varphi_\nu^i|aa^\dagger|\varphi_\nu^i\rangle \\ &= \langle\varphi_\nu^i|(N+1)|\varphi_\nu^i\rangle \\ &= (\nu+1)\langle\varphi_\nu^i|\varphi_\nu^i\rangle \end{aligned} \tag{B-28}$$

Since, according to lemma I, ν is positive or zero, the ket $a^\dagger|\varphi_\nu^i\rangle$ always has a non-zero norm and, consequently, is never zero.

(ii) The proof of the fact that $a^\dagger|\varphi_\nu^i\rangle$ is an eigenvector of N is analogous to that of lemma II; starting from relation (B-17b) between operators, we obtain:

$$\begin{aligned} [N,a^\dagger]\ |\varphi_\nu^i\rangle &= a^\dagger|\varphi_\nu^i\rangle \\ Na^\dagger|\varphi_\nu^i\rangle &= a^\dagger N|\varphi_\nu^i\rangle + a^\dagger|\varphi_\nu^i\rangle = (\nu+1)a^\dagger|\varphi_\nu^i\rangle \end{aligned} \tag{B-29}$$

B-2-b. The spectrum of N is composed of non-negative integers

Consider an arbitrary eigenvalue ν of N and a non-zero eigenvector $|\varphi_\nu^i\rangle$ associated with this eigenvalue.

According to lemma I, ν is necessarily positive or zero. First, let us assume ν to be non-integral. We are now going to show that such a hypothesis contradicts lemma I and must consequently be excluded. If ν is non-integral, we can always find an integer $n \geq 0$ such that:

$$n < \nu < n+1 \tag{B-30}$$

Now let us consider the series of vectors:

$$|\varphi_\nu^i\rangle, \quad a|\varphi_\nu^i\rangle \ldots a^n|\varphi_\nu^i\rangle \tag{B-31}$$

According to lemma *II*, each of the vectors $a^p|\varphi_\nu^i\rangle$ of this series (with $0 \leq p \leq n$) is non-zero and an eigenvector of N with the eigenvalue $\nu - p$ (*cf.* Fig. 3). The proof is by iteration: $|\varphi_\nu^i\rangle$ is non-zero by hypothesis; $a|\varphi_\nu^i\rangle$ is non-zero (since $\nu > 0$) and corresponds to the eigenvalue $\nu - 1$ of N ...; $a^p|\varphi_\nu^i\rangle$ is obtained when a acts on $a^{p-1}|\varphi_\nu^i\rangle$, an eigenvector of N with the strictly positive eigenvalue $\nu - p + 1$, since $p \leq n$ and $\nu > n$ [*cf.* (B-30)].

Now let a act on the ket $a^n|\varphi_\nu^i\rangle$. Since $\nu - n > 0$ according to (B-30), the action of a on $a^n|\varphi_\nu^i\rangle$ (an eigenvector of N with the eigenvalue $\nu - n > 0$) yields a non-zero vector (lemma *II*). Moreover, again according to lemma *II*, $a^{n+1}|\varphi_\nu^i\rangle$ is an eigenvector of N with the eigenvalue $\nu - n - 1$, which is strictly negative according to (B-30). If ν is non-integral, we can therefore construct a non-zero eigenvector of N with a strictly negative eigenvalue. Since this is impossible, according to lemma *I*, the hypothesis of non-integral ν must be rejected.

What now happens if:

$$\nu = n \tag{B-32}$$

with n a positive integer or zero? In the series of vectors (B-31), $a^n|\varphi_n^i\rangle$ is non-zero and an eigenvector of N with the eigenvalue 0. According to lemma *II* (§ (*i*)), we therefore have:

$$a^{n+1}|\varphi_n^i\rangle = 0 \tag{B-33}$$

The series of vectors obtained by repeated action of the operator a on $|\varphi_n^i\rangle$ is therefore limited when n is integral. It is then never possible to obtain a non-zero eigenvector of N which corresponds to a negative eigenvalue.

In conclusion, ν can only be a non-negative integer.

Lemma *III* can then be used to show that the spectrum of N indeed includes all positive or zero integers. We have already constructed an eigenvector of N with an eigenvalue of zero ($a^n|\varphi_n^i\rangle$). All we must do is let $(a^\dagger)^k$ act on such a vector in order to obtain an eigenvector of N of eigenvalue k, where k is an arbitrary positive integer.

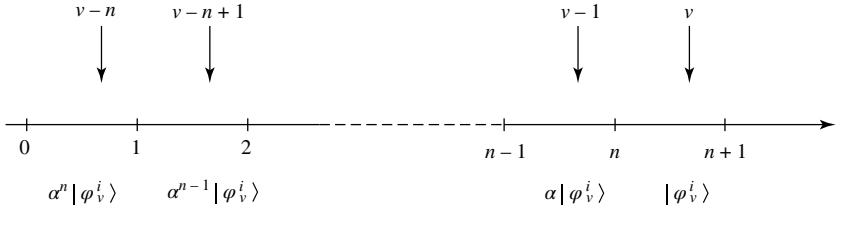


Figure 3: Letting a act several times on the ket $|\varphi_\nu^i\rangle$, we can construct eigenvectors of N with eigenvalues $\nu - 1$, $\nu - 2$ etc...

If we then refer to formula (B-19), we conclude that the eigenvalues of H are of the form:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{B-34}$$

with $n = 0, 1, 2, \ldots$ Therefore, in quantum mechanics, *the energy of the harmonic oscillator is quantized* and cannot take on any arbitrary value. Note also that the smallest value (the ground state) is not zero, but $\hbar\omega/2$ (see § D-2 below).

B-2-c. Interpretation of the a and $a^\dagger$ operators

If we start with an eigenstate $|\varphi_n^i\rangle$ of H corresponding to the eigenvalue $E_n = (n + 1/2)\hbar\omega$, application of the operator a yields an eigenvector associated with the eigenvalue $E_{n-1} = (n + 1/2)\hbar\omega - \hbar\omega$, and application of $a^\dagger$ yields, in the same way, the energy $E_{n+1} = (n + 1/2)\hbar\omega + \hbar\omega$.

For this reason, $a^\dagger$ is said to be a *creation operator* and a an *annihilation operator* (or destruction operator); their action on an eigenvector of N makes an energy quantum $\hbar\omega$ appear or disappear.

B-3. Degeneracy of the eigenvalues

We now show that the energy levels of the one-dimensional harmonic oscillator, given by equation (B-34), are not degenerate.

B-3-a. The ground state is non-degenerate

The eigenstates of H associated with the eigenvalue $E_0 = \hbar\omega/2$, that is, the eigenstates of N associated with the eigenvalue $n = 0$, according to lemma *II* of § B-2-a-β, must all satisfy the equation:

$$a|\varphi_0^i\rangle = 0 \tag{B-35}$$

To find the degeneracy of the E_0 level, all we must do is see how many linearly independent kets satisfy (B-35).

Using definition (B-6a) of a and relations (B-1), we can write (B-35) in the form:

$$\frac{1}{\sqrt{2}}\left[\sqrt{\frac{m\omega}{\hbar}}X + \frac{i}{\sqrt{m\hbar\omega}}P\right]|\varphi_0^i\rangle = 0 \tag{B-36}$$

In the $\{|x\rangle\}$ representation, this relation becomes:

$$\left(\frac{m\omega}{\hbar}x + \frac{\mathrm{d}}{\mathrm{d}x}\right)\varphi_0^i(x) = 0 \tag{B-37}$$

where:

$$\varphi_0^i(x) = \langle x|\varphi_0^i\rangle \tag{B-38}$$

Therefore we must solve a first-order differential equation. Its general solution is:

$$\varphi_0^i(x) = c\,\mathrm{e}^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2} \tag{B-39}$$

where c is the constant of integration. The various solutions of (B-37) are all proportional to each other. Consequently, to within a multiplicative factor, there exists only one ket $|\varphi_0\rangle$ that satisfies (B-35): the ground state $E_0 = \hbar\omega/2$ is not degenerate.

B-3-b. All the states are non-degenerate

We have just seen that the ground state is not degenerate. Let us show by recurrence that this is also the case for all the other states.

All we need prove is that, if the level $E_n = (n + 1/2)\hbar\omega$ is not degenerate, the level $E_{n+1} = (n + 1 + 1/2)\hbar\omega$ is not either. Let us therefore assume that there exists, to within a constant factor, only one vector $|\varphi_n\rangle$ such that:

$$N|\varphi_n\rangle = n|\varphi_n\rangle \tag{B-40}$$

Then consider an eigenvector $|\varphi^i_{n+1}\rangle$ corresponding to the eigenvalue $n + 1$:

$$N|\varphi^i_{n+1}\rangle = (n + 1)|\varphi^i_{n+1}\rangle \tag{B-41}$$

We know that the ket $a|\varphi^i_{n+1}\rangle$ is not zero and that it is an eigenvector of N with the eigenvalue n (*cf.* lemma *II*). Since this ket is not degenerate by hypothesis, there exists a number c^i such that:

$$a|\varphi^i_{n+1}\rangle = c^i|\varphi_n\rangle \tag{B-42}$$

It is simple to invert this equation by applying $a^\dagger$ to both sides:

$$a^\dagger a|\varphi^i_{n+1}\rangle = c^i a^\dagger|\varphi_n\rangle \tag{B-43}$$

that is, taking (B-13) and (B-41) into account:

$$|\varphi^i_{n+1}\rangle = \frac{c^i}{n+1}a^\dagger|\varphi_n\rangle \tag{B-44}$$

We already knew that $a^\dagger|\varphi_n\rangle$ was an eigenvector of N with the eigenvalue $(n + 1)$; we see here that all kets $|\varphi^i_{n+1}\rangle$ associated with the eigenvalue $(n + 1)$ are proportional to $a^\dagger|\varphi_n\rangle$. They are therefore proportional to each other: the eigenvalue $(n + 1)$ is not degenerate.

Thus, since the eigenvalue $n = 0$ is not degenerate (see § B-3-a), the eigenvalue $n = 1$ is not either, nor is $n = 2$, etc...: all the eigenvalues of N and, consequently, all those of H, are non-degenerate. This enables us to write simply $|\varphi_n\rangle$ for the eigenvector of H associated with the eigenvalue $E_n = (n + 1/2)\hbar\omega$.

C. Eigenstates of the Hamiltonian

In this section, we are going to study the principal properties of the eigenstates of the operator N and of the Hamiltonian H.

C-1. The $\{|\varphi_n\rangle\}$ representation

We shall assume that N and H are observables, meaning their eigenvectors constitute a basis in the space $\mathcal{E}_x$, the state space of a particle in a one-dimensional problem (this could be proved by considering the wave functions associated with the eigenstates of N, which we shall calculate in § C-2 below). Since none of the eigenvalues of N (or of H) is degenerate (see § B-3), N (or H) alone constitutes a C.S.C.O. in $\mathcal{E}_x$.

C-1-a. The basis vectors in terms of $|\varphi_0\rangle$

The vector $|\varphi_0\rangle$ associated with $n = 0$ is the vector of $\mathcal{E}_x$ that satisfies:

$$a|\varphi_0\rangle = 0 \tag{C-1}$$

It is defined to within a constant factor; we shall assume $|\varphi_0\rangle$ to be normalized, so the indeterminacy is reduced to a global phase factor of the form $e^{i\theta}$, with θ real.

According to lemma *III* of § B-2-a, the vector $|\varphi_1\rangle$ which corresponds to $n = 1$ is proportional to $a^\dagger|\varphi_0\rangle$:

$$|\varphi_1\rangle = c_1 a^\dagger|\varphi_0\rangle \tag{C-2}$$

We shall determine c_1 by requiring $|\varphi_1\rangle$ to be normalized and choosing the phase of $|\varphi_1\rangle$ (relative to $|\varphi_0\rangle$) such that c_1 is real and positive. The square of the norm of $|\varphi_1\rangle$, according to (C-2), is equal to:

$$\begin{aligned}\langle\varphi_1|\varphi_1\rangle &= |c_1|^2\langle\varphi_0|aa^\dagger|\varphi_0\rangle \\ &= |c_1|^2\langle\varphi_0|(a^\dagger a + 1)|\varphi_0\rangle\end{aligned} \tag{C-3}$$

where (B-9) has been used. Since $|\varphi_0\rangle$ is a normalized eigenstate of $N = a^\dagger a$ with the eigenvalue zero, we find:

$$\langle\varphi_1|\varphi_1\rangle = |c_1|^2 = 1 \tag{C-4}$$

With the preceding phase convention, we have $c_1 = 1$ and, consequently:

$$|\varphi_1\rangle = a^\dagger|\varphi_0\rangle \tag{C-5}$$

Similarly, we can construct $|\varphi_2\rangle$ from $|\varphi_1\rangle$:

$$|\varphi_2\rangle = c_2\, a^\dagger|\varphi_1\rangle \tag{C-6}$$

We require $|\varphi_2\rangle$ to be normalized and choose its phase such that c_2 is real and positive:

$$\begin{aligned}\langle\varphi_2|\varphi_2\rangle &= |c_2|^2\langle\varphi_1|aa^\dagger|\varphi_1\rangle \\ &= |c_2|^2\langle\varphi_1|(a^\dagger a + 1)|\varphi_1\rangle \\ &= 2|c_2|^2 = 1\end{aligned} \tag{C-7}$$

Therefore:

$$|\varphi_2\rangle = \frac{1}{\sqrt{2}}a^\dagger|\varphi_1\rangle = \frac{1}{\sqrt{2}}(a^\dagger)^2|\varphi_0\rangle \tag{C-8}$$

if we take (C-5) into account.

This procedure can easily be generalized. If we know $|\varphi_{n-1}\rangle$ (which is normalized), then the normalized vector $|\varphi_n\rangle$ is written:

$$|\varphi_n\rangle = c_n a^\dagger|\varphi_{n-1}\rangle \tag{C-9}$$

Since:

$$\begin{aligned}\langle\varphi_n|\varphi_n\rangle &= |c_n|^2\langle\varphi_{n-1}|aa^\dagger|\varphi_{n-1}\rangle \\ &= n|c_n|^2 = 1\end{aligned} \tag{C-10}$$

we choose, with the same phase conventions as above:

$$c_n = \frac{1}{\sqrt{n}} \tag{C-11}$$

With these successive phase choices, we can obtain all the $|\varphi_n\rangle$ from $|\varphi_0\rangle$:

$$\begin{aligned}|\varphi_n\rangle &= \frac{1}{\sqrt{n}}a^\dagger|\varphi_{n-1}\rangle = \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n-1}}(a^\dagger)^2|\varphi_{n-2}\rangle = ... \\ &= \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n-1}}...\frac{1}{\sqrt{2}}(a^\dagger)^n|\varphi_0\rangle\end{aligned} \tag{C-12}$$

that is:

$$|\varphi_n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|\varphi_0\rangle \tag{C-13}$$

C-1-b. Orthonormalization and closure relations

Since H is Hermitian, the kets $|\varphi_n\rangle$ corresponding to different values of n are orthogonal. Since each of them is also normalized, they satisfy the orthonormalization relation:

$$\langle\varphi_{n'}|\varphi_n\rangle = \delta_{nn'} \tag{C-14}$$

In addition, H is an observable (we shall assume this here without proof); the set of the $|\varphi_n\rangle$ therefore constitutes a basis in $\mathcal{E}_x$. This is expressed by the closure relation:

$$\sum_n |\varphi_n\rangle\langle\varphi_n| = 1 \tag{C-15}$$

Comment:

It can be verified directly from expression (C-13) that the kets $|\varphi_n\rangle$ are orthonormal:

$$\langle\varphi_{n'}|\varphi_n\rangle = \frac{1}{\sqrt{n!n'!}}\langle\varphi_0|a^{n'}\,a^{\dagger n}|\varphi_0\rangle \tag{C-16}$$

But:

$$\begin{aligned}a^{n'}\,a^{\dagger n}|\varphi_0\rangle &= a^{n'-1}(aa^\dagger)a^{\dagger n-1}|\varphi_0\rangle \\ &= a^{n'-1}(a^\dagger a+1)a^{\dagger n-1}|\varphi_0\rangle \\ &= n\,a^{n'-1}a^{\dagger n-1}|\varphi_0\rangle\end{aligned} \tag{C-17}$$

(using the fact that $a^{\dagger^{n-1}}|\varphi_0\rangle$ is an eigenstate of $N = a^\dagger a$ with the eigenvalue $n-1$). Thus can we reduce the exponents of a and $a^\dagger$ by iteration. We obtain, finally:

$$\text{if } \; n < n' : \langle\varphi_0|a^{n'}a^{\dagger^n}|\varphi_0\rangle = n \times (n-1) \times ...2 \times 1 \langle\varphi_0|a^{n'-n}|\varphi_0\rangle \tag{C-18a}$$

$$\text{if } \; n > n' : \langle\varphi_0|a^{n'}a^{\dagger^n}|\varphi_0\rangle = n \times (n-1)...(n-n'+1)\langle\varphi_0|(a^\dagger)^{n-n'}|\varphi_0\rangle \tag{C-18b}$$

$$\text{if } \; n = n' : \langle\varphi_0|a^{n'}a^{\dagger^n}|\varphi_0\rangle = n \times (n-1) \times ...2 \times 1 \langle\varphi_0|\varphi_0\rangle \tag{C-18c}$$

The expression (C-18a) is zero because $a|\varphi_0\rangle = 0$. Similarly, (C-18b) is equal to zero because $\langle\varphi_0|(a^\dagger)^{n-n'}|\varphi_0\rangle$ can be considered to be the scalar product of $|\varphi_0\rangle$ and the bra associated with $a^{n-n'}|\varphi_0\rangle$, which is zero if $n > n'$. Finally, if we substitute (C-18c) into (C-16), we see that $\langle\varphi_n|\varphi_n\rangle$ is equal to 1.

C-1-c. Action of the various operators

The observables X and P are linear combinations of the operators a and $a^\dagger$ [formulas (B-1) and (B-7)]. Consequently, all physical quantities can be expressed in terms of a and $a^\dagger$. Now, the action of a and $a^\dagger$ on the $|\varphi_n\rangle$ vectors is especially simple [see equations (C-19) below]. In most cases, it is therefore desirable to use the $\{|\varphi_n\rangle\}$ representation to calculate the matrix elements and mean values of the various observables.

With the phase conventions introduced in § C-1-a above, the action of the a and $a^\dagger$ operators on the vectors of the $\{|\varphi_n\rangle\}$ basis is given by:

$$\boxed{a^\dagger|\varphi_n\rangle = \sqrt{n+1}\ |\varphi_{n+1}\rangle} \tag{C-19a}$$

$$\boxed{a|\varphi_n\rangle = \sqrt{n}\ |\varphi_{n-1}\rangle} \tag{C-19b}$$

We have already proved (C-19a): it suffices to replace n by $n+1$ in equations (C-9) and (C-11). To obtain (C-19b), multiply both sides of (C-9) on the left by the operator a and use (C-11):

$$a|\varphi_n\rangle = \frac{1}{\sqrt{n}}\ aa^\dagger|\varphi_{n-1}\rangle = \frac{1}{\sqrt{n}}(a^\dagger a + 1)|\varphi_{n-1}\rangle = \sqrt{n}\ |\varphi_{n-1}\rangle \tag{C-20}$$

Comment:

The adjoint equations of (C-19a) and (C-19b) are:

$$\langle\varphi_n|a = \sqrt{n+1}\ \langle\varphi_{n+1}| \tag{C-21a}$$

$$\langle\varphi_n|a^\dagger = \sqrt{n}\ \langle\varphi_{n-1}| \tag{C-21b}$$

Note that a decreases or increases n by one unit depending on whether it acts on the ket $|\varphi_n\rangle$ or on the bra $\langle\varphi_n|$. Similarly, $a^\dagger$ increases or decreases n by one unit, depending on whether it acts on the ket $|\varphi_n\rangle$ or on the bra $\langle\varphi_n|$.

Starting with (C-19) and using (B-1) and (B-7), we immediately find the expressions for the kets $X|\varphi_n\rangle$ and $P|\varphi_n\rangle$:

$$X|\varphi_n\rangle = \sqrt{\frac{\hbar}{m\omega}}\frac{1}{\sqrt{2}}\,(a^\dagger + a)|\varphi_n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left[\sqrt{n+1}\;|\varphi_{n+1}\rangle + \sqrt{n}\;|\varphi_{n-1}\rangle\right] \tag{C-22a}$$

$$P|\varphi_n\rangle = \sqrt{m\hbar\omega}\frac{i}{\sqrt{2}}(a^\dagger - a)|\varphi_n\rangle = i\sqrt{\frac{m\hbar\omega}{2}}\;\left[\sqrt{n+1}\;|\varphi_{n+1}\rangle - \sqrt{n}\;|\varphi_{n-1}\rangle\right] \tag{C-22b}$$

The matrix elements of the a, $a^\dagger$, X and P operators in the $\{|\varphi_n\rangle\}$ representation are therefore:

$$\langle\varphi_{n'}|a|\varphi_n\rangle = \sqrt{n}\;\delta_{n',n-1} \tag{C-23a}$$

$$\langle\varphi_{n'}|a^\dagger|\varphi_n\rangle = \sqrt{n+1}\;\delta_{n',n+1} \tag{C-23b}$$

$$\langle\varphi_{n'}|X|\varphi_n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left[\sqrt{n+1}\;\delta_{n',n+1} + \sqrt{n}\;\delta_{n',n-1}\right] \tag{C-23c}$$

$$\langle\varphi_{n'}|P|\varphi_n\rangle = i\sqrt{\frac{m\hbar\omega}{2}}\left[\sqrt{n+1}\;\delta_{n',n+1} - \sqrt{n}\;\delta_{n',n-1}\right] \tag{C-23d}$$

The matrices representing a and $a^\dagger$ are indeed Hermitian conjugates of each other, as can be seen from their explicit expressions:

$$(a) = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & .. & 0 & .. \\ 0 & 0 & \sqrt{2} & 0 & .. & 0 & .. \\ 0 & 0 & 0 & \sqrt{3} & .. & 0 & .. \\ 0 & 0 & 0 & 0 & .. & 0 & .. \\ .. & .. & .. & .. & .. & \sqrt{n} & .. \\ 0 & 0 & 0 & 0 & 0 & 0 & .. \\ .. & .. & .. & .. & .. & .. & .. \end{pmatrix} \tag{C-24a}$$

and:

$$(a^\dagger) = \begin{pmatrix} 0 & 0 & 0 & 0 & .. & 0 & .. \\ \sqrt{1} & 0 & 0 & 0 & .. & 0 & .. \\ 0 & \sqrt{2} & 0 & 0 & .. & 0 & .. \\ 0 & 0 & \sqrt{3} & 0 & .. & 0 & .. \\ .. & .. & .. & .. & .. & 0 & .. \\ 0 & 0 & 0 & 0 & \sqrt{n+1} & 0 & .. \\ .. & .. & .. & .. & .. & .. & .. \end{pmatrix} \tag{C-24b}$$

As for the matrices representing X and P, they are both Hermitian: the matrix associated with X is, to within a constant factor, the sum of the two preceding ones; the matrix associated with P is proportional to their difference, but the presence of the factor i in (C-22b) re-establishes its Hermiticity.

C-2. Wave functions associated with the stationary states

We shall now use the $\{|x\rangle\}$ representation and write the functions $\varphi_n(x) = \langle x|\varphi_n\rangle$ which then represent the eigenstates of the Hamiltonian.

We have already determined the function $\varphi_0(x)$ which represents the ground state $|\varphi_0\rangle$ (*cf.* § B-3-a):

$$\varphi_0(x) = \langle x|\varphi_0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2} \tag{C-25}$$

The constant that appears before the exponential insures the normalization of $\varphi_0(x)$.

To obtain the functions $\varphi_n(x)$ associated with the other stationary states of the harmonic oscillator, all we need to do is use expression (C-13) for the ket $|\varphi_n\rangle$ and the fact that, in the $\{|x\rangle\}$ representation, $a^\dagger$ is represented by:

$$\frac{1}{\sqrt{2}}\left[\sqrt{\frac{m\omega}{\hbar}}x - \sqrt{\frac{\hbar}{m\omega}}\frac{\mathrm{d}}{\mathrm{d}x}\right]$$

since X is represented by multiplication by x, and P by $\frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x}$ [formula (B-6b)]. Thus we obtain:

$$\begin{aligned} \varphi_n(x) = \langle x|\varphi_n\rangle &= \frac{1}{\sqrt{n!}}\langle x|(a^\dagger)^n|\varphi_0\rangle \\ &= \frac{1}{\sqrt{n!}}\frac{1}{\sqrt{2^n}}\left[\sqrt{\frac{m\omega}{\hbar}}x - \sqrt{\frac{\hbar}{m\omega}}\frac{\mathrm{d}}{\mathrm{d}x}\right]^n \varphi_0(x) \end{aligned} \tag{C-26}$$

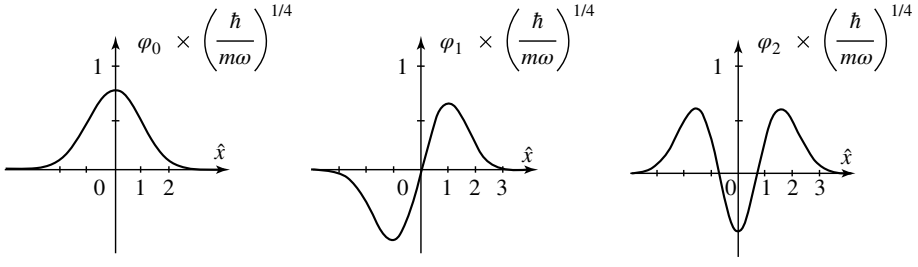


Figure 4: Wave functions associated with the first three levels of a harmonic oscillator.

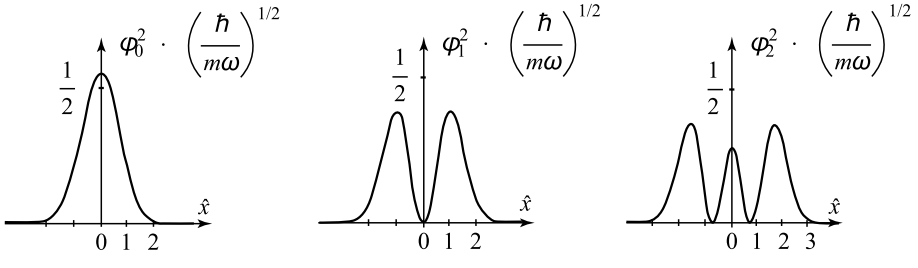


Figure 5: Probability densities associated with the first three levels of a harmonic oscillator.

that is:

$$\varphi_n(x) = \left[\frac{1}{2^n n!}\ \left(\frac{\hbar}{m\omega}\right)^n\ \right]^{1/2} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \left[\frac{m\omega}{\hbar}x - \frac{\mathrm{d}}{\mathrm{d}x}\right]^n\ \mathrm{e}^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2} \tag{C-27}$$

It is easy to see from this expression that $\varphi_n(x)$ is the product of $\mathrm{e}^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2}$ and a polynomial of degree n and parity $(-1)^n$, called a *Hermite polynomial* (*cf.* Complements $\mathrm{B_V}$ and $\mathrm{C_V}$).

A simple calculation gives the first several functions $\varphi_n(x)$:

$$\begin{aligned}
\varphi_1(x) &= \left[\frac{4}{\pi}\left(\frac{m\omega}{\hbar}\right)^3\right]^{1/4} x\,\mathrm{e}^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2} \\
\varphi_2(x) &= \left(\frac{m\omega}{4\pi\hbar}\right)^{1/4} \left[2\frac{m\omega}{\hbar}x^2 - 1\right] \mathrm{e}^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2}
\end{aligned} \tag{C-28}$$

These functions are shown in Figure 4, and the corresponding probability densities in Figure 5. Figure 6 gives the shape of the wave function $\varphi_n(x)$ and that of the probability density $|\varphi_n(x)|^2$ for $n = 10$.

We see from these figures that when n increases, the region of the Ox axis in which $\varphi_n(x)$ takes on non-negligible values becomes larger. This corresponds to the fact,

in classical mechanics, that the amplitude of the particle's motion increases with the energy [*cf.* Fig. 1 and relation (A-8)]. It follows that the mean value of the potential energy grows with n [*cf.* comment (*ii*) of § D-1], since $\varphi_n(x)$, when n is large, takes on non-negligible values in regions of the x-axis where $V(x)$ is large. Moreover, we see in these figures that the number of zeros of $\varphi_n(x)$ is n (*cf.* Complement $\mathrm{B_V}$, where this property is derived). This implies that the mean kinetic energy of the particle increases with n [*cf.* comment (*ii*) of § D-1], since this energy is given by:

$$\frac{1}{2m}\langle P^2\rangle = -\frac{\hbar^2}{2m}\int_{-\infty}^{+\infty}\varphi_n^*(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi_n(x)\,\mathrm{d}x \tag{C-29}$$

When the number of zeros of $\varphi_n(x)$ increases, the curvature of the wave function increases, and, in (C-29), the second derivative $\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi_n(x)$ takes on larger and larger values.

Finally, when n is large, we observe (see, for example, Figure 6) that the probability density $|\varphi_n(x)|^2$ is large for $x \simeq \pm x_M$ [where x_M is the amplitude of the classical motion of energy E_n; *cf.* (A-8)]. This result is related to a feature of the motion predicted by classical mechanics: the classical particle has a zero velocity at $x = \pm x_M$; therefore, on the average, it spends more time in the neighborhood of these two points than in the center of the interval $-x_M \leq x \leq x_M$.

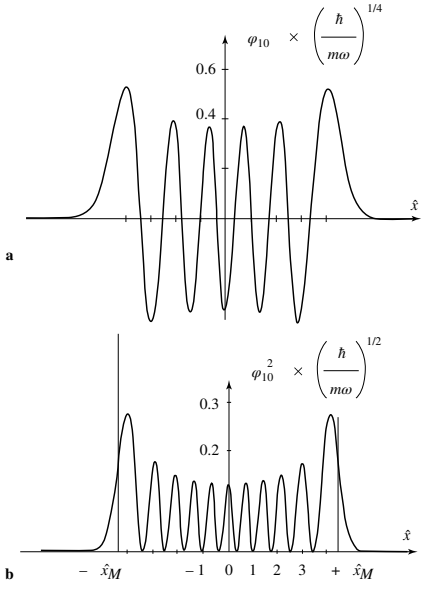


Figure 6: Shape of the wave function (fig. a) and of the probability density (fig. b) for the $n = 10$ level of a harmonic oscillator.

D. Discussion

D-1. Mean values and root mean square deviations of X and P in a state $|\varphi_n\rangle$

Neither X nor P commutes with H, and the eigenstates $|\varphi_n\rangle$ of H are not eigenstates of X or P. Consequently, if the harmonic oscillator is in a stationary state $|\varphi_n\rangle$, a measurement of the observable X or the observable P can, *a priori*, yield any result (since the spectra of X and P include all real numbers). We shall now calculate the mean values of X and P in such a stationary state and then their root mean square deviations ΔX and ΔP, which will enable us to verify the uncertainty relation.

As we indicated in § C-1-c, we shall perform these calculations with the help of the operators a and $a^\dagger$. As far as the mean values of X and P are concerned, the result follows directly from formulas (C-22), which show that neither X nor P has diagonal matrix elements:

$$\begin{aligned}\langle\varphi_n|X|\varphi_n\rangle &= 0\\ \langle\varphi_n|P|\varphi_n\rangle &= 0\end{aligned} \tag{D-1}$$

To obtain the root mean square deviations ΔX and ΔP, we must calculate the mean values of X^2 and P^2:

$$\begin{aligned}(\Delta X)^2 &= \langle\varphi_n|X^2|\varphi_n\rangle - (\langle\varphi_n|X|\varphi_n\rangle)^2 = \langle\varphi_n|X^2|\varphi_n\rangle\\ (\Delta P)^2 &= \langle\varphi_n|P^2|\varphi_n\rangle - (\langle\varphi_n|P|\varphi_n\rangle)^2 = \langle\varphi_n|P^2|\varphi_n\rangle\end{aligned} \tag{D-2}$$

Now, according to (B-1) and (B-7):

$$\begin{aligned}X^2 &= \frac{\hbar}{2m\omega}(a^\dagger + a)(a^\dagger + a)\\ &= \frac{\hbar}{2m\omega}(a^{\dagger 2} + aa^\dagger + a^\dagger a + a^2)\\ P^2 &= -\frac{m\hbar\omega}{2}(a^\dagger - a)(a^\dagger - a)\\ &= -\frac{m\hbar\omega}{2}(a^{\dagger 2} - aa^\dagger - a^\dagger a + a^2)\end{aligned} \tag{D-3}$$

The terms in a^2 and $a^{\dagger 2}$ do not contribute to the diagonal matrix elements, since $a^2|\varphi_n\rangle$ is proportional to $|\varphi_{n-2}\rangle$, and $a^{\dagger 2}|\varphi_n\rangle$ to $|\varphi_{n+2}\rangle$; both are orthogonal to $|\varphi_n\rangle$. On the other hand:

$$\begin{aligned}\langle\varphi_n|(a^\dagger a + aa^\dagger)|\varphi_n\rangle &= \langle\varphi_n|(2a^\dagger a + 1)|\varphi_n\rangle\\ &= 2n + 1\end{aligned} \tag{D-4}$$

Consequently:

$$(\Delta X)^2 = \langle\varphi_n|X^2|\varphi_n\rangle = \left(n + \frac{1}{2}\right)\frac{\hbar}{m\omega} \tag{D-5a}$$

$$(\Delta P)^2 = \langle\varphi_n|P^2|\varphi_n\rangle = \left(n + \frac{1}{2}\right)m\hbar\omega \tag{D-5b}$$

The product $\Delta X.\Delta P$ is therefore equal to:

$$\Delta X.\Delta P = \left(n + \frac{1}{2}\right)\hbar \tag{D-6}$$

We again find (*cf.* Complement C_{III}) that it is greater than or equal to $\hbar/2$. In fact, this lower bound is attained for $n = 0$, that is, for the ground state (§ D-2 below).

Comments:

(*i*) If x_M denotes the amplitude of the classical motion whose energy is given by $E_n = (n + 1/2)\hbar\omega$, it is easy to see, using (A-8) and (D-5a), that:

$$\Delta X = \frac{1}{\sqrt{2}} x_M \tag{D-7}$$

Similarly, if p_M denotes the oscillation amplitude of the corresponding classical momentum:

$$p_M = m\omega\, x_M \tag{D-8}$$

we obtain:

$$\Delta P = \frac{1}{\sqrt{2}} p_M \tag{D-9}$$

It is not surprising that ΔX is of the order of the interval $[-x_M, +x_M]$ over which the classical motion occurs (*cf.* Fig. 1): we saw at the end of § C, that it is approximately inside this interval that $\varphi_n(x)$ takes on non-negligible values. Furthermore, it is easy to understand why, when n increases, so does ΔX. For large n, the probability density $|\varphi_n(x)|^2$ has two symmetric peaks situated approximately at $x = \pm x_M$. The root mean square deviation cannot be much smaller than the distance between these peaks, even if each of them is very sharp (*cf.* Chap. III, § C-5 and the discussion of § 1-b of Complement A_{III}). An analogous argument can be set forth for ΔP (*cf.* Complement D_V).

(*ii*) The mean potential energy of a particle in the state $|\varphi_n\rangle$ is:

$$\langle V(X)\rangle = \frac{1}{2} m\omega^2 \langle X^2\rangle \tag{D-10}$$

that is, since $\langle X\rangle$ is zero [*cf.* (D-1)]:

$$\langle V(X)\rangle = \frac{1}{2} m\omega^2 (\Delta X)^2 \tag{D-11}$$

Similarly, we could find the mean kinetic energy of this particle:

$$\left\langle \frac{P^2}{2m} \right\rangle = \frac{1}{2m} (\Delta P)^2 \tag{D-12}$$

Substituting relations (D-5) into (D-11) and (D-12), we obtain:

$$\begin{aligned} \langle V(X)\rangle &= \frac{1}{2}\left(n+\frac{1}{2}\right)\hbar\omega = \frac{E_n}{2} \\ \left\langle \frac{P^2}{2m}\right\rangle &= \frac{1}{2}\left(n+\frac{1}{2}\right)\hbar\omega = \frac{E_n}{2} \end{aligned} \tag{D-13}$$

The mean potential and kinetic energies are therefore equal. This is an illustration of the virial theorem (*cf.* exercise 10 of Complement $\mathrm{L_{III}}$).

(*iii*) A stationary state $|\varphi_n\rangle$ has no equivalent in classical mechanics: its energy is not zero although the mean values $\langle X\rangle$ and $\langle P\rangle$ are. Nevertheless, there is a certain analogy between the state $|\varphi_n\rangle$ and that of a classical particle whose position is given by (A-5) [where x_M is related to the energy E_n by relation (A-8)], but for which the initial phase φ of the motion is chosen at random (all values included between 0 and 2π have the same probability). The mean values of x and p are then zero, since:

$$\begin{cases} \overline{x}_{cl} = x_M \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \cos(\omega t-\varphi)\,\mathrm{d}\varphi = 0 \\ \overline{p}_{cl} = -p_M \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \sin(\omega t-\varphi)\,\mathrm{d}\varphi = 0 \end{cases} \tag{D-14}$$

Moreover, we find, for the root mean square deviations of the position and the momentum, values identical to those of the state $|\varphi_n\rangle$ [formulas (D-7) and (D-9)]:

$$\begin{aligned} \overline{x}^2_{cl} &= x^2_M \frac{1}{2\pi}\int_0^{2\pi} \cos^2(\omega t-\varphi)\,\mathrm{d}\varphi = \frac{x^2_M}{2} \\ \overline{p}^2_{cl} &= p^2_M \frac{1}{2\pi}\int_0^{2\pi} \sin^2(\omega t-\varphi)\,\mathrm{d}\varphi = \frac{p^2_M}{2} \end{aligned} \tag{D-15}$$

that is:

$$\begin{aligned} \delta x_{cl} &= \sqrt{\overline{x}^2_{cl} - (\overline{x}_{cl})^2} = \frac{x_M}{\sqrt{2}} \\ \delta p_{cl} &= \sqrt{\overline{p}^2_{cl} - (\overline{p}_{cl})^2} = \frac{p_M}{\sqrt{2}} \end{aligned} \tag{D-16}$$

D-2. Properties of the ground state

In classical mechanics, the lowest energy of the harmonic oscillator is obtained when the particle is at rest (zero momentum and kinetic energy) at the x-origin ($x = 0$ and therefore zero potential energy). The situation is completely different in quantum mechanics: the minimum energy state is $|\varphi_0\rangle$, whose *energy is not zero*, and the associated wave function has a certain *spatial extension*, characterized by the root mean square deviation $\Delta X = \sqrt{\hbar/2m\omega}$.

This essential difference between the quantum and classical results can be seen to have its source in the uncertainty relations, which forbid the simultaneous minimization

of the kinetic energy and the potential energy. As we pointed out in Complements C_I and M_{III}, the ground state corresponds to a compromise in which the sum of these two energies is as small as possible.

In the special case of a harmonic oscillator, it is possible to state these qualitative considerations semi-quantitatively, and thus find the order of magnitude of the energy and the spatial extension of the ground state. If the distance ξ characterizes this spatial extension, the mean potential energy will be of the order of:

$$\overline{V} \simeq \frac{1}{2} m\omega^2 \xi^2 \tag{D-17}$$

But ΔP is then equal to about $\hbar/\xi$, so the mean kinetic energy is approximately:

$$\overline{T} = \frac{\overline{p^2}}{2m} \simeq \frac{\hbar^2}{2m\xi^2} \tag{D-18}$$

The order of magnitude of the total energy is therefore:

$$\overline{E} = \overline{T} + \overline{V} \simeq \frac{\hbar^2}{2m\xi^2} + \frac{1}{2} m\omega^2 \xi^2 \tag{D-19}$$

The variation of $\overline{T}$, $\overline{V}$ and $\overline{E}$ with respect to ξ is shown in Figure 7. For small values of ξ, $\overline{T}$ prevails over $\overline{V}$; the opposite occurs for large values of ξ. The ground state therefore corresponds approximately to the minimum of the function (D-19); it is easy to see that this minimum occurs at:

$$\xi_m \simeq \sqrt{\frac{\hbar}{m\omega}} \tag{D-20}$$

and is equal to:

$$\overline{E}_m \simeq \hbar\omega \tag{D-21}$$

We again find the correct orders of magnitude of E_0 and ΔX in the state $|\varphi_0\rangle$.

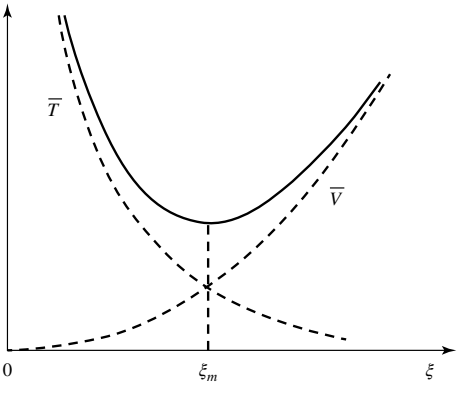


Figure 7: Variation of the potential energy $\overline{V}$ and of the kinetic energy $\overline{T}$ with respect to a parameter ξ characterizing the spatial extension of the wave function about $x = 0$. Since the potential energy is at a minimum at $x = 0$, $\overline{V}$ is a function that increases with $\xi(\overline{V} \propto \xi^2)$. On the other hand, according to Heisenberg's uncertainty relation, the kinetic energy $\overline{T}$ is a decreasing function of ξ. The lowest possible total energy, obtained for $\xi = \xi_m$, results from a compromise in which the sum $\overline{T} + \overline{V}$ (solid line) is at a minimum.

The harmonic oscillator possesses the pecularity that, because of the form of the potential $V(x)$, the product $\Delta X.\Delta P$ actually attains its lower bound, $\hbar/2$, in the ground state $|\varphi_0\rangle$ [formula (D-6)]. This is related to the fact (*cf.* Complement C_{III}) that the wave function of the ground state is Gaussian.

D-3. Time evolution of the mean values

Consider a harmonic oscillator whose state at $t = 0$ is:

$$|\psi(0)\rangle = \sum_{n=0}^{\infty} c_n(0)|\varphi_n\rangle \tag{D-22}$$

($|\psi(0)\rangle$ is assumed to be normalized). Its state $|\psi(t)\rangle$ at t can be obtained by using rule (D-54) of Chapter III:

$$\begin{aligned}|\psi(t)\rangle &= \sum_{n=0}^{\infty} c_n(0)\, \mathrm{e}^{-i\ E_n t/\hbar}|\varphi_n\rangle \\ &= \sum_{n=0}^{\infty} c_n(0)\, \mathrm{e}^{-i\left(n+\frac{1}{2}\right)\omega t}|\varphi_n\rangle\end{aligned} \tag{D-23}$$

The mean value of any physical quantity A is therefore given as a function of time by:

$$\langle\psi(t)|A|\psi(t)\rangle = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} c_m^*(0)\ c_n(0)\ A_{mn}\, \mathrm{e}^{i(m-n)\omega t} \tag{D-24}$$

with:

$$A_{mn} = \langle\varphi_m|A|\varphi_n\rangle \tag{D-25}$$

Since m and n are integers, the time evolution of the mean values involves only the frequency $\omega/2\pi$ and its various harmonics, which constitute the Bohr frequencies of the harmonic oscillator.

Let us consider, in particular, the mean values of the observables X and P. According to formulas (C-22), the only non-zero matrix elements X_{mn} and P_{mn} are those for which $m = n \pm 1$. Consequently, the mean values of X and P include only terms in $\mathrm{e}^{\pm i\omega t}$; they are sinusoidal functions of time with angular frequency ω. This obviously relates to the classical solution of the harmonic oscillator problem. Moreover, as we pointed out in the discussion of Ehrenfest's theorem (Chap. III, § D-1-d-γ), the form of the harmonic oscillator potential implies that for all $|\psi\rangle$ the mean values of X and P rigorously satisfy the classical equations of motion. Thus, according to general formulas (D-34) and (D-35) of Chapter III:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle = \frac{1}{i\hbar}\langle[X, H]\rangle = \frac{\langle P\rangle}{m} \tag{D-26a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle = \frac{1}{i\hbar}\langle[P, H]\rangle = -m\omega^2\langle X\rangle \tag{D-26b}$$

If we integrate these equations, we obtain:

$$\langle X\rangle(t) = \langle X\rangle(0)\cos\omega t + \frac{1}{m\omega}\langle P\rangle(0)\sin\omega t$$
$$\langle P\rangle(t) = \langle P\rangle(0)\cos\omega t + m\omega\langle X\rangle(0)\sin\omega t \tag{D-27}$$

We again find the sinusoidal form indicated by formula (D-24).

Comment:

It is important to note that this analogy with the classical situation appears only when $|\psi(0)\rangle$ is a superposition of states $|\varphi_n\rangle$ of the type of (D-22), where several coefficients $c_n(0)$ are non-zero. If all these coefficients except one are equal to zero, the oscillator is in a stationary state and the mean values of all the observables are constant over time.

It follows that, in a stationary state $|\varphi_n\rangle$, the behavior of a harmonic oscillator is totally different from that predicted by classical mechanics, even if n is very large (the limit of large quantum numbers). If we want to construct a wave packet whose average position oscillates over time, we must superpose different states $|\varphi_n\rangle$ (see Complement G_V).

References and suggestions for further reading:

Dirac (1.13), § 34; Messiah (1.17), Chap. XII.

A_V : SOME EXAMPLES OF HARMONIC OSCILLATORS	Demonstrates, with some examples chosen from various fields, the importance of the quantum mechanical harmonic oscillator in physics. Semi-quantitative and rather simple; recommended for a first reading.
B_V : STUDY OF THE STATIONARY STATES IN THE $\{\|x\rangle\}$ REPRESENTATION. HERMITE POLYNOMIALS	Technical study of the stationary wave functions of the harmonic oscillator. Intended to serve as a reference.
C_V : SOLVING THE EIGENVALUE EQUATION OF THE HARMONIC OSCILLATOR BY THE POLYNOMIAL METHOD	Another method that yields the results of Chapter V. Reveals the relation between energy quantization and the behavior of the wave functions at infinity. Moderately difficult.
D_V : STUDY OF THE STATIONARY STATES IN THE $\{\|p\rangle\}$ REPRESENTATION	Shows that, in a stationary state of the harmonic oscillator, the momentum probability distribution has the same form as that of the position. Fairly simple.
E_V : THE ISOTROPIC THREE-DIMENSIONAL HARMONIC OSCILLATOR	Generalisation of the results of Chapter V to three dimensions. Recommended for a first reading, since it is simple and important.
F_V : A CHARGED HARMONIC OSCILLATOR PLACED IN A UNIFORM ELECTRIC FIELD	A direct and simple application of the results of Chapter V (except for § 3, which uses the translation operator introduced in Complement E_{II}). Recommended for a first reading.
G_V : COHERENT "QUASI-CLASSICAL" STATES OF THE HARMONIC OSCILLATOR	Detailed study of the "quasi-classical" states of the harmonic oscillator, which illustrates the relation between quantum and classical mechanics. Important because of its applications to the quantum theory of radiation. Moderately difficult; can be omitted in a first reading.
H_V : NORMAL VIBRATIONAL MODES OF TWO COUPLED HARMONIC OSCILLATORS	Study, in the very simple case of two coupled harmonic oscillators, of the normal vibrational modes of a system. Recommended, since it is simple and physically important.

J_V : VIBRATIONAL MODES OF AN INFINITE LINEAR CHAIN OF COUPLED HARMONIC OSCILLATORS; PHONONS **K_V : VIBRATIONAL MODES OF A CONTINUOUS PHYSICAL SYSTEM. APPLICATION TO RADIATION; PHOTONS**	J_V, K_V: Introduction, using simplified models, of concepts that are particularly important in physics. Rather difficult (graduate level); can be reserved for later study. J_V: determination of the normal vibrational modes of a linear chain of coupled oscillators, leading to the notion of phonon, fundamental in solid state physics. K_V: normal vibrational modes of a continuous system. A simple way to introduce photons in the quantum mechanical study of the electromagnetic field.
L_V : THE ONE-DIMENSIONAL HARMONIC OSCILLATOR IN THERMODYNAMIC EQUILIBRIUM AT A TEMPERATURE T	Application of the density operator (introduced in Complement E_{III}) to a harmonic oscillator in thermal equilibrium. Important from a physical point of view, but requires knowledge of E_{III}.
M_V : EXERCISES	

Complement A_V

Some examples of harmonic oscillators

1 Vibration of the nuclei of a diatomic molecule 527
1-a Interaction energy of two atoms 527
1-b Motion of the nuclei 529
1-c Experimental observations of nuclear vibration 531
2 Vibration of the nuclei in a crystal 534
2-a The Einstein model 534
2-b The quantum mechanical nature of crystalline vibrations . . 535
3 Torsional oscillations of a molecule: ethylene 536
3-a Structure of the ethylene molecule C_2H_4 536
3-b Classical equations of motion 538
3-c Quantum mechanical treatment 539
4 Heavy muonic atoms 541
4-a Comparison with the hydrogen atom 541
4-b The heavy muonic atom treated as a harmonic oscillator . . . 542
4-c Order of magnitude of the energies and spread of the wave functions 544

We mentioned in the introduction to Chapter V that the results obtained in the study of the harmonic oscillator are applicable to numerous cases in physics, especially those concerning small oscillations of a system about a position of stable equilibrium (where the potential energy is at a minimum). The aim of this complement is to describe some examples of such oscillations and to point out their physical importance: vibration of the nuclei in a diatomic molecule or a crystalline lattice, torsional oscillations in a molecule, motion of a muon μ^- inside a heavy nucleus. We do not intend to discuss these phenomena in great detail here. We shall confine ourselves to a simple, qualitative discussion.

1. Vibration of the nuclei of a diatomic molecule

1-a. Interaction energy of two atoms

The formation of a molecule from two neutral atoms occurs because the interaction energy $V(r)$ of these two atoms has a minimum (r is the distance between them). The form of $V(r)$ is shown in Figure 1. When r is very large, the two atoms do not interact and $V(r)$ approaches a constant which we shall choose as the energy origin. Then, as r decreases, $V(r)$ varies approximately like $-1/r^6$: the corresponding attractive forces are the Van der Waals forces (which we shall study in Complement C_{XI}). When r becomes so small that the electronic wave functions overlap, $V(r)$ decreases faster and passes through a minimum at $r = r_e$; it then increases and becomes very large as r approaches zero.

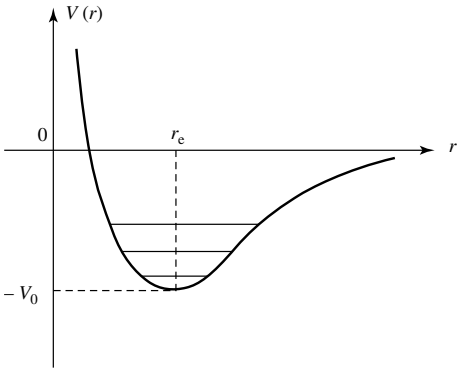


Figure 1: Form of the interaction potential between two atoms that can form a stable molecule. Classically, V_0 is the dissociation energy of the molecule and r_e, the distance between the two nuclei in the equilibrium position. In quantum mechanics, one obtains vibrational states (the horizontal lines inside the well) whose energies are all greater than $-V_0$.

The minimum of $V(r)$ is responsible for the phenomenon of the chemical bond that can form between the two atoms. We have already pointed out, in § C-2-c of Chapter IV (taking the molecule H_2^+ as an example), that the cause of this lowering of the energy is a delocalization phenomenon of the electronic states (quantum resonance) which allows the electrons to profit from the attraction of the two nuclei. The rapid rise of $V(r)$ at small distances is due to the repulsion of the nuclei.

If the nuclei were classical particles, they would have stable equilibrium positions separated by $r = r_e$. The depth V_0 of the potential well at $r = r_e$ is called, classically, the dissociation energy of the molecule: it is, in fact, the energy that must be furnished to the two atoms in order to separate them. The larger V_0, the more stable the molecule.

The theoretical and experimental determination of the curve of Figure 1 is a very important problem in atomic and molecular physics. We shall see that, by studying the vibrations of the nuclei, we get a certain amount of information about this curve.

Comment: (the Born-Oppenheimer approximation)

The quantum mechanical description of a diatomic molecule is actually a very complex problem; it involves finding the stationary states of a system of particles, the nuclei and the electrons, all interacting with each other. In general, it is impossible to solve the Schrödinger equation for such a system exactly. A significant simplification arises from the fact that the mass of the electrons is much smaller than that of the nuclei. It follows from this that one can, in a first approximation, study the two motions separately. One begins by determining the motion of the electrons for a fixed value of the distance r between the two nuclei; thus one obtains a series of stationary states for the electronic system, of energies $E_1(r)$, $E_2(r)$... Then one considers the ground state, of energy $E_1(r)$, of the electronic system; when r varies because of the motion of the nuclei, the electronic system always remains in the ground state, for all r. This means that the system's wave function adapts itself instantaneously to any change in r: the electrons, which are very mobile, are said to follow "adiabatically" the motion of the nuclei. In the study of this motion, the electronic energy $E_1(r)$ then plays the role of a potential energy of interaction between these two nuclei. This interaction potential depends on the distance between the nuclei, r, and adds to their electrostatic repulsion $Z_1 Z_2\ e^2/r$ (where Z_1 and Z_2 are the atomic numbers of the two nuclei; we have set $e^2 = q^2/4\pi\varepsilon_0$, where q is the charge of the electron). The total potential energy $V(r)$ of the system of the two nuclei, which

enables us to determine their motion, is then:

$$V(r) = E_1(r) + \frac{Z_1 Z_2\ e^2}{r} \tag{1}$$

It is this function that is shown in Figure 1.

1-b. Motion of the nuclei

α. Separation of the rotational and vibrational motions

We are thus faced with a problem that involves the motion of two particles of masses m_1 and m_2, whose interaction is described by the potential $V(r)$ of Figure 1 depending only on the distance between them. The problem is complicated by the existence of several degrees of freedom: vibrational (variation of r) and rotational (variation of the polar angles θ and φ which give the direction of the axis of the molecule). In addition, these degrees of freedom are coupled: when the molecule vibrates, its moment of inertia changes because of the variation of r, and the rotational energy is modified.

If we confine ourselves to small amplitude vibrations, it can be shown that the coupling between vibrational and rotational degrees of freedom is negligible since the relative variation of the moment of inertia is very small during the vibration. The problem is then reduced (as we shall see in detail in Complement $\mathrm{F_{VII}}$) to two independent problems: in the first place, the study of the rotation of a "dumbbell"[1] composed of two masses m_1 and m_2 separated by a fixed distance r_e; plus a one-dimensional problem (in which r is the only variable) involving a fictitious particle whose mass m is equal to the reduced mass of m_1 and m_2 (*cf.* Chap. VII, § B):

$$m = \frac{m_1 m_2}{m_1 + m_2} \tag{2}$$

moving in the potential $V(r)$ of Figure 1. We must then solve the eigenvalue equation:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + V(r)\right]\varphi(r) = E\ \varphi(r) \tag{3}$$

We shall concentrate on the latter problem here.

β. Vibrational states

If we confine ourselves to small amplitude oscillations, we can make a limited expansion of $V(r)$ in the neighborhood of its minimum, at $r = r_e$:

$$V(r) = -V_0 + \frac{1}{2}V''(r_e)(r - r_e)^2 + \frac{1}{6}V'''(r_e)(r - r_e)^3 + \ldots \tag{4}$$

The discussion in § A-2 of Chapter V shows that if we neglect higher than second-order terms in expression (4), we are left with the equation of a one-dimensional harmonic oscillator centered at $r = r_e$, of angular frequency:

$$\omega = \sqrt{\frac{V''(r_e)}{m}} \tag{5}$$

[1] We shall study this system (also called a "rigid rotator") quantum mechanically in Complement $\mathrm{C_{VI}}$, once we have introduced angular momentum.

The vibrational states $|\varphi_v\rangle$, shown by the horizontal lines in Figure 1, therefore have energies given by:

$$E_v = \left(v + \frac{1}{2}\right)\hbar\omega - V_0 \tag{6}$$

where $v = 0, 1, 2, \ldots$ (v is used instead of n in the notation of molecular vibrations).

According to the discussion in § D-3 of Chapter V, the mean value $\langle R\rangle$ of the distance between the two nuclei oscillates about r_e with a frequency of $\nu = \omega/2\pi$ which can thus be seen to be the vibrational frequency of the molecule.

Comments:

(*i*) Even in the ground state, the wave function of a harmonic oscillator has a finite spread, of the order of $\sqrt{\hbar/2m\omega}$ (*cf.* § D-2 of Chapter V). The distance between the two nuclei of the molecule in the vibrational ground state is therefore defined only to within $\sqrt{\hbar/2m\omega}$. An important condition for the decoupling of the vibrational and rotational degrees of freedom is therefore that:

$$\sqrt{\frac{\hbar}{2m\omega}} \ll r_e \tag{7}$$

(*ii*) When the reduced mass m is known, the measurement of ω yields, according to (5), the second derivative $V''(r_e)$. When the quantum number v increases, it is no longer possible to neglect terms in $(r - r_e)^3$ in expression (4) (which indicate the deviation of the potential well from a parabolic form). The oscillator then becomes anharmonic. Studying the effects of the term in $(r - r_e)^3$ of (4) by perturbation theory (as we shall do in Complement A_{XI}), one finds that the separation $E_{v+1} - E_v$ of two neighboring states is not the same for large and small values of v. Studying the variation of $E_{v+1} - E_v$ with respect to v enables us to obtain the coefficient $V'''(r_e)$ of the term in $(r - r_e)^3$. Thus we see how the study of the frequencies of molecular vibration enables us to define more precisely the form of the curve $V(r)$ in the neighborhood of its minimum.

γ. *Order of magnitude of the vibrational frequencies*

Molecular vibrational frequencies are commonly expressed in cm^{-1}, by giving the inverse of the wavelength λ (expressed in cm) of an electromagnetic wave of the same frequency ν. Note that 1 cm^{-1} corresponds to a frequency of 3×10^{10} Hertz and to an energy of 1.24×10^{-4} eV.

The vibrational frequencies of diatomic molecules fall between several tens and several thousands of cm^{-1}. The corresponding wavelengths therefore go from a few microns to a few hundred microns, consequently falling in the infrared.

Formula (5) shows that as m decreases, ω increases. This frequency also increases with $V''(r_e)$, that is, with a greater curvature of the potential well at $r = r_e$. Since r_e is always of the same order of magnitude (a few Å), $V''(r_e)$ increases with the depth V_0 of the well: ω therefore increases with the chemical stability. We shall consider some concrete illustrations of the preceding observations.

The vibrational frequencies of the hydrogen and deuterium molecules (H_2 and D_2) are, respectively (not taking into account anharmonicity corrections):

$$\begin{aligned} \nu_{H_2} &= 4\,401 \text{ cm}^{-1} \\ \nu_{D_2} &= 3\,112 \text{ cm}^{-1} \end{aligned} \tag{8}$$

The curve $V(r)$ is the same in these two cases: the chemical bond between the two atoms depends only on the electronic atmosphere. However, the reduced mass of H_2 is half as large as that of D_2. We must therefore have, according to (5), $\nu_{H_2} = \sqrt{2}\ \nu_{D_2}$. This is in agreement with the experimental values (8).

Now let us consider an example of two molecules that have about the same reduced mass but very different chemical stabilities. The molecule $^{79}Br\ ^{85}Rb$ is chemically stable (halogen-alkaline bond); its vibrational frequency is 181 cm^{-1}. Molecules of $^{84}Kr^{85}Rb$ have been observed recently in optical pumping experiments. Their chemical stability is much lower, because krypton, which is a rare gas, is practically inert from a chemical point of view (in fact, the cohesion of the molecule is due only to Van der Waals forces). These molecules have been found to have a vibrational frequency of the order of 13 cm^{-1}. The considerable difference between this figure and the preceding one is due solely to the difference in chemical stability of the two types of molecules since the reduced masses are, to within a few per cent, practically the same.

1-c. Experimental observations of nuclear vibration

We shall now explain how nuclear vibration can be detected experimentally. In particular, we shall consider the interaction of the molecule with an electromagnetic wave.

α. *Infrared absorption and emission*

First, let us assume the molecule to be *heteropolar* (composed of two different atoms). Since the electrons are attracted towards the more electronegative atom, the molecule generally has a permanent dipole moment $D(r)$ which depends on the distance r between the two nuclei. Expanding $D(r)$ in the neighborhood of the equilibrium position $r = r_e$, we obtain:

$$D(r) = d_0 + d_1(r - r_e) + ... \tag{9}$$

where d_0 and d_1 are real constants.

When the molecule is in a linear superposition $|\psi(t)\rangle$ of several stationary vibrational states $|\varphi_v\rangle$, the mean value $\langle\psi(t)|D(R)|\psi(t)\rangle$ of its electric dipole moment oscillates about the value d_0 with a frequency of $\omega/2\pi$. The oscillatory term arises from the mean value of the term $d_1(R - r_e)$ of (9) ($R - r_e$ plays the same role in our problem as the observable X of the harmonic oscillator studied in § D-3 of Chapter V). Now $(R - r_e)$ has a non-zero matrix element between two states $|\varphi_v\rangle$ and $|\varphi_{v'}\rangle$ only when $v - v' = \pm 1$. This selection rule enables us to understand why only one Bohr frequency $\nu = \omega/2\pi$ appears in the motion of $\langle D(R)\rangle(t)$ [the harmonic frequencies evidently appear when one takes into account the anharmonicity of the potential and terms of higher order in expansion (9); their intensity is however much weaker].

This vibration of the electric dipole moment results in a coupling between the molecule and the electromagnetic field; the molecule can consequently absorb or emit radiation of frequency ν. In terms of photons, the molecule can absorb a photon of energy $h\nu$ and move from the state $|\varphi_v\rangle$ to the state $|\varphi_{v+1}\rangle$ (Fig. 2-a) or emit a photon $h\nu$ by going from $|\varphi_v\rangle$ to $|\varphi_{v-1}\rangle$ (Fig. 2-b).

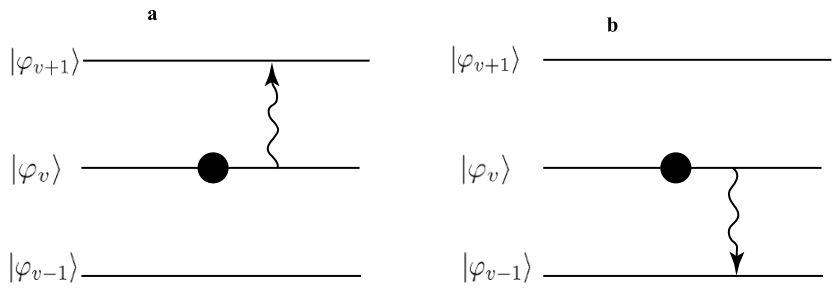


Figure 2: Absorption (fig. a) or emission (fig. b) of a photon by a heteropolar molecule going from the vibrational state $|\varphi_v\rangle$ to the state $|\varphi_{v+1}\rangle$ or $|\varphi_{v-1}\rangle$.

β. The Raman effect

Now let us consider a *homopolar* molecule (consisting of two identical atoms). Because of symmetry, the permanent electric dipole moment is then zero for all r, and the molecule is "inactive" in the infrared.

Imagine that an optical wave of frequency $\Omega/2\pi$ strikes this molecule. This frequency, much higher than those considered previously, is able to excite the electrons of the molecule; under the effect of the optical wave, the electrons will undergo forced oscillation and re-emit radiation of the same frequency in all directions. This is the well-known phenomenon of the molecular scattering of light (Rayleigh scattering)[2]. What new phenomena are produced by the vibration of the molecule?

What happens can be explained qualitatively in the following way. The electronic susceptibility[3] of the molecule is generally a function of the distance r between the two nuclei. When r varies (recall that this variation is slow compared to the motion of the electrons), the amplitude of the *induced* electric dipole moment, which vibrates at a frequency of $\Omega/2\pi$, varies. The time dependence of the dipole moment is therefore that of a sinusoid of frequency $\Omega/2\pi$ whose amplitude is modulated at the frequency of the molecular vibration $\omega/2\pi$, which is much smaller (Fig. 3). The frequency distribution of the light emitted by the molecule is given by the Fourier transform of the motion of the electric dipole shown in Figure 3. It is easy to see (Fig. 4) that there exists a central line of frequency $\Omega/2\pi$ (Rayleigh scattering) and two shifted lines, of frequency $(\Omega-\omega)/2\pi$ (Stokes Raman scattering) and frequency $(\Omega+\omega)/2\pi$ (anti-Stokes Raman scattering).

It is very simple to interpret these lines in terms of photons. Consider an optical photon of energy $\hbar\Omega$ which strikes the molecule when it is in the state $|\varphi_v\rangle$ (Fig. 5-a). If the molecule does not change vibrational states during the scattering process, the scattering is elastic. Because of conservation of energy, the scattered photon has the

[2] In Complement A_{XIII}, we shall use quantum mechanics to study the forced motion of the electrons of an atom under the effect of incident light waves.

[3] Under the effect of the field $\mathbf{E}_0\, e^{i\Omega t}$ of the incident optical wave, the electronic cloud of the molecule acquires an induced dipole moment $\mathbf{D}$ given by:

$$\mathbf{D} = \chi(\Omega)\mathbf{E}_0\, e^{i\Omega t}$$

$\chi(\Omega)$ is, by definition, the electronic susceptibility of the molecule. The important point here is that χ depends on r.

same energy as the incident photon (Fig. 5-b: Rayleigh line). However, the molecule, during the scattering process, can make a transition from the state $|\varphi_v\rangle$ to the state $|\varphi_{v+1}\rangle$. The molecule acquires an energy $\hbar\omega$ at the expense of the scattered photon, whose energy therefore is $\hbar(\Omega - \omega)$ (Fig. 5-c): the scattering is *inelastic* (Stokes Raman line). Finally, the molecule may move from the state $|\varphi_v\rangle$ to the state $|\varphi_{v-1}\rangle$, in which case the scattered photon will have an energy of $\hbar(\Omega + \omega)$ (Fig. 5-d: anti-Stokes Raman line).

Comments:

(*i*) The Raman effect can also be observed with heteropolar molecules.

(*ii*) The Raman effect has enjoyed a revival of interest because of the development of lasers. If, in the cavity of a laser oscillating at a frequency of $\Omega/2\pi$, one places a cell filled with a substance that exhibits the Raman effect, one can, in certain cases, obtain an amplification (stimulated Raman effect) and hence a laser oscillation at the frequency $(\Omega - \omega)/2\pi$, where ω is the vibrational frequency of the molecules in the cell (Raman laser). Thus, by varying this substance, one can vary the oscillation frequency of the laser.

(*iii*) The study of Raman and infrared spectra of molecules is useful in chemistry because it permits the identification of the various bonds which exist in a complex molecule.

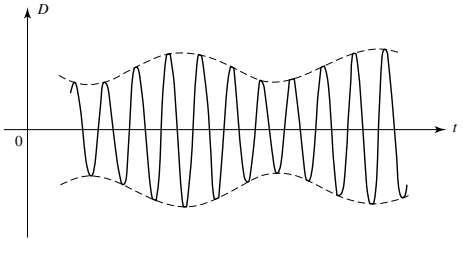


Figure 3: The vibration of a molecule modulates the amplitude of the oscillating electric dipole induced by an incident light wave.

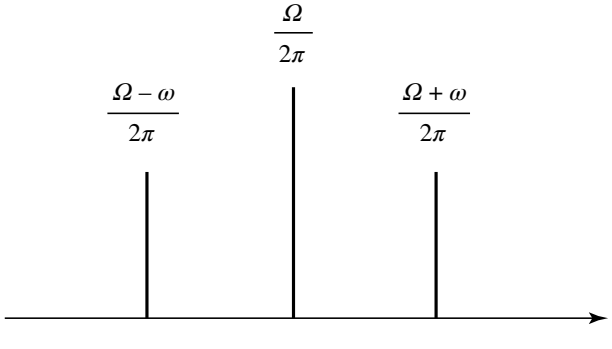


Figure 4: Spectrum of the oscillations shown in Figure 3. In addition to the central line, whose frequency is the same as that of the incident light wave (Rayleigh line), two shifted lines appear (the Stokes and anti-Stokes Raman lines). The frequency shift is equal to the vibrational frequency of the molecule.

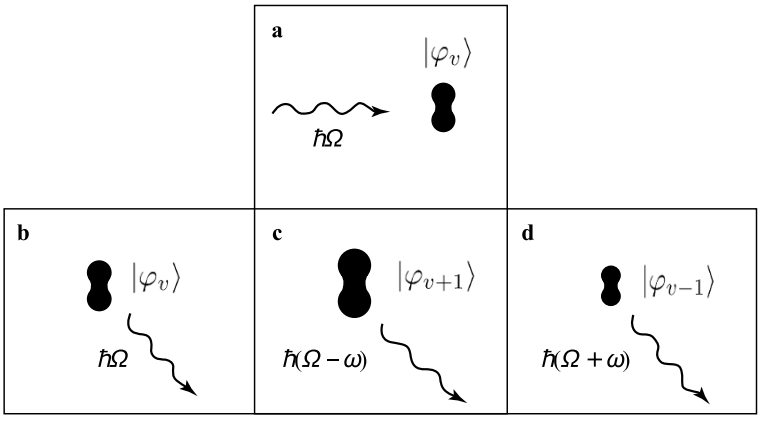


Figure 5: Schematic representation of the scattering of a photon of energy $\hbar\Omega$ by a molecule which is initially in the vibrational state $|\varphi_v\rangle$ (fig. a): Rayleigh scattering without a change in the vibrational state (fig. b); Stokes or anti-Stokes Raman scattering with a change in the molecule's state from $|\varphi_v\rangle$ to $|\varphi_{v+1}\rangle$ (fig. c) or to $|\varphi_{v-1}\rangle$ (fig. d).

For example, the vibration frequency of a group of two carbon atoms depends on whether the bond between them is single, double or triple.

2. Vibration of the nuclei in a crystal

2-a. The Einstein model

A crystal consists of a system of atoms (or ions) which are regularly distributed in space, forming a periodic lattice. For simplicity, let us choose a one-dimensional model in which we consider a linear chain of atoms.

The average position of the nucleus of the qth atom is:

$$x_q^0 = q\ d \tag{10}$$

where d is the distance between adjacent atoms (on the order of a few Å).

Let $U(x_1,\ x_2,\ ...,\ x_q,\ ...)$ be the total potential energy of the crystal nuclei, which depends on their positions $x_1,\ x_2,\ ...,\ x_q...$ If $x_q - x_q^0$ is not too large, that is, if each nucleus is not too far from its equilibrium position, $U(x_1,\ x_2,\ ...,\ x_q,\ ...)$ has, in certain cases, the following simple form:

$$U(x_1,\ x_2,\ ...,\ x_q,\ ...) \simeq U_0 + \sum_q \frac{1}{2}(x_q - x_q^0)^2 U_0'' + ... \tag{11}$$

where U_0 and U_0'' are real constants (with $U_0'' > 0$). The absence of terms linear in $x_q - x_q^0$ shows that x_q^0 is a stable equilibrium position for the nucleus (q) (a minimum of U). We add to (11) the total kinetic energy:

$$T = \sum_q \frac{p_q^2}{2m} \tag{12}$$

where p_q is the momentum of the nucleus (q) of mass m. The total Hamiltonian H of the system is, to within the constant U_0, a sum of Hamiltonians of one-dimensional harmonic oscillators centered at each nucleus (q):

$$H = U_0 + \sum_q \left[\frac{p_q^2}{2m} + \frac{1}{2}(x_q - x_q^0)^2 U_0'' \right] \tag{13}$$

Consequently, in this simplified model, each nucleus vibrates about its equilibrium position independently of its neighbors, with an angular frequency:

$$\omega = \sqrt{\frac{U_0''}{m}} \tag{14}$$

As in the case of the diatomic molecule, ω increases when m decreases and when the curvature of the potential attracting the nucleus towards its equilibrium position increases.

Comment:

In the simple model we have just presented, each nucleus vibrates independently of the others. This is because the proposed potential U does not contain any terms which are simultaneously dependent on more than one of the variables x_q, as it would if it described internuclear interactions. This model is not realistic since such interactions do, in fact, exist. In Complement J_V, we shall present a more elaborate model which takes into account the coupling between each nucleus and its two nearest neighbors. We shall see that it is still possible, in this model, to put the total Hamiltonian of the system in the form of a sum of Hamiltonians of independent harmonic oscillators.

2-b. The quantum mechanical nature of crystalline vibrations

Despite its very schematic character, the Einstein model enables us to understand a certain number of phenomena related to the quantum mechanical nature of crystalline vibrations. The low temperature behavior of the constant volume specific heat, which cannot be explained using classical mechanics, will be described in Complement L_V in connection with the study of the properties of a harmonic oscillator in thermodynamic equilibrium. In the present complement, we shall discuss a spectacular effect related to the finite spread of the wave functions associated with the position of each atom in the ground state.

At absolute zero, under a pressure of one atmosphere, all substances except helium are solids. To solidify helium, it is necessary to apply a pressure of at least 25 atmospheres. Can this peculiarity be explained qualitatively?

First let us try to understand the phenomenon of the melting of an ordinary substance. At absolute zero, the atoms are practically localized at their equilibrium positions; the spread of their wave functions about the x_q^0 is given by [*cf.* formula (D-5a) of Chapter V]:

$$\Delta X \simeq \sqrt{\frac{\hbar}{2m\omega}} = \left[\frac{\hbar^2}{4mU_0''} \right]^{1/4} \tag{15}$$

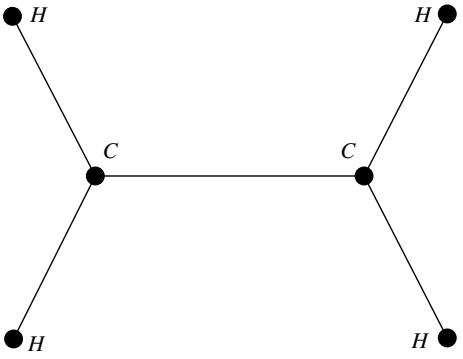


Figure 6: Plane structure of the ethylene molecule.

[where we have used expression (14) for ω]. ΔX is, in general, very small. When the crystal is heated, the nuclei move into higher and higher vibrational states: in classical language, they vibrate with a larger and larger amplitude; in quantum mechanical language, the spread of their wave functions increases [with the square root of the vibrational quantum number – see formula (D-5a) of Chapter V]. When this spread is no longer negligible with respect to the interatomic distance d, the crystal melts (see § 4-c of Complement L_V, in which this phenomenon is studied more quantitatively).

It is impossible to solidify helium at ordinary pressures. This corresponds to the fact that, even at absolute zero, the spread of the wave function given by (15) is not negligible compared to d. This results from the fact that the mass of helium is very small and its chemical affinity, very weak (the curvature U_0'' of the potential in the neighborhood of each minimum is very small, since the potential wells are very shallow). Both factors produce the same effect in formula (15): a large spread ΔX. Now, an increase in the pressure results in an increase in U_0'' and therefore, in ω; consequently, ΔX decreases. This is due to the fact that, at high pressures, each helium atom is "wedged" between its neighbors: the smaller the average distance between these neighbors (the higher the pressure), the sharper the potential minimum (the greater U_0''). Thus we see how an increase in the pressure makes the solidification of helium possible.

3. Torsional oscillations of a molecule: ethylene

3-a. Structure of the ethylene molecule C_2H_4

The structure of the molecule C_2H_4 is well-known: the six atoms of the molecule are in the same plane (Fig. 6) and the angles between the various $C-H$ and $C-C$ bonds are close to 120°.

Now imagine that, without changing the relative positions of the bonds of each carbon atom, we rotate one of the CH_2 groups, about the $C-C$ axis, through an angle α with respect to the other one. Figure 7 represents the molecule as seen along the $C-C$ axis: the $C-H$ bonds of one CH_2 group are shown in solid lines and those of the other one, in dashed lines. How does the potential energy $V(\alpha)$ of the molecule vary with respect to α?

Since the stable structure of the molecule is planar, the angle $\alpha = 0$ must correspond to a minimum of $V(\alpha)$. It is also clear that $\alpha = \pi$ corresponds to another minimum

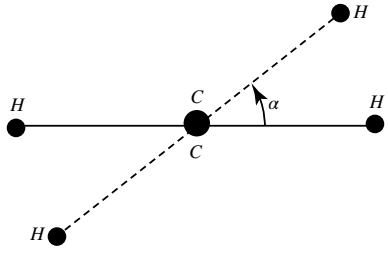


Figure 7: Torsion of the ethylene molecule (seen along the $C-C$ *axis): one of the* CH_2 *groups has rotated with respect to the other one through an angle* α *about the* $C-C$ *axis.*

of $V(\alpha)$, since the two structures associated with $\alpha = 0$ and $\alpha = \pi$ are undistinguishable. $V(\alpha)$ therefore has the form shown in Figure 8 [α varies from $-\pi/2$ to $3\pi/2$; $V(0)$ is chosen as the energy origin].

The two stable positions $\alpha = 0$ and $\alpha = \pi$ are separated by a potential barrier of height V_0. The potential of Figure 8 is often approximated by the simple formula:

$$V(\alpha) = \frac{V_0}{2}(1 - \cos 2\alpha) \tag{16}$$

Comment:

Quantum mechanics enables us to interpret all the features of the C_2H_4 molecule which we have just described. In this molecule, each carbon atom has four valence electrons. Three of these electrons (σ electrons) are found to have wave functions that are symmetrical about three coplanar lines making angles of 120° with each other, and defining the directions of the chemical bonds (Fig. 6). These wave functions overlap those of the electrons of the neighboring atoms to a considerable extent, and it is this overlap that insures the stability of the $C-H$ bonds and of part of the $C-C$ bond (this phenomenon is called "sp^2 hybridation" and will be studied in greater detail in Complement E_{VII}). The last valence electron of each carbon atom (π electron) has a wave function which is symmetrical about a line passing through C and perpendicular to the plane defined by C and its three neighbors. The overlap of the wave functions of the two π electrons is maximum and, consequently, the chemical stability of the double bond is greatest when

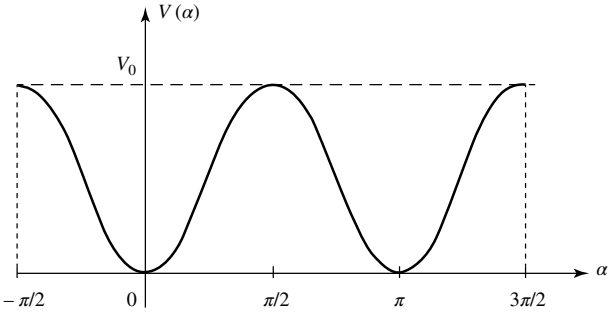


Figure 8: The potential energy of the molecule depends on the torsion angle α*;* $V(\alpha)$ *is minimal for* $\alpha = 0$ *and* $\alpha = \pi$ *(planar structures).*

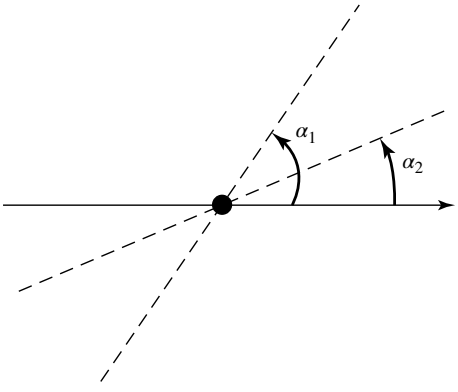


Figure 9: To write the classical equations of motion, we denote by α_1 and α_2 the angles formed by the planes of the two CH_2 groups with a fixed plane.

the two lines associated with the π electrons are parallel, that is, when the six atoms of the molecules are in the same plane. The structure of Figure 6 is thus entirely explained.

Since $V(\alpha)$ can be approximated by a parabola in the neighborhood of each of its two minima, the molecule performs torsional oscillations about its two stable equilibrium positions. We now examine them. First, we shall review rapidly the corresponding classical equations.

3-b. Classical equations of motion

We denote by α_1 and α_2 the angles formed by the planes of the two CH_2 groups with a fixed plane passing through the C – C axis (Fig. 9). The angle in Figure 7 is obviously:

$$\alpha = \alpha_1 - \alpha_2 \tag{17}$$

Let I be the moment of inertia of one of the CH_2 groups with respect to the C – C axis. Since the potential energy depends only on $\alpha = \alpha_1 - \alpha_2$, the dynamical equations describing the rotation of each group are written:

$$\begin{cases} I\ \dfrac{\mathrm{d}^2\alpha_1}{\mathrm{d}t^2} = -\dfrac{\partial}{\partial\alpha_1}V(\alpha_1-\alpha_2) = -\dfrac{\mathrm{d}}{\mathrm{d}\alpha}V(\alpha) \\ I\ \dfrac{\mathrm{d}^2\alpha_2}{\mathrm{d}t^2} = -\dfrac{\partial}{\partial\alpha_2}V(\alpha_1-\alpha_2) = +\dfrac{\mathrm{d}}{\mathrm{d}\alpha}V(\alpha) \end{cases} \tag{18}$$

Adding and subtracting these two equations we obtain:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\alpha_1+\alpha_2) = 0 \tag{19a}$$

$$I\ \frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} = -2\ \frac{\mathrm{d}}{\mathrm{d}\alpha}V(\alpha) \tag{19b}$$

Equation (19a) indicates that the entire molecule can rotate freely about the C – C axis independently of the torsional motion: the angle $(\alpha_1 + \alpha_2)/2$ of the plane bisecting the planes of the two CH_2 groups is a linear function of time. Equation (19b) describes the torsional motion (rotation of one group with respect to the other). Let us consider this motion in the immediate neighborhood of one of the stable equilibrium positions, $\alpha = 0$. We expand expression (16) in the neighborhood of $\alpha = 0$:

$$V(\alpha) \simeq V_0\alpha^2 \tag{20}$$

Substituting (20) into (19b), we obtain:

$$\frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} + \frac{4V_0}{I}\alpha = 0 \tag{21}$$

We recognize (21) as the equation of a one-dimensional harmonic oscillator (α is the only variable) of angular frequency:

$$\omega_t = 2\sqrt{\frac{V_0}{I}} \tag{22}$$

For the C_2H_4 molecule, ω_t is of the order of 825 cm^{-1}.

3-c. Quantum mechanical treatment

In the neighborhood of its two equilibrium positions $\alpha = 0$ and $\alpha = \pi$, the molecule possesses "torsional states" of quantized energy $E_n = (n+1/2)\hbar\omega_t$, with $n = 0, 1, 2, \ldots$ In a first approximation, each energy level $E_n = (n+1/2)\hbar\omega_t$ is therefore doubly degenerate, since for each one there are two states $|\varphi_n\rangle$ and $|\varphi'_n\rangle$ whose wave functions $\varphi_n(\alpha)$ and $\varphi'_n(\alpha)$ differ only in that one is centered at $\alpha = 0$ and the other, at $\alpha = \pi$ (Fig. 10-a and 10-b).

In fact, we must also take into account a typically quantum mechanical effect: the tunnel effect across the potential barrier separating the two minima (Fig. 8). We have already encountered a situation of this type, in Complement G_{IV}, in connection with the inversion of the NH_3 molecule. Calculations analogous to those in that complement could show that the degeneracy between the two states $|\varphi_n\rangle$ and $|\varphi'_n\rangle$ is removed by the tunnel effect. Thus, for each value of n, two stationary states, $|\psi^n_+\rangle$ and $|\psi^n_-\rangle$, appear (to a first approximation, they are symmetrical and antisymmetrical linear combinations of $|\varphi_n\rangle$ and $|\varphi'_n\rangle$). The larger n (that is, the closer the initial energy E_n is to V_0, and hence, the more important the tunnel effect), the greater their energy difference $\hbar\delta_n$. However, $\hbar\delta_n$ is always much smaller than the distance $\hbar\omega_t$ between adjacent levels n and $n \pm 1$ (Fig. 11).

For the mean value of the angle α, quantum mechanics therefore predicts the following motion: rapid oscillations of frequency ω_t about one of the two values $\alpha = 0$ and $\alpha = \pi$, upon which are superposed much slower oscillations between $\alpha = 0$ and $\alpha = \pi$, at the Bohr frequencies $\delta_0/2\pi$, $\delta_1/2\pi$, $\delta_2/2\pi$...

Comment:

States of course exist for which the energy is greater than the maximum height V_0 of the potential barrier of Figure 8. These states correspond to a rotational kinetic energy

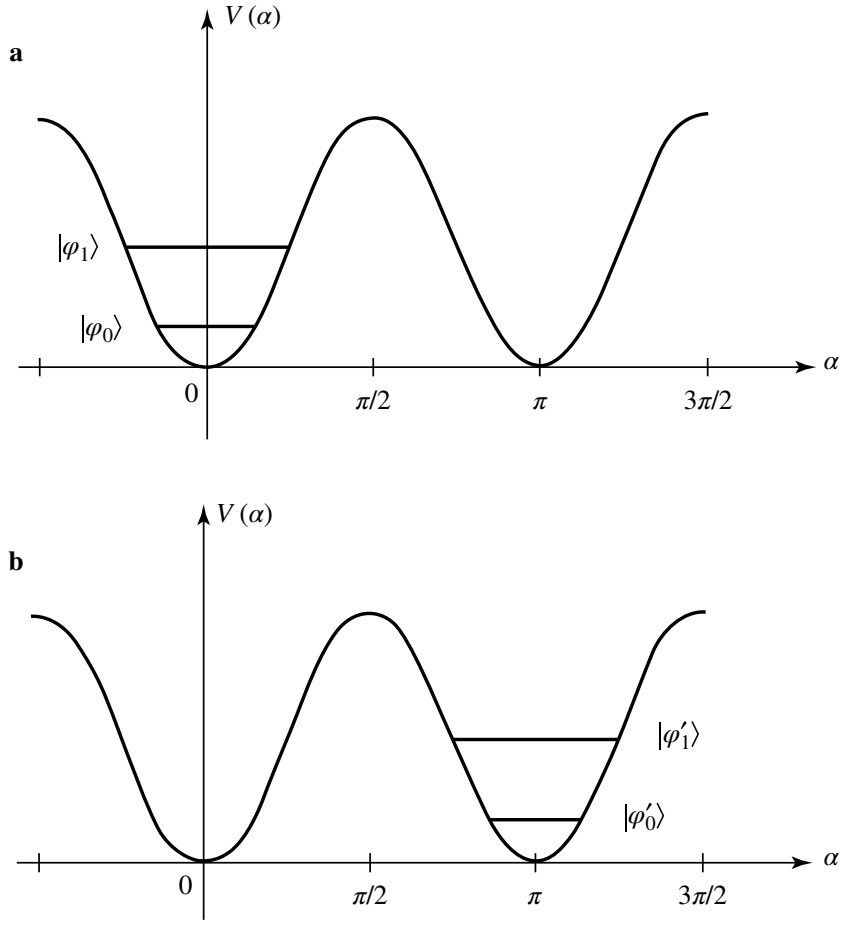


Figure 10: When one neglects the tunnel effect across the potential barriers at $\alpha = \pi/2$ and $\alpha = 3\pi/2$, one can find torsional states of the molecule localized in the wells centered at $\alpha = 0$ (fig. a) and $\alpha = \pi$ (fig. b).

that is large enough for one of the CH_2 groups to be considered as rotating almost freely with respect to the other one (while being, nevertheless, periodically slowed down and accelerated by the potential of Figure 8).

The ethane molecule C_2H_6 behaves in this way. The absence of π electrons in this molecule permits a much freer rotation of one of the CH_3 groups with respect to the other (the potential barrier V_0 is much lower). In this case, the potential $V(\alpha)$, which tends to oppose the free rotation of one of the CH_3 groups with respect to the other, has a period of $2\pi/3$ because of symmetry.

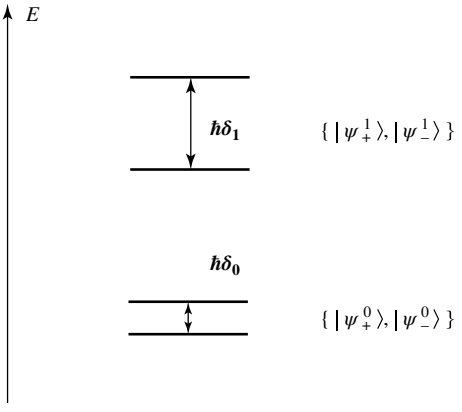


Figure 11: The tunnel effect removes the degeneracy of the energy levels shown in Figure 10. As one approaches the top of the barrier, this phenomenon becomes more important ($\delta_1 > \delta_0$). $|\psi_+^0\rangle$, $|\psi_-^0\rangle$, $|\psi_+^1\rangle$, $|\psi_-^1\rangle$ are the new stationary states.

4. Heavy muonic atoms

The muon μ^- (sometimes called, for historical reasons, the "μ meson") is a particle which has the same properties as the electron except that its mass is about 207 times greater[4]. In particular, it is not sensitive to strong interactions, and its coupling with nuclei is essentially electromagnetic. A muon μ^- which has been slowed down in matter can be attracted by the Coulomb field of an atomic nucleus and can form a bound state with that nucleus. The system thus constituted is called a muonic atom.

4-a. Comparison with the hydrogen atom

In Chapter VII (§ C), we shall study the bound states of two particles of opposite charge and, in particular, those of the hydrogen atom. We shall see that the results of quantum mechanics concerning the energies of bound states are the same as those of the Bohr model (Chap. VII, § C-2). Similarly, the spread of the wave functions which describe these bound states is of the order of the Bohr orbital radius. Let us therefore begin by using this simple model to calculate the energies and spreads of the first bound states of a muon μ^- in the Coulomb field of a heavy atom such as lead ($Z = 82$, $A = 207$).

If we consider the nucleus to be infinitely heavy, the nth Bohr orbital has an energy of:

$$E_n = -\frac{Z^2 m e^4}{2\hbar^2}\frac{1}{n^2} \tag{23}$$

where Z is the atomic number of the nucleus, $e^2 = q^2/4\pi\varepsilon_0$ (where q is the electron charge), and m represents the mass of the electron or of the muon, depending on the case. When one goes from hydrogen to the muonic atom under study here, E_n is multiplied by a factor of $Z^2 m_\mu/m_e \simeq (82)^2 \times 207 \simeq 1.4 \times 10^6$. From this we deduce that, for the

[4]The muon is unstable: it decays into an electron and two neutrinos.

muonic atom:

$$\begin{cases} E_1 \simeq -19 \text{ MeV} \\ \\ E_2 \simeq -4.7 \text{ MeV} \end{cases} \tag{24}$$

The radius of the nth Bohr orbital is given by:

$$r_n = -\frac{n^2\hbar^2}{Zme^2} \tag{25}$$

For hydrogen, $r_1 \simeq 0.5$ Å. Here, this number must be divided by Zm_μ/m_e, which gives:

$$\begin{cases} r_1 \simeq 3 \times 10^{-13} \text{ cm} \\ \\ r_2 \simeq 12 \times 10^{-13} \text{ cm} \end{cases} \tag{26}$$

In the preceding calculations, we have implicitly assumed the nucleus to be point-like (in the Bohr model and in the theory presented in Chapter VII, § C, the potential energy is taken equal to $-Z\ e^2/r$). The small values found for r_1 and r_2 [formulas (26)] show us that this viewpoint is not at all valid for a heavy muonic atom. The lead nucleus has a non-negligible radius ρ_0, on the order of 8.5×10^{-13} cm (recall that the radius of a nucleus increases with $A^{1/3}$). The preceding qualitative calculation therefore leaves us with the impression that the spread of the wave functions of the muon may be smaller than the nucleus[5]. Consequently, we must reconsider the problem completely and first calculate the potential "seen" by the muon on the *inside* as well as on the outside of the nuclear charge distribution.

4-b. The heavy muonic atom treated as a harmonic oscillator

We shall use a rough model of the lead nucleus: we shall assume its charge to be evenly distributed throughout a sphere of radius $\rho_0 \simeq 8.5 \times 10^{-13}$ cm.

When the distance r of the muon from the center of this sphere is greater than ρ_0, its potential energy is given by:

$$V(r) = -\frac{Ze^2}{r} \quad \text{for} \quad r \geq \rho_0 \tag{27}$$

For $r < \rho_0$, one can calculate the electrostatic force acting on the muon, using Gauss's theorem; it is directed towards the center of the sphere and its absolute value is:

$$Ze^2 \left(\frac{r}{\rho_0}\right)^3 \frac{1}{r^2} = \frac{Ze^2}{\rho_0^3} r \tag{28}$$

[5]For hydrogen, the spread of the wave functions, on the order of an Angström, is about 10^5 times larger than the dimensions of the proton, which can therefore be treated like a point. The new situation encountered here results from several factors which reinforce each other: increased m and increased Z, which results in a greater electrostatic force and a larger nuclear radius.

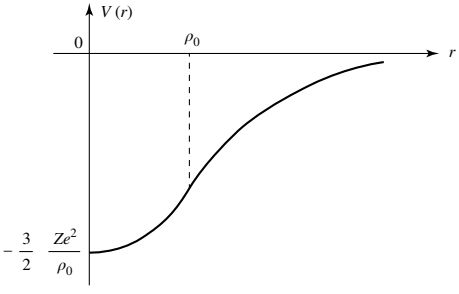


Figure 12: Form of the potential $V(r)$ seen by a muon μ^- attracted by a nucleus of radius ρ_0 situated at $r = 0$. When $r < \rho_0$, the variation of the potential is parabolic (if the charge density of the nucleus is uniform); when $r > \rho_0$, $V(r)$ varies like $1/r$ (Coulomb's law).

This force is derived from the potential energy:

$$V(r) = \frac{1}{2}\frac{Ze^2}{\rho_0^3}r^2 + C \quad \text{for} \quad r \leq \rho_0 \tag{29}$$

The constant C is determined by the condition that expressions (27) and (29) be identical for $r = \rho_0$:

$$C = -\frac{3}{2}\frac{Ze^2}{\rho_0} \tag{30}$$

Figure 12 represents the potential energy of the muon, plotted with respect to r.

Inside the nucleus, the potential is parabolic. The orders of magnitude we calculated in § 4-a indicate that it would not be realistic to choose a pure Coulomb potential for the ground state of the muonic lead atom since the wave function is actually concentrated in the region where the potential is parabolic. It is therefore certainly preferable to consider the muon to be "elastically bound" to the nucleus in this case. We then have a three-dimensional harmonic oscillator (Complement E_V) whose angular frequency is:

$$\omega = \sqrt{\frac{Ze^2}{m_\mu \rho_0^3}} \tag{31}$$

In fact, we shall see that the wave function of the ground state of this harmonic oscillator is not zero outside the nucleus, so the harmonic approximation is not perfect either.

Comment:

It is interesting that the physical system studied here presents many analogies with the first atomic model, proposed by J. J. Thomson. This physicist assumed the positive charge of the atom to be distributed in a sphere whose radius was of the order of a few Angströms, with the electrons moving in the parabolic potential existing inside this charge distribution (model of the elastically bound electron). We know from Rutherford's experiments that the nucleus is much smaller and that such a model does not correspond to reality for atoms.

4-c. Order of magnitude of the energies and spread of the wave functions

If we substitute into expression (31) the numerical values:

$$Z = 82 \qquad c \simeq 3 \times 10^8 \text{ m/sec}$$

$$\frac{e^2}{\hbar c} \simeq \frac{1}{137} \qquad m_\mu \simeq 207\ m_e \simeq 1.86 \times 10^{-28} \text{ kg}$$

$$\hbar \simeq 1.05 \times 10^{-34} \text{ Joule sec} \qquad \rho_0 \simeq 8.5 \times 10^{-15} \text{ m}$$

we find:

$$\omega \simeq 1.3 \times 10^{22} \text{ rad/sec} \tag{32}$$

which corresponds to an energy $\hbar\omega$ on the order of:

$$\hbar\omega \simeq 8.4 \text{ MeV} \tag{33}$$

We can compare $\hbar\omega$ to the total depth of the well $\frac{3}{2}\frac{Ze^2}{\rho_0}$, which is equal to:

$$\frac{3}{2}\frac{Ze^2}{\rho_0} \simeq 21 \text{ MeV} \tag{34}$$

We see that $\hbar\omega$ is smaller than this depth, but not small enough for us to be able to neglect completely the non-parabolic part of $V(r)$.

Similarly, the spread of the ground state, if the well were perfectly parabolic, would be on the order of:

$$\sqrt{\frac{\hbar}{2m_\mu\omega}} \simeq 4.7 \times 10^{-13} \text{ cm} \tag{35}$$

The qualitative predictions of § 4-a are therefore confirmed: a large part of the wave function of the muon is inside the nucleus. Nevertheless, what happens outside cannot be completely neglected.

The exact calculation of the energies and the wave functions is therefore more complicated than it would be for a simple harmonic oscillator. The Schrödinger equation corresponding to the potential of Figure 12 must be solved (taking into account, in addition, spin, relativistic corrections, etc...). Such a calculation is important: the study of the energy of photons emitted by a heavy muonic atom contributes information about the structure of the nucleus, for example concerning the real charge distribution inside the nuclear volume.

Comment:

In the case of ordinary atoms (with an electron instead of a muon), it is valid to neglect the effects of the deviation of the potential from the $-Ze^2/r$ law. However, one can take this deviation into account by using perturbation theory (*cf.* Chap. XI). In Complement D_{XI}, we shall study this "volume effect" of the nucleus on the atomic energy levels.

References and suggestions for further reading:

Molecular vibrations: Karplus and Porter (12.1), Chap. 7; Pauling and Wilson (1.9), Chap. X; Herzberg (12.4), Vol. I, Chap. III, § 1; Landau and Lifshitz (1.19), Chaps. XI and XIII.

Stimulated Raman effect: Baldwin (15.19), § 5.2; see also Schawlow's article (15.17).

Torsion oscillations: Herzberg (12.4), Vol. II, Chap. II, § 5d; Kondratiev (11.6), § 37

The Einstein model: Kittel (13.2), Chap. 6; Seitz (13.4), Chap. III; Ziman (13.3), Chap. 2; see also Bertman and Guyer's article (13.20).

Muonic atoms: Cagnac and Pebay-Peyroula (11.2), § XIX-7; Weissenberg (16.19), § 4-2; see also De Benedetti's article (11.21).

Complement B_V

Study of the stationary states in the $\{|x\rangle\}$ representation. Hermite polynomials

1 Hermite polynomials 547
1-a Definition and simple properties 547
1-b Generating function 548
1-c Recurrence relations; differential equation 549
1-d Examples 550
2 The eigenfunctions of the harmonic oscillator Hamiltonian 550
2-a Generating function 550
2-b $\varphi_n(x)$ in terms of the Hermite polynomials 551
2-c Recurrence relations 552

We now intend to study, in a little more detail than in § C-2 of Chapter V, the wave functions $\varphi_n(x) = \langle x|\varphi_n\rangle$ associated with the stationary states $|\varphi_n\rangle$ of the harmonic oscillator. Before undertaking this study, we shall define the Hermite polynomials and mention their principal properties.

1. Hermite polynomials

1-a. Definition and simple properties

Consider the Gaussian function:

$$F(z) = \mathrm{e}^{-z^2} \tag{1}$$

represented by the bell-shaped curve in Figure 1. The successive derivatives of F are given by:

$$F'(z) = -2z\ \mathrm{e}^{-z^2} \tag{2}$$

$$F''(z) = (4z^2 - 2)\ \mathrm{e}^{-z^2} \tag{3}$$

The nth-order derivative, $F^{(n)}(z)$, can be written:

$$F^{(n)}(z) = (-1)^n\, H_n(z)\ \mathrm{e}^{-z^2} \tag{4}$$

where $H_n(z)$ is an nth-degree polynomial in z. The proof is by recurrence. This relation is valid for $n = 1$, 2 [*cf.* equations (2) and (3)]. Assume it is true for $n - 1$:

$$F^{(n-1)}(z) = (-1)^{n-1}\, H_{n-1}(z)\ \mathrm{e}^{-z^2} \tag{5}$$

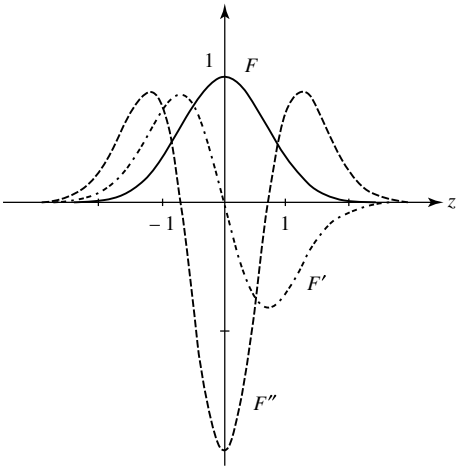


Figure 1: Shape of the Gaussian function $F(z)$ and of its first and second derivatives $F'(z)$ and $F''(z)$.

where $H_{n-1}(z)$ is a polynomial of degree $n-1$. We then obtain relation (4) directly by differentiation, if we set:

$$H_n(z) = \left(2z - \frac{\mathrm{d}}{\mathrm{d}z}\right) H_{n-1}(z) \tag{6}$$

Since $H_{n-1}(z)$ is a polynomial of degree $n-1$ in z, we see from this last relation that $H_n(z)$ is indeed an nth-degree polynomial. The polynomial $H_n(z)$ is called the *nth-degree Hermite polynomial*. Its definition is therefore:

$$H_n(z) = (-1)^n \, \mathrm{e}^{z^2} \frac{\mathrm{d}^n}{\mathrm{d}z^n} \mathrm{e}^{-z^2} \tag{7}$$

We see from (2) and (3) that $H_1(z)$ and $H_2(z)$ are, respectively, even and odd. Moreover, relation (6) shows that if $H_{n-1}(z)$ has a definite parity, $H_n(z)$ has the opposite parity. From this, we deduce that the parity of $H_n(z)$ is $(-1)^n$.

The zeros of $H_n(z)$ correspond to those of the nth-order derivative of the function $F(z)$. We are going to show that $H_n(z)$ has n real zeros, between which one finds those of H_{n-1}. It can be seen from Figure 1 and from relations (1), (2) and (3) that this is true for $n = 0, 1, 2$. Arguing by recurrence, we can generalize this result: assume that $H_{n-1}(z)$ has $n-1$ real zeros; if z_1 and z_2 are two consecutive zeros of $H_{n-1}(z)$ and therefore of $F^{(n-1)}(z)$, Rolle's theorem shows that the derivative $F^{(n)}(z)$ of $F^{(n-1)}(z)$ goes to zero at a point z_3 between z_1 and z_2; therefore, $H_n(z_3) = 0$. Since, in addition, $F^{(n-1)}(z)$ goes to zero when $z \longrightarrow -\infty$ and when $z \longrightarrow +\infty$, $F^{(n)}(z)$ and $H_n(z)$ have at least n real zeros [and not more, because $H_n(z)$ is nth-degree] between which are interposed those of $H_{n-1}(z)$.

1-b. Generating function

Consider the function of z and λ:

$$F(z+\lambda) = \mathrm{e}^{-(z+\lambda)^2} \tag{8}$$

Taylor's formula enables us to write:

$$F(z+\lambda) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} F^{(n)}(z)$$
$$= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} (-1)^n H_n(z) \mathrm{e}^{-z^2} \tag{9}$$

Multiplying this relation by e^{z^2} and replacing λ by $-\lambda$ we obtain:

$$\mathrm{e}^{z^2} F(z-\lambda) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} H_n(z) \tag{10}$$

that is, if we replace $F(z-\lambda)$ by its value:

$$\mathrm{e}^{-\lambda^2+2\lambda z} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} H_n(z) \tag{11}$$

The Hermite polynomials can therefore be obtained from the series expansion in λ of the function $\mathrm{e}^{-\lambda^2+2\lambda z}$, which for this reason, is called the "generating function" of the Hermite polynomials.

Relation (11) gives us another definition of the Hermite polynomials $H_n(z)$:

$$H_n(z) = \left\{ \frac{\partial^n}{\partial \lambda^n} \mathrm{e}^{-\lambda^2+2\lambda z} \right\}_{\lambda=0} \tag{12}$$

1-c. Recurrence relations; differential equation

We have already obtained, in (6), one recurrence relation. It is easy to obtain others by differentiating relation (11). A differentiation with respect to z yields:

$$2\lambda\ \mathrm{e}^{-\lambda^2+2\lambda z} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \frac{\mathrm{d}}{\mathrm{d}z} H_n(z) \tag{13}$$

that is, replacing $\mathrm{e}^{-\lambda^2+2\lambda z}$ by the expansion (11) and setting equal terms of the same power in λ:

$$\frac{\mathrm{d}}{\mathrm{d}z} H_n(z) = 2n\, H_{n-1}(z) \tag{14}$$

Similarly, if we differentiate (11) with respect to λ, an analogous argument yields:

$$H_n(z) = 2z\ H_{n-1}(z) - 2(n-1)\, H_{n-2}(z) \tag{15}$$

Finally, it is not difficult to obtain a differential equation satisfied by the polynomials $H_n(z)$. Differentiating (14) and using (6), we get:

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2} H_n(z) = 2n \frac{\mathrm{d}}{\mathrm{d}z} H_{n-1}(z)$$
$$= 2n\, [2z H_{n-1}(z) - H_n(z)] \tag{16}$$

that is, replacing $H_{n-1}(z)$ by its value as given in (14):

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}z^2} - 2z \frac{\mathrm{d}}{\mathrm{d}z} + 2n \right] H_n(z) = 0 \tag{17}$$

1-d. Examples

Definition (7) or the recurrence relation (6) (which amounts to the same thing) enables us to calculate the first Hermite polynomials easily:

$$\begin{aligned} H_0(z) &= 1 \\ H_1(z) &= 2z \\ H_2(z) &= 4z^2 - 2 \\ H_3(z) &= 8z^3 - 12z \end{aligned} \tag{18}$$

In general:

$$H_n(z) = \left(2z - \frac{\mathrm{d}}{\mathrm{d}z}\right)^n (1) \tag{19}$$

2. The eigenfunctions of the harmonic oscillator Hamiltonian

2-a. Generating function

Consider the function:

$$K(\lambda, x) = \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} \langle x|\varphi_n\rangle \tag{20}$$

Using the relation [*cf.* Chap. V, formula (C-13)]:

$$|\varphi_n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|\varphi_0\rangle \tag{21}$$

we obtain:

$$\begin{aligned} K(\lambda, x) &= \sum_{n=0}^{\infty} \langle x| \frac{(\lambda a^\dagger)^n}{n!} |\varphi_0\rangle \\ &= \langle x|\mathrm{e}^{\lambda a^\dagger}|\varphi_0\rangle \end{aligned} \tag{22}$$

We now introduce, as in Chapter V, the dimensionless operators $\hat{X}$ and $\hat{P}$:

$$\begin{cases} \hat{X} = \beta X \\ \\ \hat{P} = \dfrac{P}{\beta\hbar} \end{cases} \tag{23}$$

where the parameter β, which has the dimensions of an inverse length, is defined by:

$$\beta = \sqrt{\frac{m\omega}{\hbar}} \tag{24}$$

The operator:

$$\mathrm{e}^{\lambda a^\dagger} = \mathrm{e}^{\frac{\lambda}{\sqrt{2}}(\hat{X} - i\hat{P})} \tag{25}$$

can be calculated by using formula (63) of Complement B_{II}, where we set:

$$\begin{cases} A = \dfrac{\lambda}{\sqrt{2}}\hat{X} \\ B = -\dfrac{i\lambda}{\sqrt{2}}\hat{P} \end{cases} \tag{26}$$

We obtain:

$$\begin{aligned} e^{\lambda a^\dagger} &= e^{\frac{\lambda}{\sqrt{2}}\hat{X}}\ e^{-\frac{i\lambda}{\sqrt{2}}\hat{P}}\ e^{\frac{i}{4}\lambda^2[\hat{X},\hat{P}]} \\ &= e^{\frac{\lambda}{\sqrt{2}}\hat{X}}\ e^{-\frac{i\lambda}{\sqrt{2}}\hat{P}}\ e^{-\lambda^2/4} \end{aligned} \tag{27}$$

Substituting this result into (22), we find:

$$\begin{aligned} K(\lambda, x) &= e^{-\lambda^2/4}\langle x|e^{(\lambda/\sqrt{2})\,\hat{X}}e^{(-i\lambda/\sqrt{2})\,\hat{P}}|\varphi_0\rangle \\ &= e^{-\lambda^2/4}\ e^{\beta\lambda x/\sqrt{2}}\langle x|e^{(-i\lambda/\sqrt{2})\,P/\beta\hbar}|\varphi_0\rangle \end{aligned} \tag{28}$$

Now, we have [*cf.* Complement E_{II}, formula (15)]:

$$\langle x|e^{-i\frac{\lambda P}{\beta\hbar\sqrt{2}}} = \langle x - \lambda/\beta\sqrt{2}| \tag{29}$$

and (28) can be written:

$$\begin{aligned} K(\lambda, x) &= e^{-\lambda^2/4}\ e^{\beta\lambda x/\sqrt{2}}\langle x - \lambda/\beta\sqrt{2}|\varphi_0\rangle \\ &= e^{-\lambda^2/4}\ e^{\beta\lambda x/\sqrt{2}}\varphi_0(x - \lambda/\beta\sqrt{2}) \end{aligned} \tag{30}$$

Using formula (C-25) of Chapter V, we finally obtain:

$$K(\lambda, x) = \left(\frac{\beta^2}{\pi}\right)^{1/4} \exp\left\{-\frac{\beta^2 x^2}{2} + \beta\lambda x\sqrt{2} - \frac{\lambda^2}{2}\right\} \tag{31}$$

According to definition (20), all we must do to find the wave functions $\varphi_n(x) = \langle x|\varphi_n\rangle$ is expand this expression in powers of λ:

$$K(\lambda, x) = \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}}\varphi_n(x) \tag{32}$$

$K(\lambda, x)$ is called the generating function of the $\varphi_n(x)$.

2-b. $\varphi_n(x)$ in terms of the Hermite polynomials

Replacing, in formula (11), λ by $\lambda/\sqrt{2}$ and z by βx, we obtain:

$$\exp\left\{-\frac{\lambda^2}{2} + \beta\lambda x\sqrt{2}\right\} = \sum_{n=0}^{\infty}\left(\frac{\lambda}{\sqrt{2}}\right)^n \frac{1}{n!}H_n(\beta x) \tag{33}$$

Substituting this expression into (31):

$$K(\lambda, x) = \left(\frac{\beta^2}{\pi}\right)^{1/4} \sum_{n=0}^{\infty} \left(\frac{\lambda}{\sqrt{2}}\right)^n \frac{1}{n!} \mathrm{e}^{-\beta^2 x^2/2}\ H_n(\beta x) \tag{34}$$

Setting equal the coefficients of the various powers of λ in (32) and (34), we obtain:

$$\varphi_n(x) = \left(\frac{\beta^2}{\pi}\right)^{1/4} \frac{1}{\sqrt{2^n\ n!}} \mathrm{e}^{-\beta^2 x^2/2}\ H_n(\beta x) \tag{35}$$

The shape of the function $\varphi_n(x)$ is therefore analogous to that of the nth-order derivative of the Gaussian function $F(x)$ considered in § 1 above; $\varphi_n(x)$ is of parity $(-1)^n$ and possesses n zeros interposed between those of $\varphi_{n+1}(x)$. We mentioned in § C-2 of Chapter V that this property is related to the increase in the average kinetic energy of the states $|\varphi_n\rangle$ when n increases.

2-c. Recurrence relations

Let us write the equations:

$$\begin{cases} a|\varphi_n\rangle = \sqrt{n}\ |\varphi_{n-1}\rangle \\ \\ a^\dagger|\varphi_n\rangle = \sqrt{n+1}\ |\varphi_{n+1}\rangle \end{cases} \tag{36}$$

in the $\{|x\rangle\}$ representation. Using the definitions of a and $a^\dagger$ [*cf.* Chap. V, relations (B-6)], we see that in the $\{|x\rangle\}$ representation, the action of these operators is given by:

$$a \Longrightarrow \frac{\beta}{\sqrt{2}}\left[x + \frac{1}{\beta^2}\frac{\mathrm{d}}{\mathrm{d}x}\right] \qquad a^\dagger \Longrightarrow \frac{\beta}{\sqrt{2}}\left[x - \frac{1}{\beta^2}\frac{\mathrm{d}}{\mathrm{d}x}\right] \tag{37}$$

Equations (36) therefore become:

$$\begin{cases} \dfrac{\beta}{\sqrt{2}}\left[x + \dfrac{1}{\beta^2}\dfrac{\mathrm{d}}{\mathrm{d}x}\right]\varphi_n(x) = \sqrt{n}\ \varphi_{n-1}(x) \\ \dfrac{\beta}{\sqrt{2}}\left[x - \dfrac{1}{\beta^2}\dfrac{\mathrm{d}}{\mathrm{d}x}\right]\varphi_n(x) = \sqrt{n+1}\ \varphi_{n+1}(x) \end{cases} \tag{38}$$

Let us take the sum and difference of these equations:

$$\begin{cases} x\beta\sqrt{2}\ \varphi_n(x) = \sqrt{n}\ \varphi_{n-1}(x) + \sqrt{n+1}\ \varphi_{n+1}(x) & (39) \\ \\ \dfrac{\sqrt{2}}{\beta}\ \dfrac{\mathrm{d}}{\mathrm{d}x}\varphi_n(x) = \sqrt{n}\ \varphi_{n-1}(x) - \sqrt{n+1}\ \varphi_{n+1}(x) & (40) \end{cases}$$

Comment:

If we replace the functions $\varphi_n(x)$ in (39) and (40) by their expressions given in (35), we obtain, after simplification (setting $\hat{x} = \beta x$):

$$2\hat{x}H_n(\hat{x}) = 2n\ H_{n-1}(\hat{x}) + H_{n+1}(\hat{x}) \tag{41}$$

$$2\left[-\hat{x}H_n(\hat{x}) + \frac{\mathrm{d}}{\mathrm{d}\hat{x}}H_n(\hat{x})\right] = 2n\ H_{n-1}(\hat{x}) - H_{n+1}(\hat{x}) \tag{42}$$

By taking the sum and the difference of these equations, we obtain relations (6) and (14) of § 1-a.

References

Messiah (1.17), App. B, § III; Arfken (10.4), Chap. 13, § 1; Angot (10.2), § 7.8.

Complement C_V

Solving the eigenvalue equation of the harmonic oscillator by the polynomial method

1 Changing the function and the variable 555
2 The polynomial method 557
2-a The asymptotic form of $\hat{\varphi}(\hat{x})$ 557
2-b The calculation of $h(\hat{x})$ in the form of a series expansion . . . 558
2-c Quantization of the energy 558
2-d Stationary wave functions 560

In § B of Chapter V, the method used to calculate the energies of the harmonic oscillator stationary states $|\varphi_n\rangle$ is based on the use of the operators a, $a^\dagger$ and N, as well as their their commutation relations. It is possible to obtain the same results in a completely different way, by solving the eigenvalue equation of the Hamiltonian H in the $\{|x\rangle\}$ representation. This is what we are going to do in this complement.

1. Changing the function and the variable

In the $\{|x\rangle\}$ representation, the eigenvalue equation of H is written:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+\frac{1}{2}m\omega^2x^2\right]\varphi(x)=E\,\varphi(x) \tag{1}$$

As in Chapter V, let us introduce the dimensionless operators $\hat{X}$ and $\hat{P}$:

$$\begin{cases}\hat{X}=\beta X\\ \\ \hat{P}=\dfrac{P}{\beta\hbar}\end{cases} \tag{2}$$

where the parameter β, which has the dimensions of an inverse length, is defined by:

$$\beta=\sqrt{\frac{m\omega}{\hbar}} \tag{3}$$

Let us denote by $|\xi_{\hat{x}}\rangle$ the eigenvector of $\hat{X}$ with eigenvalue is $\hat{x}$:

$$\hat{X}|\xi_{\hat{x}}\rangle=\hat{x}|\xi_{\hat{x}}\rangle \tag{4}$$

The orthonormalization and closure relations of the kets $|\xi_{\hat{x}}\rangle$ are written:

$$\langle\xi_{\hat{x}}|\xi_{\hat{x}'}\rangle=\delta(\hat{x}-\hat{x}') \tag{5}$$

$$\int_{-\infty}^{+\infty}\mathrm{d}\hat{x}\;|\xi_{\hat{x}}\rangle\langle\xi_{\hat{x}}|=1 \tag{6}$$

The ket $|\xi_{\hat{x}}\rangle$ is obviously an eigenvector of X, with the eigenvalue $\hat{x}/\beta$. Therefore, when:

$$\hat{x} = \beta x \tag{7}$$

the two kets $|x\rangle$ and $|\xi_{\hat{x}}\rangle$ are proportional. However, they are not equal. Writing the closure relation for the kets $|x\rangle$:

$$\int_{-\infty}^{+\infty} \mathrm{d}x \; |x\rangle \langle x| = 1 \tag{8}$$

and making the change of variables given in (7), we obtain:

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}\hat{x}}{\beta} \; |x = \frac{\hat{x}}{\beta}\rangle \langle x = \frac{\hat{x}}{\beta}| = 1 \tag{9}$$

Comparison with (6) shows that we can, for example, set:

$$|x = \frac{\hat{x}}{\beta}\rangle = \sqrt{\beta} \; |\xi_{\hat{x}}\rangle \tag{10}$$

to orthonormalize the kets $|\xi_{\hat{x}}\rangle$ with respect to $\hat{x}$, since the kets $|x\rangle$ are orthonormal with respect to x.

Let $|\varphi\rangle$ be an arbitrary ket, $\varphi(x) = \langle x|\varphi\rangle$ its wave function in the $\{|x\rangle\}$ representation, and $\hat{\varphi}(\hat{x}) = \langle \hat{x}|\varphi\rangle$ its wave function in the $\{|\xi_{\hat{x}}\rangle\}$ representation. According to (10):

$$\hat{\varphi}(\hat{x}) = \langle \xi_{\hat{x}}|\varphi\rangle = \frac{1}{\sqrt{\beta}} \; \langle x = \frac{\hat{x}}{\beta}|\varphi\rangle \tag{11}$$

that is:

$$\hat{\varphi}(\hat{x}) = \frac{1}{\sqrt{\beta}} \; \varphi(x = \frac{\hat{x}}{\beta}) \tag{12}$$

If $|\varphi\rangle$ is normalized, relation (8) yields:

$$\langle\varphi|\varphi\rangle = \langle\varphi| \left(\int_{-\infty}^{+\infty} \mathrm{d}x \; |x\rangle\langle x| \right) |\varphi\rangle = \int_{-\infty}^{+\infty} \varphi^*(x) \, \varphi(x) \; \mathrm{d}x = 1 \tag{13}$$

and relation (6) gives:

$$\langle\varphi|\varphi\rangle = \langle\varphi| \left(\int_{-\infty}^{+\infty} \mathrm{d}\hat{x} \; |\xi_{\hat{x}}\rangle \langle \xi_{\hat{x}}| \right) |\varphi\rangle = \int_{-\infty}^{+\infty} \hat{\varphi}^*(\hat{x}) \, \hat{\varphi}(\hat{x}) \; \mathrm{d}\hat{x} = 1 \tag{14}$$

The wave function $\varphi(x)$ is therefore normalized with respect to the variable x, as is $\hat{\varphi}(\hat{x})$ with respect to the variable $\hat{x}$. [This can be shown by using the integral in (13), in which we make the change of variables indicated in (7)].

Now, substituting (7) and (12) into (1), we obtain:

$$\frac{1}{2} \left[-\frac{\mathrm{d}^2}{\mathrm{d}\hat{x}^2} + \hat{x}^2 \right] \hat{\varphi}(\hat{x}) = \varepsilon \, \hat{\varphi}(\hat{x}) \tag{15}$$

setting:

$$\varepsilon = \frac{E}{\hbar\omega} \tag{16}$$

Equation (15) is more convenient than equation (1), since all the quantities appearing in it are dimensionless.

2. The polynomial method

2-a. The asymptotic form of $\hat{\varphi}(\hat{x})$

Equation (15) can be written:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\hat{x}^2} - (\hat{x}^2 - 2\varepsilon)\right]\hat{\varphi}(\hat{x}) = 0 \tag{17}$$

Let us try to predict intuitively the behavior of $\hat{\varphi}(\hat{x})$ for very large $\hat{x}$. To do this, consider the functions:

$$G_{\pm}(\hat{x}) = \mathrm{e}^{\pm\hat{x}^2/2} \tag{18}$$

They are solutions of the differential equations:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\hat{x}^2} - (\hat{x}^2 \pm 1)\right] G_{\pm}(\hat{x}) = 0 \tag{19}$$

When $\hat{x}$ approaches infinity:

$$\hat{x}^2 \pm 1 \sim \hat{x}^2 \sim \hat{x}^2 - 2\varepsilon \tag{20}$$

and equations (17) and (19) take on the same form asymptotically. We should therefore expect the solutions of equation (17) to behave[1], for large $\hat{x}$, either like $\mathrm{e}^{\hat{x}^2/2}$ or like $\mathrm{e}^{-\hat{x}^2/2}$. From a physical point of view, the only functions $\hat{\varphi}(\hat{x})$ of interest to us are those that are bounded everywhere. This restricts us to solutions of (17) that behave like $\mathrm{e}^{-\hat{x}^2/2}$ (if they exist). This is why we shall set:

$$\hat{\varphi}(\hat{x}) = \mathrm{e}^{-\hat{x}^2/2} h(\hat{x}) \tag{21}$$

Substituting (21) into (17), we obtain:

$$\frac{\mathrm{d}^2}{\mathrm{d}\hat{x}^2} h(\hat{x}) - 2\hat{x}\frac{\mathrm{d}}{\mathrm{d}\hat{x}} h(\hat{x}) + (2\varepsilon - 1)h(\hat{x}) = 0 \tag{22}$$

We are going to show how this equation can be solved by expanding $h(\hat{x})$ in a power series. Then we shall impose the condition that its solutions be physically acceptable.

[1] The solutions of equation (17) are not necessarily equivalent to $\mathrm{e}^{\hat{x}^2/2}$ or $\mathrm{e}^{-\hat{x}^2/2}$ when $\hat{x} \to \infty$: the intuitive arguments which we have given do not exclude, for example, the possibility that $\hat{\varphi}(\hat{x})$ may behave like the product of $\mathrm{e}^{\hat{x}^2/2}$ or $\mathrm{e}^{-\hat{x}^2/2}$ by a power of $\hat{x}$.

2-b. The calculation of $h(\hat{x})$ in the form of a series expansion

As we pointed out in § A-3 of Chapter V, the solutions of equation (1) [or, which amounts to the same thing, of (17)] can be sought amongst either even or odd functions. Since the function $e^{-\hat{x}^2/2}$ is even, we can therefore set:

$$h(\hat{x}) = \hat{x}^p \left(a_0 + a_2\hat{x}^2 + a_4\hat{x}^4 + ... + a_{2m}\,\hat{x}^{2m} + ...\right) \tag{23}$$

with $a_0 \neq 0$ (where $a_0\hat{x}^p$ is, by definition, the first non-zero term of the expansion). Writing (23) in the form:

$$h(\hat{x}) = \sum_{m=0}^{\infty} a_{2m}\,\hat{x}^{2m+p} \tag{24}$$

we easily obtain:

$$\frac{\mathrm{d}}{\mathrm{d}\hat{x}}h(\hat{x}) = \sum_{m=0}^{\infty}(2m+p)a_{2m}\,\hat{x}^{(2m+p-1)} \tag{25}$$

and:

$$\frac{\mathrm{d}^2}{\mathrm{d}\hat{x}^2}h(\hat{x}) = \sum_{m=0}^{\infty}(2m+p)(2m+p-1)a_{2m}\,\hat{x}^{(2m+p-2)} \tag{26}$$

Let us substitute (24), (25) and (26) into (22). For this equation to be satisfied, each term of the series expansion of the left-hand side must be equal to zero. For the general term in $\hat{x}^{2m+p}$, this condition is written:

$$(2m+p+2)(2m+p+1)a_{2m+2} = (4m+2p-2\varepsilon+1)a_{2m} \tag{27}$$

The term of lowest degree is in $\hat{x}^{p-2}$; its coefficient will be zero if:

$$p(p-1)a_0 = 0 \tag{28}$$

Since a_0 is not zero, we therefore have either $p = 0$ [the function $\hat{\varphi}(\hat{x})$ is then even] or $p = 1$ [$\hat{\varphi}(\hat{x})$ is then odd].

Relation (27) can be written:

$$a_{2m+2} = \frac{4m+2p+1-2\varepsilon}{(2m+p+2)(2m+p+1)}a_{2m} \tag{29}$$

which is a recurrence relation between the coefficients a_{2m}. Since a_0 is not zero, relation (29) enables us to calculate a_2 in terms of a_0, then a_4 in terms of a_2, and so on.

For arbitrary ε, we therefore have the series expansion of two linearly independent solutions of equation (22), corresponding respectively to $p = 0$ and $p = 1$.

2-c. Quantization of the energy

We must now choose, from amongst all the solutions found in the preceding section, those which satisfy the physical conditions that $\hat{\varphi}(\hat{x})$ be bounded everywhere.

For most values of ε, the numerator of (29) does not go to zero for any positive or zero integer m. Since none of the coefficients a_{2m} is then zero, the series has an infinite number of terms.

It can be shown that the asymptotic behavior of such a series makes it physically unacceptable. We see from (29) that:

$$\frac{a_{2m+2}}{a_{2m}} \underset{m\to\infty}{\sim} \frac{1}{m} \tag{30}$$

Now consider the power series expansion of the function $\mathrm{e}^{\lambda\hat{x}^2}$ (where λ is a real parameter):

$$\mathrm{e}^{\lambda\hat{x}^2} = \sum_{m=0}^{\infty} b_{2m}\,\hat{x}^{2m} \tag{31}$$

with:

$$b_{2m} = \frac{\lambda^m}{m!} \tag{32}$$

For this second series, we therefore have:

$$\frac{b_{2m+2}}{b_{2m}} = \frac{m!}{(m+1)!}\frac{\lambda^{m+1}}{\lambda^m} = \frac{\lambda}{m+1} \underset{m\to\infty}{\sim} \frac{\lambda}{m} \tag{33}$$

If we choose the value of the parameter λ such that:

$$0 < \lambda < 1 \tag{34}$$

we see from (30) and (33) that there exists an integer M such that the condition $m > M$ implies:

$$\frac{a_{2m+2}}{a_{2m}} > \frac{b_{2m+2}}{b_{2m}} > 0 \tag{35}$$

We can deduce from this that, when condition (34) is satisfied, we have:

$$\left|\hat{x}^{-p}\ h(\hat{x}) - P(\hat{x})\right| \geq \left|\frac{a_{2M}}{b_{2M}}\right| \left|\mathrm{e}^{\lambda\hat{x}^2} - Q(\hat{x})\right| \tag{36}$$

where $P(\hat{x})$ and $Q(\hat{x})$ are polynomials of degree $2M$ given by the first $M+1$ terms of series (23) and (31). When $\hat{x}$ approaches infinity, (36) gives:

$$|h(\hat{x})| \underset{\hat{x}\to\infty}{\geq} \left|\frac{a_{2M}}{b_{2M}}\right| \hat{x}^p\,\mathrm{e}^{\lambda\hat{x}^2} \tag{37}$$

and therefore:

$$|\hat{\varphi}(\hat{x})| \underset{\hat{x}\to\infty}{\geq} \left|\frac{a_{2M}}{b_{2M}}\right| \hat{x}^p\,\mathrm{e}^{(\lambda-1/2)\hat{x}^2} \tag{38}$$

Since we can choose λ such that:

$$1/2 < \lambda < 1 \tag{39}$$

$|\hat{\varphi}(\hat{x})|$ is not bounded when $\hat{x} \to \infty$. We must therefore reject this solution, which makes no sense physically.

There is only one possibility left: that the numerator of (29) goes to zero for a value m_0 of m. We then have:

$$\begin{cases} a_{2m} \neq 0 \quad \text{if} \quad m \leq m_0 \\ \\ a_{2m} = 0 \quad \text{if} \quad m > m_0 \end{cases} \tag{40}$$

and the power series expansion of $h(\hat{x})$ reduces to a polynomial of degree $2m_0 + p$. The behavior at infinity of $\hat{\varphi}(\hat{x})$ is then determined by that of the exponential $\mathrm{e}^{-\hat{x}^2/2}$, and $\hat{\varphi}(\hat{x})$ is physically acceptable (since it is square-integrable).

The fact that the numerator of (29) goes to zero at $m = m_0$ imposes the condition:

$$2\varepsilon = 2(2m_0 + p) + 1 \tag{41}$$

If we set:

$$2m_0 + p = n \tag{42}$$

equation (41) can be written:

$$\varepsilon = \varepsilon_n = n + 1/2 \tag{43}$$

where n is an arbitrary positive integer or zero (m is an arbitrary positive integer or zero, and p is equal to 0 or 1). Condition (43) introduces the quantization of the harmonic oscillator energy, since it implies [*cf.* (16)]:

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega \tag{44}$$

We have thus obtained relation (B-34) of Chapter V.

2-d. Stationary wave functions

The polynomial method also yields the eigenfunctions associated with the various energies E_n, in the form:

$$\hat{\varphi}_n(\hat{x}) = \mathrm{e}^{-\hat{x}^2/2}\ h_n(\hat{x}) \tag{45}$$

where $h_n(\hat{x})$ is an nth degree polynomial. According to (23) and (42), $h_n(\hat{x})$ is an even function if n is even and an odd function if n is odd.

The ground state is obtained for $n = 0$, that is, for $m_0 = p = 0$; $h_0(\hat{x})$ is then a constant, and:

$$\hat{\varphi}_0(\hat{x}) = a_0\, \mathrm{e}^{-\hat{x}^2/2} \tag{46}$$

A simple calculation shows that, to normalize $\hat{\varphi}_0(\hat{x})$ with respect to the variable $\hat{x}$, it suffices to choose:

$$a_0 = \pi^{-1/4} \tag{47}$$

Then, using (12), we find:

$$\varphi_0(x) = \left(\frac{\beta^2}{\pi}\right)^{1/4} \mathrm{e}^{-\beta^2 x^2/2} \tag{48}$$

which is indeed the expression given in Chapter V [formula (C-25)].

To the first excited state $E_1 = 3\hbar\omega/2$ corresponds $n = 1$, that is, $m_0 = 0$ and $p = 1$; $h_1(\hat{x})$ then has only one term, obtained by a calculation analogous to the preceding one:

$$\hat{\varphi}_1(\hat{x}) = \left(\frac{4}{\pi}\right)^{1/4} \hat{x}\, \mathrm{e}^{-\hat{x}^2/2} \tag{49a}$$

that is:

$$\varphi_1(x) = \left(\frac{4\beta^6}{\pi}\right)^{1/4} x\ \mathrm{e}^{-\beta^2 x^2/2} \tag{49b}$$

For $n = 2$, we have $m_0 = 1$ and $p = 0$. Relation (29) then yields:

$$a_2 = -2a_0 \tag{50}$$

which finally leads to:

$$\hat{\varphi}_2(\hat{x}) = \left(\frac{1}{4\pi}\right)^{1/4} (2\hat{x}^2 - 1)\ \mathrm{e}^{-\hat{x}^2/2} \tag{51a}$$

that is:

$$\varphi_2(x) = \left(\frac{\beta^2}{4\pi}\right)^{1/4} (2\beta^2 x^2 - 1)\ \mathrm{e}^{-\beta^2 x^2/2} \tag{51b}$$

For arbitrary n, $h_n(\hat{x})$ is the polynomial solution of equation (22), which can be written, taking the quantization condition (43) into account:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\hat{x}^2} - 2\hat{x}\frac{\mathrm{d}}{\mathrm{d}\hat{x}} + 2n\right] h(\hat{x}) = 0 \tag{52}$$

We recognize (52) to be the differential equation satisfied by the Hermite polynomial $H_n(\hat{x})$ [see equation (17) of Complement $\mathrm{B_V}$]. The polynomial $h_n(\hat{x})$ is therefore proportional to $H_n(\hat{x})$, where the proportionality factor is determined by normalization of $\hat{\varphi}(\hat{x})$. This is in agreement with formula (35) of complement $\mathrm{B_V}$.

References

Mathematical treatment of differential equations: Morse and Feshbach (10.13), Chaps. 5 and 6; Courant and Hilbert (10.11), § V-11.

Complement D_V

Study of the stationary states in the $\{|p\rangle\}$ representation

1 Wave functions in momentum space 563
1-a Changing the variable and the function 563
1-b Determination of $\hat{\overline{\varphi}}_n(\hat{p})$ 564
1-c Calculation of the phase factor 564
2 Discussion 565

The distribution of the possible momenta of a particle in the state $|\varphi_n\rangle$ is given by the wave function $\overline{\varphi}_n(p)$ in the $\{|\mathbf{p}\rangle\}$ representation, the Fourier transform of the wave function $\varphi_n(x)$ in the $\{|x\rangle\}$ representation. We shall show in this complement that in the case of the harmonic oscillator, the functions φ_n and $\overline{\varphi}_n$ are the same (to within multiplicative factors). Thus, in a stationary state, the probability distributions of the momentum and the position have similar forms.

1. Wave functions in momentum space

1-a. Changing the variable and the function

In Complement C_V, we introduced, for simplicity, the operator:

$$\hat{X} = \beta X \tag{1}$$

where:

$$\beta = \sqrt{\frac{m\omega}{\hbar}} \tag{2}$$

as well as the eigenkets $|\xi_{\hat{x}}\rangle$ of $\hat{X}$ and the wave function $\hat{\varphi}(\hat{x})$ in the $\{|\xi_{\hat{x}}\rangle\}$ representation. We shall follow a similar procedure for the operator:

$$\hat{P} = \frac{P}{\beta\hbar} \tag{3}$$

We shall therefore call $|\pi_{\hat{p}}\rangle$ the eigenkets of $\hat{P}$:

$$\hat{P}|\pi_{\hat{p}}\rangle = \hat{p}|\pi_{\hat{p}}\rangle \tag{4}$$

and denote by $\hat{\overline{\varphi}}(\hat{p})$ the wave function in the $\{|\pi_{\hat{p}}\rangle\}$ representation:

$$\hat{\overline{\varphi}}(\hat{p}) = \langle\pi_{\hat{p}}|\varphi\rangle \tag{5}$$

Just as the ket $|\xi_{\hat{x}}\rangle$ is proportional to the ket $|x = \hat{x}/\beta\rangle$, the ket $|\pi_{\hat{p}}\rangle$ is proportional to the ket $|p = \beta\hbar\hat{p}\rangle$. If we change β to $1/\beta\hbar$ [*cf.* (1) and (3)], equation (10) of Complement C_V shows that:

$$|\pi_{\hat{p}}\rangle = \sqrt{\beta\hbar}\ \ |p = \beta\hbar\hat{p}\rangle \tag{6}$$

The wave function $\hat{\overline{\varphi}}(\hat{p})$ in the $\{|\pi_{\hat{p}}\rangle\}$ representation is therefore related to the wave function $\overline{\varphi}(p)$ in the $\{|p\rangle\}$ representation by:

$$\hat{\overline{\varphi}}(\hat{p}) = \sqrt{\beta\hbar}\ \overline{\varphi}(p = \beta\hbar\hat{p}) \tag{7}$$

Furthermore, we can use (6) and relation (10) of Complement C_V to obtain:

$$\langle \xi_{\hat{x}} | \pi_{\hat{p}} \rangle = \frac{e^{i\hat{p}\hat{x}}}{\sqrt{2\pi}} \tag{8}$$

We therefore have, using definition (5) and the closure relation for the $\{|\xi_{\hat{x}}\rangle\}$ basis:

$$\begin{aligned} \hat{\overline{\varphi}}(\hat{p}) &= \int_{-\infty}^{+\infty} \langle \pi_{\hat{p}} | \xi_{\hat{x}} \rangle \, \langle \xi_{\hat{x}} | \varphi \rangle \, d\hat{x} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-i\hat{p}\hat{x}} \ \hat{\varphi}(\hat{x})\, d\hat{x} \end{aligned} \tag{9}$$

The function $\hat{\overline{\varphi}}$ is therefore the Fourier transform of $\hat{\varphi}$.

1-b. Determination of $\hat{\overline{\varphi}}_n(\hat{p})$

We have seen [*cf.* equation (15) of Complement C_V] that the stationary wave functions $\hat{\varphi}(\hat{x})$ of the harmonic oscillator satisfy the equation:

$$\frac{1}{2}\left[-\frac{d^2}{d\hat{x}^2} + \hat{x}^2\right] \hat{\varphi}(\hat{x}) = \varepsilon\, \hat{\varphi}(\hat{x}) \tag{10}$$

Now, the Fourier transform of $\dfrac{d^2}{d\hat{x}^2}\, \hat{\varphi}(\hat{x})$ is $-\hat{p}^2 \hat{\overline{\varphi}}(\hat{p})$ and that of $\hat{x}^2\, \hat{\varphi}(\hat{x})$ is $-\dfrac{d^2}{d\hat{p}^2}\hat{\overline{\varphi}}(\hat{p})$. The Fourier transform of equation (10) is therefore:

$$\frac{1}{2}\left[\hat{p}^2 - \frac{d^2}{d\hat{p}^2}\right] \hat{\overline{\varphi}}(\hat{p}) = \varepsilon\, \hat{\overline{\varphi}}(\hat{p}) \tag{11}$$

If we compare (10) and (11), we see that the functions $\hat{\varphi}_n$ and $\hat{\overline{\varphi}}_n$ satisfy the same differential equation. Moreover, we know that this equation, when $\varepsilon = n + 1/2$ (where n is a positive integer or zero), has only one square-integrable solution (the eigenvalues ε_n are non-degenerate; *cf.* Chapter V, § B-3). We can conclude that $\hat{\varphi}_n$ and $\hat{\overline{\varphi}}_n$ are proportional. Since these two functions are normalized, the proportionality factor is a complex number of modulus 1, so that:

$$\hat{\overline{\varphi}}_n(\hat{p}) = e^{i\theta_n} \hat{\varphi}_n(\hat{x} = \hat{p}) \tag{12}$$

where $e^{i\theta_n}$ is a phase factor which we shall now determine.

1-c. Calculation of the phase factor

The wave function of the ground state is given by [*cf.* Complement C_V, formulas (46) and (47)]:

$$\hat{\varphi}_0(\hat{x}) = \pi^{-1/4}\, e^{-\hat{x}^2/2} \tag{13}$$

This is a Gaussian function; its Fourier transform is therefore [*cf.* Appendix I, relation (50)]:

$$\hat{\overline{\varphi}}_0(\hat{p}) = \pi^{-1/4}\,\mathrm{e}^{-\hat{p}^2/2} \tag{14}$$

This implies that θ_0 is zero.

To find θ_n, let us write, in the $\{|\xi_{\hat{x}}\rangle\}$ and $\{|\pi_{\hat{p}}\rangle\}$ representations, the relation:

$$a^\dagger|\varphi_n\rangle = \sqrt{n+1}\;|\varphi_{n+1}\rangle \tag{15}$$

In the $|\xi_{\hat{x}}\rangle$ representation, $\hat{X}$ and $\hat{P}$ act like $\hat{x}$ and $\dfrac{1}{i}\dfrac{\mathrm{d}}{\mathrm{d}\hat{x}}$; $a^\dagger$ therefore acts like $\dfrac{1}{\sqrt{2}}\left(\hat{x} - \dfrac{\mathrm{d}}{\mathrm{d}\hat{x}}\right)$. In the $\{|\hat{p}\rangle\}$ representation, $\hat{X}$ acts like $i\dfrac{\mathrm{d}}{\mathrm{d}\hat{p}}$ and $\hat{P}$ like $\hat{p}$; $a^\dagger$ therefore acts like $\dfrac{i}{\sqrt{2}}\left(\dfrac{\mathrm{d}}{\mathrm{d}\hat{p}} - \hat{p}\right)$.

In the $\{|\xi_{\hat{x}}\rangle\}$ representation, relation (15) therefore becomes:

$$\hat{\varphi}_{n+1}(\hat{x}) = \frac{1}{\sqrt{2(n+1)}}\left(\hat{x} - \frac{\mathrm{d}}{\mathrm{d}\hat{x}}\right)\hat{\varphi}_n(\hat{x}) \tag{16}$$

while in the $\{|\pi_{\hat{p}}\rangle\}$ representation, it becomes:

$$\hat{\overline{\varphi}}_{n+1}(\hat{p}) = \frac{i}{\sqrt{2(n+1)}}\left(\frac{\mathrm{d}}{\mathrm{d}\hat{p}} - \hat{p}\right)\hat{\overline{\varphi}}_n(\hat{p}) \tag{17}$$

We therefore have:

$$\mathrm{e}^{i\theta_{n+1}} = -i\,\mathrm{e}^{i\theta_n} \tag{18}$$

that is, knowing that $\theta_0 = 0$:

$$\mathrm{e}^{i\theta_n} = (-i)^n \tag{19}$$

Thus we obtain:

$$\hat{\overline{\varphi}}_n(\hat{p}) = (-i)^n\,\hat{\varphi}_n(\hat{x} = \hat{p}) \tag{20}$$

or, returning to the functions φ_n and $\overline{\varphi}_n$:

$$\overline{\varphi}_n(p) = (-i)^n \frac{1}{\beta\sqrt{\hbar}}\;\varphi_n\left(x = \frac{p}{\beta^2\hbar}\right) \tag{21}$$

2. Discussion

Consider a particle in the state $|\varphi_n\rangle$. When the position of the particle is measured, one has a probability $\rho_n(x)\,\mathrm{d}x$ of finding a result between x and $x + \mathrm{d}x$, where $\rho_n(x)$ is given by:

$$\rho_n(x) = |\varphi_n(x)|^2 \tag{22}$$

Similarly, in a measurement of the momentum of the particle, one has a probability $\overline{\rho}_n(p)\,\mathrm{d}p$ of finding a result between p and $p+\mathrm{d}p$, with:

$$\overline{\rho}_n(p) = |\overline{\varphi}_n(p)|^2 \tag{23}$$

Relation (21) then yields:

$$\overline{\rho}_n(p) = \frac{1}{m\omega}\ \rho_n\left(x = \frac{p}{m\omega}\right) \tag{24}$$

which shows that the momentum distribution in a stationary state has the same form as the position distribution.

We see, for example (*cf.* Fig. 6 of Chapter V), that if n is large, $\overline{\rho}_n(p)$ has a peak at each of the two values:

$$p = \pm\ m\omega x_M = \pm\ p_M \tag{25}$$

where p_M is the maximum momentum of the classical particle moving in the potential well with an energy E_n. An argument analogous to the one set forth at the end of § C-2 of Chapter V enables us to understand this result. When the momentum of the classical particle is equal to $\pm p_M$, its acceleration is zero (its velocity is stationary), and the values $\pm p_M$ of the momentum are, averaging over time, the most probable ones. Comment *(i)* of § D-1 of Chapter V concerning the probability density $\rho_n(x)$ can easily be transposed to this context; for example, when n is large, the root mean square deviation ΔP can be interpreted as being of the order of magnitude of the distance between the peaks of $\overline{\rho}_n(p)$ situated at $p = \pm p_M$.

It is also possible to understand directly from Figure 6-a of Chapter V why these values of the momentum are highly probable when n is large. The wave function then performs a large number of oscillations between the two peaks, analogous to those of a sinusoid. This happens because the differential equation for the wave function [*cf.* formula (A-17) of Chapter V] when $E \gg m\omega^2 x^2/2$ becomes:

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi(x) + \frac{2mE}{\hbar^2}\varphi(x) \simeq 0 \tag{26}$$

which yields, according to the definition of p_M:

$$\varphi(x) \simeq A\,\mathrm{e}^{ip_M\cdot x/\hbar} + A'\,\mathrm{e}^{-ip_M\cdot x/\hbar} \tag{27}$$

The wave function (when n is large) therefore looks like a sinusoid of wavelength h/p_M over a relatively large region of the Ox axis. This sinusoid can be considered to be the sum of two progressive waves [*cf.* (27)] associated with the two opposite momenta $\pm p_M$ (corresponding to the to-and-fro motion of the particle in the well). It is not surprising, therefore, that the probability density $\overline{\rho}_n(p)$ should be large in the neighborhood of the values $p = \pm p_M$.

An analogous argument also enables us to understand the order of magnitude of the product $\Delta X.\Delta P$. This product is equal to [*cf.* Chap. V, relations (D-6), (D-7) and (D-9)]:

$$\Delta X.\Delta P = \left(n + \frac{1}{2}\right)\hbar = \frac{x_M.p_M}{2} \tag{28}$$

When n increases, the amplitudes x_M and p_M of the oscillations increase, and the product $\Delta X.\Delta P$ takes on values much greater than its minimum value $\hbar/2$. We might wonder why this is the case, since we have seen in several examples that when the width ΔX of a function increases, the width ΔP of its Fourier transform decreases. This is indeed what would happen for the functions $\varphi_n(x)$ if, in the interval $-x_M \leq x \leq +x_M$ where they take on non-negligible values, they varied slowly, reaching, for example, a single maximum or minimum. This is in fact the case for small values of n, for which the value of the product $\Delta X.\Delta P$ is indeed near its minimum. However, when n is large, the functions $\varphi_n(x)$ perform numerous oscillations in the interval $-x_M \leq x \leq +x_M$, where they have n zeros. One can therefore associate with them wavelengths of the order of $\lambda \simeq x_M/n \simeq \Delta X/n$, corresponding to momenta of the particle situated in a domain of dimension ΔP given by:

$$\Delta P \simeq \frac{h}{\lambda} \simeq \frac{nh}{\Delta X} \tag{29}$$

We thus find again that:

$$\Delta X.\Delta P \simeq nh \tag{30}$$

This situation is somewhat analogous to the one studied in § 1 of Complement A_{III}, in connection with the infinite one-dimensional well.

Complement E_V

The isotropic three-dimensional harmonic oscillator

1 The Hamiltonian operator 569
2 Separation of the variables in Cartesian coordinates 570
3 Degeneracy of the energy levels 572

In Chapter V, we studied the one-dimensional harmonic oscillator. We now show how to use the results of this study to treat the three-dimensional harmonic oscillator.

1. The Hamiltonian operator

Consider a spinless particle of mass m which can move in three-dimensional space. The particle is subjected to a central force (i.e. a force that is constantly directed towards the coordinate origin O) whose absolute value is proportional to the distance of the particle from the point O:

$$\mathbf{F} = -k\,\mathbf{r} \tag{1}$$

(k is a positive constant).

This force field is derived from the potential energy:

$$V(\mathbf{r}) = \frac{1}{2}k\,\mathbf{r}^2 = \frac{1}{2}m\omega^2\mathbf{r}^2 \tag{2}$$

where the angular frequency ω is defined as for the one-dimensional harmonic oscillator:

$$\omega = \sqrt{\frac{k}{m}} \tag{3}$$

The classical Hamiltonian is therefore:

$$\mathcal{H}(\mathbf{r},\mathbf{p}) = \frac{\mathbf{p}^2}{2m} + \frac{1}{2}m\omega^2\mathbf{r}^2 \tag{4}$$

Using the quantization rules (Chap. III, § B-5), we immediately deduce the Hamiltonian operator:

$$H = \frac{\mathbf{P}^2}{2m} + \frac{1}{2}m\omega^2\mathbf{R}^2 \tag{5}$$

Since the Hamiltonian H is time-independent, we shall solve its eigenvalue equation:

$$H|\psi\rangle = E|\psi\rangle \tag{6}$$

where $|\psi\rangle$ belongs to the state space $\mathcal{E}_\mathbf{r}$ of a particle in three-dimensional space.

Comment:

Since $V(\mathbf{r})$ depends only on the distance $r = |\mathbf{r}|$ of the particle from the origin [$V(\mathbf{r})$ is consequently invariant under an arbitrary rotation], this harmonic oscillator is said to be *isotropic*. Nevertheless, the calculations which follow can easily be generalized to the case of an anisotropic harmonic oscillator, for which:

$$V(\mathbf{r}) = \frac{m}{2}\left(\omega_x^2 x^2 + \omega_y^2 y^2 + \omega_z^2 z^2\right) \tag{7}$$

where the three constants ω_x, ω_y and ω_z are different.

2. Separation of the variables in Cartesian coordinates

Recall that the state space $\mathcal{E}_\mathbf{r}$ can be considered (*cf.* Chap. II, § F) to be the tensor product:

$$\mathcal{E}_\mathbf{r} = \mathcal{E}_x \otimes \mathcal{E}_y \otimes \mathcal{E}_z \tag{8}$$

where $\mathcal{E}_x$ is the state space of a particle moving along Ox, that is, the space associated with the wave functions $\varphi(x)$. $\mathcal{E}_y$ and $\mathcal{E}_z$ are defined analogously.

Now, expression (5) for the Hamiltonian H can be written in the form:

$$\begin{aligned} H &= \frac{1}{2m}\left(P_x^2 + P_y^2 + P_z^2\right) + \frac{1}{2}m\omega^2\left(X^2 + Y^2 + Z^2\right) \\ &= H_x + H_y + H_z \end{aligned} \tag{9}$$

with:

$$H_x = \frac{P_x^2}{2m} + \frac{1}{2}m\omega^2 X^2 \tag{10}$$

and similar definitions for H_y and H_z. H_x is a function only of X and P_x: H_x is therefore the extension into $\mathcal{E}_\mathbf{r}$ of an operator that actually acts in $\mathcal{E}_x$. Similarly, H_y and H_z act only in $\mathcal{E}_y$ and $\mathcal{E}_z$ respectively. In $\mathcal{E}_x$, H_x is a one-dimensional harmonic oscillator Hamiltonian. The same is true for H_y and H_z in $\mathcal{E}_y$ and $\mathcal{E}_z$.

H_x, H_y and H_z commute. Each of them therefore commutes with their sum H. Consequently, the eigenvalue equation (6) can be solved by seeking the eigenvectors of H that are also eigenvectors of H_x, H_y and H_z. Now, we already know the eigenvectors and eigenvalues of H_x in $\mathcal{E}_x$, as well as those of H_y in $\mathcal{E}_y$ and of H_z in $\mathcal{E}_z$:

$$H_x|\varphi_{n_x}\rangle = \left(n_x + \frac{1}{2}\right)\hbar\omega \quad |\varphi_{n_x}\rangle \ ; \ |\varphi_{n_x}\rangle \in \mathcal{E}_x \tag{11a}$$

$$H_y|\varphi_{n_y}\rangle = \left(n_y + \frac{1}{2}\right)\hbar\omega \quad |\varphi_{n_y}\rangle \ ; \ |\varphi_{n_y}\rangle \in \mathcal{E}_y \tag{11b}$$

$$H_z|\varphi_{n_z}\rangle = \left(n_z + \frac{1}{2}\right)\hbar\omega \quad |\varphi_{n_z}\rangle \ ; \ |\varphi_{n_z}\rangle \in \mathcal{E}_z \tag{11c}$$

(n_x, n_y and n_z are positive integers or zero). It follows (*cf.* Chap. II, § F) that the eigenstates common to H, H_x, H_y and H_z are of the form:

$$|\psi_{n_x,n_y,n_z}\rangle = |\varphi_{n_x}\rangle|\varphi_{n_y}\rangle|\varphi_{n_z}\rangle \tag{12}$$

According to equations (9) and (11):

$$H|\psi_{n_x,n_y,n_z}\rangle = \left(n_x + n_y + n_z + \frac{3}{2}\right)\hbar\omega|\psi_{n_x,n_y,n_z}\rangle \tag{13}$$

The eigenvectors of H are seen to be *tensor products of eigenvectors* of H_x, H_y and H_z respectively, and the eigenvalues of H, to be *sums of eigenvalues* of these three operators.

According to equation (13), the energy levels E_n of the isotropic three-dimensional harmonic oscillator are of the form:

$$E_n = \left(n + \frac{3}{2}\right)\hbar\omega \tag{14}$$

with:

$$n \text{ a positive integer or zero} \tag{15}$$

since n is the sum $n_x + n_y + n_z$ of three numbers, each of which can take on any non-negative integral value.

Furthermore, formula (12) enables us to deduce the properties of the vectors $|\psi_{n_x,n_y,n_z}\rangle$, common eigenstates of H, H_x, H_y and H_z, from those derived in § C-1 of Chapter V for $|\varphi_{n_x}\rangle$ (which are also valid for $|\varphi_{n_y}\rangle$ and $|\varphi_{n_z}\rangle$).

Let us introduce three pairs of creation and annihilation operators:

$$a_x = \sqrt{\frac{m\omega}{2\hbar}}X + \frac{i}{\sqrt{2m\hbar\omega}}P_x \qquad a_x^\dagger = \sqrt{\frac{m\omega}{2\hbar}}X - \frac{i}{\sqrt{2m\hbar\omega}}P_x \tag{16a}$$

$$a_y = \sqrt{\frac{m\omega}{2\hbar}}Y + \frac{i}{\sqrt{2m\hbar\omega}}P_y \qquad a_y^\dagger = \sqrt{\frac{m\omega}{2\hbar}}Y - \frac{i}{\sqrt{2m\hbar\omega}}P_y \tag{16b}$$

$$a_z = \sqrt{\frac{m\omega}{2\hbar}}Z + \frac{i}{\sqrt{2m\hbar\omega}}P_z \qquad a_z^\dagger = \sqrt{\frac{m\omega}{2\hbar}}Z - \frac{i}{\sqrt{2m\hbar\omega}}P_z \tag{16c}$$

These operators are the extensions into $\mathcal{E}_\mathbf{r}$ of operators acting in $\mathcal{E}_x$, $\mathcal{E}_y$ and $\mathcal{E}_z$. The canonical commutation relations between the components of $\mathbf{R}$ and $\mathbf{P}$ imply that the only non-zero commutators of the six operators defined in (16) are:

$$[a_x, a_x^\dagger] = [a_y, a_y^\dagger] = [a_z, a_z^\dagger] = 1 \tag{17}$$

Note that two operators with different indices always commute, as is logical because they act in different spaces. The action of the operators a_x and $a_x^\dagger$ on the states $|\psi_{n_x,n_y,n_z}\rangle$ is given by the formulas:

$$\begin{aligned} a_x|\psi_{n_x,n_y,n_z}\rangle &= (a_x|\varphi_{n_x}\rangle)\ |\varphi_{n_y}\rangle|\varphi_{n_z}\rangle \\ &= \sqrt{n_x}\ |\varphi_{n_x-1}\rangle|\varphi_{n_y}\rangle|\varphi_{n_z}\rangle \\ &= \sqrt{n_x}\ |\varphi_{n_x-1,n_y,n_z}\rangle \end{aligned} \tag{18a}$$

$$\begin{aligned} a_x^\dagger|\psi_{n_x,n_y,n_z}\rangle &= (a_x^\dagger|\varphi_{n_x}\rangle)\ |\varphi_{n_y}\rangle|\varphi_{n_z}\rangle \\ &= \sqrt{n_x+1}\ |\varphi_{n_x+1}\rangle|\varphi_{n_y}\rangle|\varphi_{n_z}\rangle \\ &= \sqrt{n_x+1}\ |\varphi_{n_x+1,n_y,n_z}\rangle \end{aligned} \tag{18b}$$

For a_y, $a_y^\dagger$ and a_z, $a_z^\dagger$, we have analogous relations.

We also know [*cf.* equation (C-13) of Chapter V] that:

$$|\varphi_{n_x}\rangle = \frac{1}{\sqrt{n_x!}}(a_x^\dagger)^{n_x}|\varphi_0\rangle \tag{19}$$

where $|\varphi_0\rangle$ is the vector of $\mathcal{E}_x$ that satisfies the condition:

$$a_x|\varphi_0\rangle = 0 \tag{20}$$

In $\mathcal{E}_y$ and $\mathcal{E}_z$, there are analogous expressions for $|\varphi_{n_y}\rangle$ and $|\varphi_{n_z}\rangle$. Consequently, according to (12):

$$|\psi_{n_x,n_y,n_z}\rangle = \frac{1}{\sqrt{n_x!n_y!n_z!}}(a_x^\dagger)^{n_x}(a_y^\dagger)^{n_y}(a_z^\dagger)^{n_z}|\psi_{0,0,0}\rangle \tag{21}$$

where $|\psi_{0,0,0}\rangle$ is the tensor product of the ground states of the three one-dimensional oscillators, so that:

$$a_x|\psi_{0,0,0}\rangle = a_y|\psi_{0,0,0}\rangle = a_z|\psi_{0,0,0}\rangle = 0 \tag{22}$$

Finally, recall that, since $|\psi_{n_x,n_y,n_z}\rangle$ is a tensor product, the associated wave function is of the form:

$$\langle\mathbf{r}|\psi_{n_x,n_y,n_z}\rangle = \varphi_{n_x}(x)\ \varphi_{n_y}(y)\ \varphi_{n_z}(z) \tag{23}$$

where φ_{n_x}, φ_{n_y} and φ_{n_z} are stationary wave functions of the one-dimensional harmonic oscillator (Chap. V, § C-2). For example:

$$\langle\mathbf{r}|\psi_{0,0,0}\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{3/4} \mathrm{e}^{-\frac{m\omega}{2\hbar}(x^2+y^2+z^2)} \tag{24}$$

3. Degeneracy of the energy levels

We showed in § B-3 of Chapter V that H_x constitutes a C.S.C.O. in $\mathcal{E}_x$; the same is true for H_y in $\mathcal{E}_y$ and for H_z in $\mathcal{E}_z$. According to § F of Chapter II, $\{H_x, H_y, H_z\}$ is thus a C.S.C.O. in $\mathcal{E}_\mathbf{r}$. Therefore, there exists (to within a constant factor) a unique ket $|\psi_{n_x,n_y,n_z}\rangle$ of $\mathcal{E}_\mathbf{r}$ corresponding to a given set of eigenvalues of H_x, H_y and H_z, that is, to fixed non-negative integers n_x, n_y and n_z.

However, H alone does not form a C.S.C.O. since *the energy levels E_n are degenerate*. If we choose an eigenvalue of H, $E_n = (n + 3/2)\hbar\omega)$ (which amounts to fixing a non-negative integer n), all the kets of the $\{|\psi_{n_x,n_y,n_z}\rangle\}$ basis that satisfy:

$$n_x + n_y + n_z = n \tag{25}$$

are eigenvectors of H with the eigenvalue E_n.

The degree of degeneracy g_n of E_n is therefore equal to the number of different sets $\{n_x, n_y, n_z\}$ satisfying condition (25). To find g_n, we can proceed as follows. With n fixed, choose n_x first, giving it one of the values:

$$n_x = 0, 1, 2, \ldots n \tag{26}$$

With n_x thus fixed, we must have:

$$n_y + n_z = n - n_x \tag{27}$$

There are then $(n - n_x + 1)$ possibilities for the pair $\{n_y, n_z\}$:

$$\{n_y, n_z\} = \{0, n - n_x\},\ \{1, n - n_x - 1\},\ \ldots \{n - n_x, 0\} \tag{28}$$

The degree of degeneracy g_n of E_n is therefore equal to:

$$g_n = \sum_{n_x=0}^{n} (n - n_x + 1) \tag{29}$$

This sum is easy to calculate:

$$g_n = (n+1) \sum_{n_x=0}^{n} 1 - \sum_{n_x=0}^{n} n_x = \frac{(n+1)(n+2)}{2} \tag{30}$$

Consequently, only the ground state $E_0 = \frac{3}{2}\hbar\omega$ is non-degenerate.

Comment:

The kets $|\psi_{n_x,n_y,n_z}\rangle$ constitute an orthonormal system of eigenvectors of H, which forms a basis in $\mathcal{E}_{\mathbf{r}}$. Since the eigenvalues E_n of H are degenerate, this system is not unique. We shall see in particular in Complement $\mathrm{B_{VII}}$ that, in order to solve equation (6), it is possible to use a set of constants of the motion other than $\{H_x, H_y, H_z\}$: thus we obtain a basis of $\mathcal{E}_{\mathbf{r}}$ which is different from the preceding one, although still consisting of eigenvectors of H. The kets of this new basis are orthonormal linear combinations of the $|\psi_{n_x,n_y,n_z}\rangle$ belonging to each of the eigensubspaces of H, that is, corresponding to a fixed value of the sum $n_x+n_y+n_z$.

Complement F_V

A charged harmonic oscillator in a uniform electric field

1 Eigenvalue equation of $H'(\mathscr{E})$ in the $\{|x\rangle\}$ representation . . 575

2 Discussion 577

2-a Electrical susceptibility of an elastically bound electron . . . 577

2-b Interpretation of the energy shift 578

3 Use of the translation operator 579

The one-dimensional harmonic oscillator studied in Chapter V consists of a particle of mass m having a potential energy:

$$V(X) = \frac{1}{2}m\omega^2 X^2 \tag{1}$$

Assume, in addition, that this particle has a charge q and that it is placed in a uniform electric field $\mathscr{E}$ parallel to Ox. What are its new stationary states and the corresponding energies?

The classical potential energy of a particle placed in a uniform field $\mathscr{E}$ is equal to[1]:

$$w(\mathscr{E}) = -q\mathscr{E}x \tag{2}$$

To obtain the quantum mechanical Hamiltonian operator $H'(\mathscr{E})$ in the presence of the field $\mathscr{E}$, we must therefore add to the potential energy (1) of the harmonic oscillator the term:

$$W(\mathscr{E}) = -q\mathscr{E}X \tag{3}$$

which gives:

$$H'(\mathscr{E}) = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2 - q\mathscr{E}X \tag{4}$$

We must now find the eigenvalues and eigenvectors of this operator. To this end, we shall use two different methods. First, we shall consider directly the eigenvalue equation of $H'(\mathscr{E})$ in the $\{|x\rangle\}$ representation, as it is very simple to interpret the results obtained. Then we shall show how the problem can be solved in a purely operator formalism.

1. Eigenvalue equation of $H'(\mathscr{E})$ in the $\{|x\rangle\}$ representation

Let $|\varphi'\rangle$ be an eigenvector of $H'(\mathscr{E})$:

$$H'(\mathscr{E})|\varphi'\rangle = E'|\varphi'\rangle \tag{5}$$

[1] We use the convention of zero potential energy for the particle at $x = 0$.

Using (4), we can write this equation in the $\{|x\rangle\}$ representation:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+\frac{1}{2}m\omega^2x^2-q\mathscr{E}x\right]\varphi'(x)=E'\varphi'(x) \tag{6}$$

Completing the square with respect to x on the left-hand side of (6), we get:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+\frac{1}{2}m\omega^2\left(x-\frac{q\mathscr{E}}{m\omega^2}\right)^2-\frac{q^2\mathscr{E}^2}{2m\omega^2}\right]\varphi'(x)=E'\varphi'(x) \tag{7}$$

Let us now replace the variable x by a new variable u, setting:

$$u=x-\frac{q\mathscr{E}}{m\omega^2} \tag{8}$$

Through the intermediary of x, φ' is then a function of u, and equation (7) becomes:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}u^2}+\frac{1}{2}m\omega^2u^2\right]\varphi'(u)=E''\varphi'(u) \tag{9}$$

with:

$$E''=E'+\frac{q^2\mathscr{E}^2}{2m\omega^2} \tag{10}$$

Thus we see that equation (9) is the same as the one used to obtain the stationary states of the harmonic oscillator in the absence of an electric field in the $\{|x\rangle\}$ representation [*cf.* Chap. V, relation (A-17)]. Therefore, we have already solved this equation, and we know that the acceptable values of E'' are given by:

$$E''_n=\left(n+\frac{1}{2}\right)\hbar\omega \tag{11}$$

(where n is a positive integer or zero).

Relations (10) and (11) show that, in the presence of the electric field, the energies E' of the stationary states of the harmonic oscillator are modified:

$$E'_n(\mathscr{E})=\left(n+\frac{1}{2}\right)\hbar\omega-\frac{q^2\mathscr{E}^2}{2m\omega^2} \tag{12}$$

The entire spectrum of the harmonic oscillator is therefore shifted by the quantity $q^2\mathscr{E}^2/2m\omega^2$.

Now, let us show that the eigenfunctions $\varphi'_n(x)$ associated with the energies (12) can all be obtained from the $\varphi_n(x)$ by a translation along Ox. The solution of (9) corresponding to a given value of n is $\varphi_n(u)$ [where the function φ_n is given, for example, by formula (35) of Complement B_V]. According to (8), we have:

$$\varphi'_n(x)=\varphi_n\left(x-\frac{q\mathscr{E}}{m\omega^2}\right) \tag{13}$$

This translation comes from the fact that the electric field exerts a force on the particle[2].

[2]It can be seen from (13) that the function $\varphi'_n(x)$ is obtained from $\varphi_n(x)$ by a translation of $q\mathscr{E}/m\omega^2$; if the product $q\mathscr{E}$ is positive, the translation is performed in the positive x-direction, which is indeed the direction of the force exerted by $\mathscr{E}$.

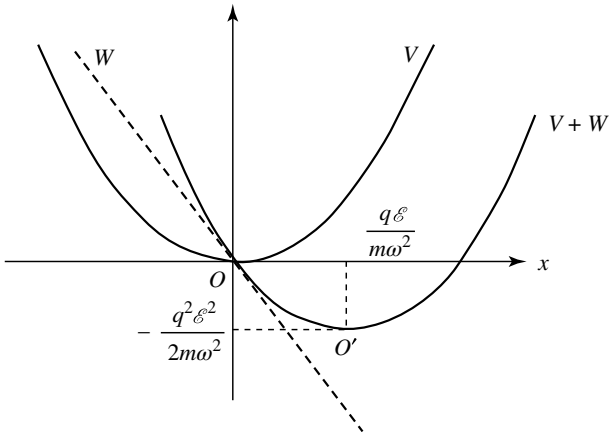


Figure 1: The presence of a uniform electric field has the effect of adding a linear term W to the potential energy V of the harmonic oscillator; the total potential $V+W$ is then represented by a displaced parabola.

Comment:

The change of variable given in (8) allows us to reduce the case of an arbitrary electric field to an already solved problem, the one in which $\mathscr{E}$ was zero. The only effect of the electric field is to change the x-origin [*cf.* (13)] and the energy origin [*cf.* (12)]. This result can easily be understood graphically (*cf.* Fig. 1). When $\mathscr{E}$ is zero, the potential energy $V(x)$ is represented by a parabola centered at O. When $\mathscr{E}$ is not zero, it is necessary to add to this potential energy the quantity $-q\mathscr{E}x$, which corresponds to the dashed line in this figure; the curve representing $V+W$ is again a parabola. Thus, in the presence of the field $\mathscr{E}$ we still have a harmonic oscillator. Since the two parabolas are superposable, they correspond to the same value of ω and therefore to the same energy difference between the levels. However, their minima O and O' are different, as is consistent with formulas (12) and (13).

2. Discussion

2-a. Electrical susceptibility of an elastically bound electron

In certain situations, the electrons of an atom or a molecule behave, to a good approximation, as if they were "elastically bound", that is, as if each of them were a harmonic oscillator. We shall prove this for atoms in Complement $\mathrm{A_{XIII}}$, using time-dependent perturbation theory.

The contribution of each electron to the electric dipole moment of the atom is described by the operator:

$$D = qX \tag{14}$$

where q is the charge of the electron ($q < 0$) and X the corresponding position observable. We are going to examine the mean value of D in the model of the elastically bound electron.

In the absence of an electric field, the mean value of the electric dipole moment in a stationary state of the oscillator is zero:

$$\langle D \rangle = q \langle \varphi_n | X | \varphi_n \rangle = 0 \tag{15}$$

[see formulas (D-1) of Chapter V].

Now, let us assume that the field $\mathscr{E}$ is turned on so slowly that the state of the electron changes gradually from $|\varphi_n\rangle$ to $|\varphi'_n\rangle$ (n remaining the same). The mean dipole moment is now different from zero, since:

$$\langle D \rangle' = q \, \langle \varphi'_n | X | \varphi'_n \rangle = q \int_{-\infty}^{+\infty} \mathrm{d}x \; x \; |\varphi'_n(x)|^2 \tag{16}$$

Using (8) and (13), we obtain:

$$\langle D \rangle' = q \int_{-\infty}^{+\infty} u \; |\varphi_n(u)|^2 \; \mathrm{d}u + \frac{q^2 \mathscr{E}}{m\omega^2} \int_{-\infty}^{+\infty} |\varphi_n(u)|^2 \; \mathrm{d}u = \frac{q^2 \mathscr{E}}{m\omega^2} \tag{17}$$

since the first integral is zero by symmetry. $\langle D \rangle'$ is therefore proportional to $\mathscr{E}$. In this model, the electrical susceptibility of the atomic electron under consideration is equal to:

$$\chi = \frac{\langle D \rangle'}{\mathscr{E}} = \frac{q^2}{m\omega^2} \tag{18}$$

It is positive, whatever the sign of q.

It is simple to interpret result (18) physically. The effect of the electric field is to shift the classical equilibrium position of the electron, that is, the mean value of its position in quantum mechanics [see formula (13)]. This results in the appearance of an induced dipole moment. χ decreases when ω increases because the oscillator is less easily deformable when the restoring force (which is proportional to ω^2) is larger.

2-b. Interpretation of the energy shift

Using the model just described, we can interpret formula (12) by calculating the variation in the mean kinetic and potential energies of the electron when it passes from the state $|\varphi_n\rangle$ to the state $|\varphi'_n\rangle$.

The variation in the kinetic energy is, in fact, zero (as can be understood intuitively from Figure 1, for example):

$$\left\langle \frac{P^2}{2m} \right\rangle' - \left\langle \frac{P^2}{2m} \right\rangle = -\frac{\hbar^2}{2m} \left[\int_{-\infty}^{+\infty} \varphi'^*_n(x) \; \frac{\mathrm{d}^2}{\mathrm{d}x^2} \varphi'_n(x) \; \mathrm{d}x \right. \\ \left. - \int_{-\infty}^{+\infty} \varphi^*_n(x) \; \frac{\mathrm{d}^2}{\mathrm{d}x^2} \varphi_n(x) \; \mathrm{d}x \right] = 0 \tag{19}$$

according to formula (13).

The variation in the potential energy can be treated in two terms:

- the first term, $\langle W(\mathscr{E})\rangle'$, corresponds to the electrical potential energy of the dipole in the field $\mathscr{E}$. Since the dipole is parallel to the field, we have, according to (17):

$$\langle W(\mathscr{E})\rangle' = -\mathscr{E}\langle D\rangle' = -\frac{q^2\mathscr{E}^2}{m\omega^2} \tag{20}$$

- the second term, $\langle V(X)\rangle' - \langle V(X)\rangle$, arises from the electric field modification of the wave function of the level labeled by the quantum number n. The "elastic" potential energy of the particle therefore changes by a quantity:

$$\langle V(X)\rangle' - \langle V(X)\rangle = \frac{1}{2}m\omega^2\left[\int_{-\infty}^{+\infty} x^2\,|\varphi_n'(x)|^2\,\mathrm{d}x - \int_{-\infty}^{+\infty} x^2\,|\varphi_n(x)|^2\,\mathrm{d}x\right] \tag{21}$$

The first integral can be calculated by using (13) and the change of variable (8):

$$\int_{-\infty}^{+\infty} x^2\,|\varphi_n'(x)|^2\,\mathrm{d}x = \int_{-\infty}^{+\infty} u^2\,|\varphi_n(u)|^2\,\mathrm{d}u + \frac{2q\mathscr{E}}{m\omega^2}\int_{-\infty}^{+\infty} u\,|\varphi_n(u)|^2\,\mathrm{d}u + \left(\frac{q\mathscr{E}}{m\omega^2}\right)^2\int_{-\infty}^{+\infty}|\varphi_n(u)|^2\,\mathrm{d}u \tag{22}$$

Since $\varphi_n(u)$ is normalized, and since the integral of $u|\varphi_n(u)|^2$ is zero by symmetry, we obtain, finally:

$$\langle V(X)\rangle' - \langle V(X)\rangle = \frac{q^2\mathscr{E}^2}{2m\omega^2} \tag{23}$$

We see why this result should be positive, since the electric field moves the particle away from the point O and attracts it into a region where the "elastic" potential energy $V(x)$ is larger.

Adding (20) to (23), we again find that the energy of the state $|\varphi_n'\rangle$ is less than that of the state $|\varphi_n\rangle$ by $q^2\mathscr{E}^2/2m\omega^2$.

3. Use of the translation operator

We shall see in this section that, instead of using the $\{|x\rangle\}$ representation, as we have done until now, we can argue directly in terms of the operator $H'(\mathscr{E})$ given in (4). More precisely, we are going to show that a unitary transformation (which corresponds to a translation of the wave function along the x-axis) transforms the operator $H = H'(\mathscr{E} = 0)$ into the operator $H'(\mathscr{E})$ (to within an additive constant which does not change the eigenvectors). Since the eigenvectors and eigenvalues of H were determined in Chapter V, this approach enables us to solve our problem.

Therefore, consider the operator:

$$U(\lambda) = \mathrm{e}^{-\lambda(a-a^\dagger)} \tag{24}$$

where λ is a real constant. Its adjoint $U^\dagger(\lambda)$ is:

$$U^\dagger(\lambda) = \mathrm{e}^{\lambda(a-a^\dagger)} \tag{25}$$

It is clear that:

$$U(\lambda)\,U^{\dagger}(\lambda) = U^{\dagger}(\lambda)\,U(\lambda) = 1 \tag{26}$$

$U(\lambda)$ is therefore a unitary operator. Under the corresponding unitary transformation, H becomes:

$$\begin{aligned}\tilde{H} &= U(\lambda)\,H\,U^{\dagger}(\lambda) \\ &= \hbar\omega\left[\frac{1}{2} + U(\lambda)\,a^{\dagger}a\,U^{\dagger}(\lambda)\right]\end{aligned} \tag{27}$$

We must now calculate the operator:

$$U(\lambda)\,a^{\dagger}a\,U^{\dagger}(\lambda) = \tilde{a}^{\dagger}\tilde{a} \tag{28}$$

with:

$$\begin{aligned}\tilde{a} &= U(\lambda)\,a\,U^{\dagger}(\lambda) \\ \tilde{a}^{\dagger} &= U(\lambda)\,a^{\dagger}\,U^{\dagger}(\lambda)\end{aligned} \tag{29}$$

To obtain $\tilde{a}$ and $\tilde{a}^{\dagger}$, we use formula (63) of Complement B_{II}, (which can be applied here since the commutator of a and $a^{\dagger}$ is equal to 1), which yields:

$$\begin{aligned}U(\lambda) &= \mathrm{e}^{-\lambda a+\lambda a^{\dagger}} = \mathrm{e}^{-\lambda a}\,\mathrm{e}^{\lambda a^{\dagger}}\mathrm{e}^{\lambda^2/2} \\ U^{\dagger}(\lambda) &= \mathrm{e}^{-\lambda a^{\dagger}+\lambda a} = \mathrm{e}^{-\lambda a^{\dagger}}\,\mathrm{e}^{\lambda a}\mathrm{e}^{-\lambda^2/2}\end{aligned} \tag{30}$$

Also, formula (51) of Complement B_{II} enables us to write:

$$\begin{cases}\left[\mathrm{e}^{-\lambda a}, a^{\dagger}\right] = -\lambda\,\mathrm{e}^{-\lambda a} \\ \\ \left[\mathrm{e}^{\lambda a^{\dagger}}, a\right] = -\lambda\,\mathrm{e}^{\lambda a^{\dagger}}\end{cases} \tag{31}$$

that is:

$$\begin{aligned}\mathrm{e}^{-\lambda a}a^{\dagger}\,\mathrm{e}^{\lambda a} &= a^{\dagger} - \lambda \\ \mathrm{e}^{\lambda a^{\dagger}}\,a\,\mathrm{e}^{-\lambda a^{\dagger}} &= a - \lambda\end{aligned} \tag{32}$$

Thus it follows that:

$$\begin{aligned}\tilde{a} &= \mathrm{e}^{-\lambda a}\,\mathrm{e}^{\lambda a^{\dagger}}\,a\,\mathrm{e}^{-\lambda a^{\dagger}}\,\mathrm{e}^{\lambda a} \\ &= \mathrm{e}^{-\lambda a}(a-\lambda)\mathrm{e}^{\lambda a} = a - \lambda\end{aligned} \tag{33}$$

and, similarly:

$$\tilde{a}^{\dagger} = a^{\dagger} - \lambda \tag{34}$$

$\tilde{H}$ is therefore given by:

$$\begin{aligned}\tilde{H} &= \hbar\omega\left[\frac{1}{2} + (a^\dagger - \lambda)(a - \lambda)\right] \\ &= \hbar\omega\left[\frac{1}{2} + a^\dagger a - \lambda(a + a^\dagger) + \lambda^2\right] \\ &= H - \lambda\hbar\omega(a + a^\dagger) + \lambda^2\hbar\omega \end{aligned} \tag{35}$$

Since $(a+a^\dagger)$ is proportional to the operator X [formulas (B-1) and (B-7) of Chapter V], it suffices to set:

$$\lambda = \frac{q\mathscr{E}}{\omega}\sqrt{\frac{1}{2m\hbar\omega}} \tag{36}$$

to obtain:

$$\tilde{H} = H - q\mathscr{E}X + \frac{q^2\mathscr{E}^2}{2m\omega^2} = H'(\mathscr{E}) + \frac{q^2\mathscr{E}^2}{2m\omega^2} \tag{37}$$

The two operators $\tilde{H}$ and $H'(\mathscr{E})$ therefore have the same eigenvectors, and their eigenvalues differ by $q^2\mathscr{E}^2/2m\omega^2$. Now, we know (*cf.* Complement $\mathrm{C_{II}}$, § 2) that if the eigenvectors of H are the kets $|\varphi_n\rangle$, those of $\tilde{H}$ are the kets:

$$|\tilde{\varphi}_n\rangle = U(\lambda)|\varphi_n\rangle \tag{38}$$

and the corresponding eigenvalues of H and $\tilde{H}$ are the same. The stationary states $|\varphi'_n\rangle$ of the harmonic oscillator in the presence of the field $\mathscr{E}$ are therefore the states $|\tilde{\varphi}_n\rangle$ given by (38). The associated eigenvalue of $H'(\mathscr{E})$ is, according to (37):

$$E'_n(\mathscr{E}) = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{q^2\mathscr{E}^2}{2m\omega^2} \tag{39}$$

which is the same as formula (12) of the preceding section. Expression (38) for the eigenvectors can be put into the form:

$$|\varphi'_n\rangle = |\tilde{\varphi}_n\rangle = \mathrm{e}^{-i\frac{q\mathscr{E}}{m\hbar\omega^2}P}|\varphi_n\rangle \tag{40}$$

using (24) and (36), as well as formulas (B-1) and (B-7) of Chapter V. We interpreted, in Complement $\mathrm{E_{II}}$, the operator $\mathrm{e}^{-ibP/\hbar}$ as being the translation operator over an algebraic distance b along Ox. $|\varphi'_n\rangle$ is therefore the state obtained from $|\varphi_n\rangle$ by a translation $q\mathscr{E}/m\omega^2$, just as is indicated by formula (13).

References:

The elastically bound electron: see references of Complement $\mathrm{A_{XIII}}$.

Complement G_V

Coherent "quasi-classical" states of the harmonic oscillator

1 Quasi-classical states 584
1-a Introducing the parameter α_0 to characterize a classical motion 584
1-b Conditions defining quasi-classical states 586
1-c Quasi-classical states are eigenvectors of the operator a . . . 587
2 Properties of the $|\alpha\rangle$ states 588
2-a Expansion of $|\alpha\rangle$ on the basis of the stationary states $|\varphi_n\rangle$ . . 588
2-b Possible values of the energy in an $|\alpha\rangle$ state 589
2-c Calculation of $\langle X\rangle$, $\langle P\rangle$, ΔX and ΔP in an $|\alpha\rangle$ state 591
2-d The operator $D(\alpha)$: the wave functions $\Psi_\alpha(x)$ 591
2-e The scalar product of two $|\alpha\rangle$ states. Closure relation 593
3 Time evolution of a quasi-classical state 594
3-a A quasi-classical state always remains an eigenvector of a . . 595
3-b Evolution of physical properties 595
3-c Motion of the wave packet 596
4 Example: quantum mechanical treatment of a macroscopic oscillator 596

The properties of the stationary states $|\varphi_n\rangle$ of the harmonic oscillator were studied in Chapter V; for example, in § D, we saw that the mean values $\langle X\rangle$ and $\langle P\rangle$ of the position and the momentum of the oscillator are zero in such a state. Now, in classical mechanics, it is well-known that the position x and the momentum p are oscillating functions of time, which always remain zero only if the energy of the motion is also zero [*cf.* Chap. V, relations (A-5) and (A-8)]. Furthermore, we know that quantum mechanics must yield the same results as classical mechanics in the limiting case where the harmonic oscillator has an energy much greater than the quantum $\hbar\omega$ (limit of large quantum numbers).

Thus, we may ask the following question: is it possible to construct quantum mechanical states leading to physical predictions which are almost identical to the classical ones, at least for a macroscopic oscillator? We shall see in this complement that such quantum states exist: they are coherent linear superpositions of all the states $|\varphi_n\rangle$. We shall call them "quasi-classical states" or "coherent states of the harmonic oscillator".

The problem we are considering here is of great general interest in quantum mechanics. As we saw in the introduction to Chapter V and in Complement A_V, many physical systems can be likened to a harmonic oscillator, at least to a first approximation. For all these systems, it is important to understand, in the framework of quantum mechanics, how to move gradually from the case in which the results given by the classical approximation are sufficient to the case in which quantum effects are preponderant. Electromagnetic radiation is a very important example of such a system. Depending on the experiment, it either reveals its quantum mechanical nature (as is the case in the

experiment discussed in § A-2-a of Chapter I, in which the light intensity is very low) or else can be treated classically. "Coherent states" of electromagnetic radiation were introduced by Glauber and are in current use in the domain of quantum optics.

The position, the momentum, and the energy of a harmonic oscillator are described in quantum mechanics by operators which do not commute; they are incompatible physical quantities. It is not possible, therefore, to construct a state in which they are all perfectly well-defined. We shall thus only look for a state vector such that, for all t, the mean values $\langle X \rangle$, $\langle P \rangle$ and $\langle H \rangle$ are as close as possible to the corresponding classical values. This will lead us to a compromise in which none of these three observables is perfectly known. We shall see, nevertheless, that the root mean square deviations ΔX, ΔP and ΔH are, in the macroscopic limit, completely negligible.

1. Quasi-classical states

1-a. Introducing the parameter α_0 to characterize a classical motion

The classical equations of motion of a one-dimensional harmonic oscillator, of mass m and angular frequency ω, are written:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} x(t) = \dfrac{1}{m} p(t) & \text{(1a)} \\ \\ \dfrac{\mathrm{d}}{\mathrm{d}t} p(t) = -m\omega^2 \, x(t) & \text{(1b)} \end{cases}$$

The quantum mechanical calculations we shall perform later will be simplified by the introduction of the dimensionless quantities:

$$\begin{cases} \hat{x}(t) = \beta \, x(t) \\ \\ \hat{p}(t) = \dfrac{1}{\hbar\beta} \, p(t) \end{cases} \tag{2}$$

where:

$$\beta = \sqrt{\frac{m\omega}{\hbar}} \tag{3}$$

Equations (1) can then be written:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \hat{x}(t) = \omega \, \hat{p}(t) & \text{(4a)} \\ \\ \dfrac{\mathrm{d}}{\mathrm{d}t} \hat{p}(t) = -\omega \, \hat{x}(t) & \text{(4b)} \end{cases}$$

The classical state of the harmonic oscillator is determined at time t when we know its position $x(t)$ and its momentum $p(t)$, that is, $\hat{x}(t)$ and $\hat{p}(t)$. We shall therefore combine these two real numbers into a single dimensionless complex number $\alpha(t)$ defined by:

$$\alpha(t) = \frac{1}{\sqrt{2}} \left[\hat{x}(t) + i \, \hat{p}(t) \right] \tag{5}$$

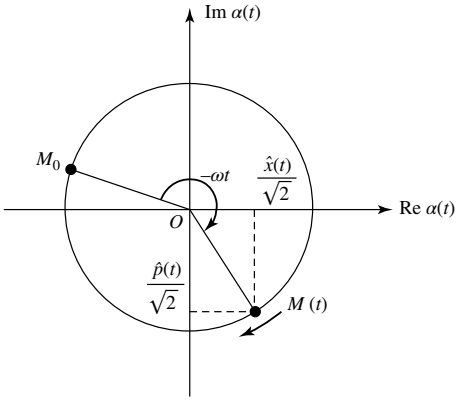


Figure 1: The point $M(t)$, which corresponds to the complex number $\alpha(t)$, characterizes the state of the harmonic oscillator at each instant. M moves in a circle with an angular velocity $-\omega$. The abscissa and ordinate of M give the position and momentum of the oscillator.

The set of two equations (4) is equivalent to the single equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha(t) = -i\omega\,\alpha(t) \tag{6}$$

whose solution is:

$$\alpha(t) = \alpha_0\,\mathrm{e}^{-i\omega t} \tag{7}$$

where we have set:

$$\alpha_0 = \alpha(0) = \frac{1}{\sqrt{2}}\left[\hat{x}(0) + i\,\hat{p}(0)\right] \tag{8}$$

Now consider the points M_0 and M in the complex plane that correspond to the complex numbers α_0 and $\alpha(t)$ [Fig. 1]. M is at M_0 at $t = 0$ and describes, with an angular velocity $-\omega$, a circle centered at O of radius OM_0.

Since, according to (5), the coordinates of M are equal to $\hat{x}(t)/\sqrt{2}$ and $\hat{p}(t)/\sqrt{2}$, we thus obtain a very simple geometrical representation of the time evolution of the state of the system. Every possible motion corresponding to given initial conditions is entirely characterized by the point M_0, that is, by the complex number α_0 (the modulus of α_0 gives the amplitude of the oscillation and the argument of α_0, its phase). According to (5) and (7), we have:

$$\begin{cases} \hat{x}(t) = \dfrac{1}{\sqrt{2}}\left[\alpha_0\,\mathrm{e}^{-i\omega t} + \alpha_0^*\,\mathrm{e}^{i\omega t}\right] & \text{(9a)} \\[2ex] \hat{p}(t) = -\dfrac{i}{\sqrt{2}}\left[\alpha_0\,\mathrm{e}^{-i\omega t} - \alpha_0^*\,\mathrm{e}^{i\omega t}\right] & \text{(9b)} \end{cases}$$

As for the classical energy $\mathcal{H}$ of the system, it is constant in time and equal to:

$$\begin{aligned} \mathcal{H} &= \frac{1}{2m}[p(0)]^2 + \frac{1}{2}m\omega^2[x(0)]^2 \\ &= \frac{\hbar\omega}{2}\left\{[\hat{x}(0)]^2 + [\hat{p}(0)]^2\right\} \end{aligned} \tag{10}$$

which yields, taking (8) into account:

$$\mathcal{H} = \hbar\omega|\alpha_0|^2 \tag{11}$$

For a macroscopic oscillator, the energy $\mathcal{H}$ is much greater than the quantum $\hbar\omega$, so:

$$|\alpha_0| \gg 1 \tag{12}$$

1-b. Conditions defining quasi-classical states

We are looking for a quantum mechanical state for which at every instant the mean values $\langle X\rangle$, $\langle P\rangle$ and $\langle H\rangle$ are practically equal to the values x, p and $\mathcal{H}$ which correspond to a given classical motion.

To calculate $\langle X\rangle$, $\langle P\rangle$ and $\langle H\rangle$, we use the expressions:

$$\begin{aligned}\hat{X} &= \beta X = \frac{1}{\sqrt{2}}(a + a^\dagger)\\ \hat{P} &= \frac{1}{\hbar\beta}P = -\frac{i}{\sqrt{2}}(a - a^\dagger)\end{aligned} \tag{13}$$

and:

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) \tag{14}$$

For an arbitrary state $|\psi(t)\rangle$, the time evolution of the matrix element $\langle a\rangle(t) = \langle\psi(t)|a|\psi(t)\rangle$ is given by (*cf.* § D-1-d of Chapter III):

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle a\rangle(t) = \langle[a, H]\rangle(t) \tag{15}$$

Now:

$$[a, H] = \hbar\omega[a, a^\dagger a] = \hbar\omega\, a \tag{16}$$

which implies:

$$i\frac{\mathrm{d}}{\mathrm{d}t}\langle a\rangle(t) = \omega\langle a\rangle(t) \tag{17}$$

that is:

$$\langle a\rangle(t) = \langle a\rangle(0)\,\mathrm{e}^{-i\omega t} \tag{18}$$

The evolution of $\langle a^\dagger\rangle(t) = \langle\psi(t)|a^\dagger|\psi(t)\rangle$ obeys the complex conjugate equation:

$$\begin{aligned}\langle a^\dagger\rangle(t) &= \langle a^\dagger\rangle(0)\,\mathrm{e}^{i\omega t}\\ &= \langle a\rangle^*(0)\,\mathrm{e}^{i\omega t}\end{aligned} \tag{19}$$

(18) and (19) are analogous to the classical equation (7).

Substituting (18) and (19) into (13), we obtain:

$$\begin{cases} \langle \hat{X} \rangle(t) = \frac{1}{\sqrt{2}} \left[\langle a \rangle(0)\, \mathrm{e}^{-i\omega t} + \langle a \rangle^*(0)\, \mathrm{e}^{i\omega t} \right] \\ \\ \langle \hat{P} \rangle(t) = -\dfrac{i}{\sqrt{2}} \left[\langle a \rangle(0)\, \mathrm{e}^{-i\omega t} - \langle a \rangle^*(0)\, \mathrm{e}^{i\omega t} \right] \end{cases} \tag{20}$$

Comparing these results with (9), we see that, in order to have at all times t:

$$\begin{cases} \langle \hat{X} \rangle(t) = \hat{x}(t) \\ \\ \langle \hat{P} \rangle(t) = \hat{p}(t) \end{cases} \tag{21}$$

it is necessary and sufficient to set, at the instant $t = 0$, the condition:

$$\langle a \rangle(0) = \alpha_0 \tag{22}$$

where α_0 is the complex parameter characterizing the classical motion which we are trying to reproduce quantum mechanically. The normalized state vector $|\psi(t)\rangle$ of the oscillator must therefore satisfy the condition:

$$\langle \psi(0)|a|\psi(0)\rangle = \alpha_0 \tag{23}$$

We must now require the mean value:

$$\langle H \rangle = \hbar\omega \langle a^\dagger a \rangle(0) + \frac{\hbar\omega}{2} \tag{24}$$

to be equal to the classical energy $\mathcal{H}$ given by (11). Since, for a classical oscillator, $|\alpha_0|$ is much greater than 1 [*cf.* (12)], we shall neglect the term $\hbar\omega/2$ (of purely quantum mechanical origin; see § D-2 of Chapter V) with respect to $\hbar\omega|\alpha_0|^2$. The second condition on the state vector can now be written:

$$\langle a^\dagger a \rangle(0) = |\alpha_0|^2 \tag{25}$$

that is:

$$\langle \psi(0)|a^\dagger a|\psi(0)\rangle = |\alpha_0|^2 \tag{26}$$

We shall see that conditions (23) and (26) are sufficient to determine the normalized state vector $|\psi(0)\rangle$ (to within a constant phase factor).

1-c. Quasi-classical states are eigenvectors of the operator a

We introduce the operator $b(\alpha_0)$ defined by:

$$b(\alpha_0) = a - \alpha_0 \tag{27}$$

We then have:

$$b^\dagger(\alpha_0)\, b(\alpha_0) = a^\dagger a - \alpha_0 a^\dagger - \alpha_0^* a + \alpha_0^* \alpha_0 \tag{28}$$

and the square of the norm of the ket $b(\alpha_0)|\psi(0)\rangle$ is:

$$\langle\psi(0)|b^\dagger(\alpha_0)\, b(\alpha_0)|\psi(0)\rangle = \langle\psi(0)|a^\dagger a|\psi(0)\rangle - \alpha_0\langle\psi(0)|a^\dagger|\psi(0)\rangle - \alpha_0^*\langle\psi(0)|a|\psi(0)\rangle + \alpha_0^*\alpha_0 \tag{29}$$

Substituting into this relation conditions (23) and (26), we obtain:

$$\langle\psi(0)|b^\dagger(\alpha_0)\, b(\alpha_0)|\psi(0)\rangle = \alpha_0^*\alpha_0 - \alpha_0\alpha_0^* - \alpha_0^*\alpha_0 + \alpha_0^*\alpha_0 = 0 \tag{30}$$

The ket $b(\alpha_0)|\psi(0)\rangle$, whose norm is zero, is therefore zero:

$$b(\alpha_0)|\psi(0)\rangle = 0 \tag{31}$$

that is:

$$a|\psi(0)\rangle = \alpha_0|\psi(0)\rangle \tag{32}$$

Conversely, if the normalized vector $|\psi(0)\rangle$ satisfies this relation, it is obvious that conditions (23) and (26) are satisfied.

We therefore arrive at the following result: the quasi-classical state, associated with a classical motion characterized by the parameter α_0, is such that $|\psi(0)\rangle$ is an eigenvector of the operator a with the eigenvalue α_0.

In what follows, we shall denote the eigenvector of a with eigenvalue α by $|\alpha\rangle$:

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{33}$$

[we shall show later that the solution of (33) is unique to within a constant factor].

2. Properties of the $|\alpha\rangle$ states

2-a. Expansion of $|\alpha\rangle$ on the basis of the stationary states $|\varphi_n\rangle$

Let us determine the ket $|\alpha\rangle$ which is a solution of (33) by using an expansion on the states $|\varphi_n\rangle$:

$$|\alpha\rangle = \sum_n c_n(\alpha)|\varphi_n\rangle \tag{34}$$

We then have:

$$a|\alpha\rangle = \sum_n c_n(\alpha)\sqrt{n}|\varphi_{n-1}\rangle \tag{35}$$

and, substituting this relation into (33), we obtain:

$$c_{n+1}(\alpha) = \frac{\alpha}{\sqrt{n+1}} c_n(\alpha) \tag{36}$$

This relation enables us to determine by recurrence all the coefficients $c_n(\alpha)$ in terms of $c_0(\alpha)$:

$$c_n(\alpha) = \frac{\alpha^n}{\sqrt{n!}} c_0(\alpha) \tag{37}$$

It follows that, when $c_0(\alpha)$ is fixed, all the $c_n(\alpha)$ are also fixed. The vector $|\alpha\rangle$ is therefore unique to within a multiplicative factor. We shall choose $c_0(\alpha)$ real and positive and normalize the ket $|\alpha\rangle$, which determines it completely. In this case, the coefficients $c_n(\alpha)$ satisfy:

$$\sum_n |c_n(\alpha)|^2 = 1 \tag{38}$$

that is:

$$|c_0(\alpha)|^2 \sum_n \frac{|\alpha|^{2n}}{n!} = |c_0(\alpha)|^2 \, \mathrm{e}^{|\alpha|^2} = 1 \tag{39}$$

With the convention we have chosen:

$$c_0(\alpha) = \mathrm{e}^{-|\alpha|^2/2} \tag{40}$$

and, finally:

$$|\alpha\rangle = \mathrm{e}^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |\varphi_n\rangle \tag{41}$$

2-b. Possible values of the energy in an $|\alpha\rangle$ state

Let us consider an oscillator in the state $|\alpha\rangle$. We see from (41) that a measurement of the energy can yield the result $E_n = (n + 1/2)\hbar\omega$ with the probability:

$$\mathcal{P}_n(\alpha) = |c_n(\alpha)|^2 = \frac{|\alpha|^{2n}}{n!} \mathrm{e}^{-|\alpha|^2} \tag{42}$$

The probability distribution obtained, $\mathcal{P}_n(\alpha)$, is therefore a Poisson distribution. Since:

$$\mathcal{P}_n(\alpha) = \frac{|\alpha|^2}{n} \mathcal{P}_{n-1}(\alpha) \tag{43}$$

it is easy to verify that $\mathcal{P}_n(\alpha)$ reaches its maximum value when:

$$n = \text{ the integral part of } |\alpha|^2 \tag{44}$$

To calculate the mean value $\langle H \rangle_\alpha$ of the energy, we can use (42) and the expression:

$$\langle H \rangle_\alpha = \sum_n \mathcal{P}_n(\alpha) \left[n + \frac{1}{2} \right] \hbar\omega \tag{45}$$

Nevertheless, it is quicker to notice that, since the adjoint relation of (33) is:

$$\langle \alpha | a^\dagger = \alpha^* \langle \alpha | \tag{46}$$

we have:

$$\langle\alpha|a^\dagger a|\alpha\rangle = \alpha^*\alpha \tag{47}$$

and therefore:

$$\langle H\rangle_\alpha = \hbar\omega\langle\alpha|\left[a^\dagger a + \frac{1}{2}\right]|\alpha\rangle = \hbar\omega\left[|\alpha|^2 + \frac{1}{2}\right] \tag{48}$$

Comparing this result to (44), we see that, when $|\alpha| \gg 1$, $\langle H\rangle_\alpha$ is not very different, in relative value, from the energy E_n which corresponds to the maximum value of $\mathcal{P}_n(\alpha)$.

Let us calculate the mean value $\langle H^2\rangle_\alpha$:

$$\langle H^2\rangle_\alpha = \hbar^2\omega^2\langle\alpha|\left(a^\dagger a + \frac{1}{2}\right)^2|\alpha\rangle \tag{49}$$

Using (33) and the fact that $[a, a^\dagger] = 1$, we easily obtain:

$$\langle H^2\rangle_\alpha = \hbar^2\omega^2\left[|\alpha|^4 + 2|\alpha|^2 + \frac{1}{4}\right] \tag{50}$$

from which we get:

$$\Delta H_\alpha = \hbar\omega|\alpha| \tag{51}$$

If we compare (48) and (51), we see that, if $|\alpha|$ is very large, we have:

$$\frac{\Delta H_\alpha}{\langle H\rangle_\alpha} \simeq \frac{1}{|\alpha|} \ll 1 \tag{52}$$

The energy of the state $|\alpha\rangle$ is very well-defined, since its relative uncertainty is very small.

Comment:

Since:

$$H = \left(N + \frac{1}{2}\right)\hbar\omega \tag{53}$$

we immediately obtain from (48) and (51):

$$\begin{cases} \langle N\rangle_\alpha = |\alpha|^2 \\ \Delta N_\alpha = |\alpha| \end{cases} \tag{54}$$

Thus we see that, to obtain a quasi-classical state, we must linearly superpose a very large number of states $|\varphi_n\rangle$ since $\Delta N_\alpha \gg 1$. However, the relative value of the dispersion over N is very small:

$$\frac{\Delta N_\alpha}{\langle N\rangle_\alpha} = \frac{1}{|\alpha|} \ll 1 \tag{55}$$

2-c. Calculation of $\langle X\rangle$, $\langle P\rangle$, ΔX and ΔP in an $|\alpha\rangle$ state

The mean values $\langle X\rangle$ and $\langle P\rangle$ can be obtained by expressing X and P in terms of a and $a^\dagger$ [formula (13)] and using (33) and (46). We obtain:

$$\begin{aligned}\langle X\rangle_\alpha &= \langle\alpha|X|\alpha\rangle = \sqrt{\frac{2\hbar}{m\omega}}\,\mathrm{Re}(\alpha)\\ \langle P\rangle_\alpha &= \langle\alpha|P|\alpha\rangle = \sqrt{2m\hbar\omega}\,\mathrm{Im}(\alpha)\end{aligned} \tag{56}$$

An analogous calculation yields:

$$\begin{aligned}\langle X^2\rangle_\alpha &= \frac{\hbar}{2m\omega}\left[(\alpha+\alpha^*)^2+1\right]\\ \langle P^2\rangle_\alpha &= \frac{m\hbar\omega}{2}\left[1-(\alpha-\alpha^*)^2\right]\end{aligned} \tag{57}$$

and therefore:

$$\begin{aligned}\Delta X_\alpha &= \sqrt{\frac{\hbar}{2m\omega}}\\ \Delta P_\alpha &= \sqrt{\frac{m\hbar\omega}{2}}\end{aligned} \tag{58}$$

Neither ΔX_α nor ΔP_α, depends on α. Note also that $\Delta X\cdot\Delta P$ takes on its minimum value:

$$\Delta X_\alpha\cdot\Delta P_\alpha = \hbar/2 \tag{59}$$

2-d. The operator $D(\alpha)$: the wave functions $\Psi_\alpha(x)$

Consider the operator $D(\alpha)$ defined by:

$$D(\alpha) = \mathrm{e}^{\alpha a^\dagger - \alpha^* a} \tag{60}$$

This operator is unitary, since:

$$D^\dagger(\alpha) = \mathrm{e}^{\alpha^* a - \alpha a^\dagger} \tag{61}$$

immediately implies:

$$D(\alpha)\,D^\dagger(\alpha) = D^\dagger(\alpha)D(\alpha) = 1 \tag{62}$$

Since the commutator of the operators $\alpha a^\dagger$ and $-\alpha^* a$ is equal to $\alpha^*\alpha$, which is a number, we can use relation (63) of Complement $\mathrm{B_{II}}$ to write:

$$D(\alpha) = \mathrm{e}^{-|\alpha|^2/2}\,\mathrm{e}^{\alpha a^\dagger}\,\mathrm{e}^{-\alpha^* a} \tag{63}$$

Now let us calculate the ket $D(\alpha)|\varphi_0\rangle$; since:

$$\begin{aligned}\mathrm{e}^{-\alpha^* a}|\varphi_0\rangle &= \left[1-\alpha^* a+\frac{\alpha^{*2}}{2!}a^2+\ldots\right]|\varphi_0\rangle\\ &= |\varphi_0\rangle\end{aligned} \tag{64}$$

then:

$$\begin{aligned} D(\alpha)|\varphi_0\rangle &= \mathrm{e}^{-|\alpha|^2/2}\,\mathrm{e}^{\alpha a^\dagger}|\varphi_0\rangle \\ &= \mathrm{e}^{-|\alpha|^2/2}\sum_n \frac{(\alpha a^\dagger)^n}{n!}|\varphi_0\rangle \\ &= \mathrm{e}^{-|\alpha|^2/2}\sum_n \frac{\alpha^n}{\sqrt{n!}}|\varphi_n\rangle \end{aligned} \tag{65}$$

Comparing (41) and (65), we see that:

$$|\alpha\rangle = D(\alpha)|\varphi_0\rangle \tag{66}$$

$D(\alpha)$ is therefore the unitary transformation which transforms the ground state $|\varphi_0\rangle$ into the quasi-classical state $|\alpha\rangle$.

Formula (66) will enable us to obtain the wave function:

$$\psi_\alpha(x) = \langle x|\alpha\rangle \tag{67}$$

which characterizes the quasi-classical state $|\alpha\rangle$ in the $\{|x\rangle\}$ representation. To calculate:

$$\psi_\alpha(x) = \langle x|D(\alpha)|\varphi_0\rangle \tag{68}$$

we shall write the operator $\alpha a^\dagger - \alpha^* a$ in terms of X and P:

$$\alpha a^\dagger - \alpha^* a = \sqrt{\frac{m\omega}{\hbar}}\left(\frac{\alpha-\alpha^*}{\sqrt{2}}\right)X - \frac{i}{\sqrt{m\hbar\omega}}\left(\frac{\alpha+\alpha^*}{\sqrt{2}}\right)P \tag{69}$$

Using formula (63) of Complement B_{II} again, we obtain:

$$D(\alpha) = \mathrm{e}^{\alpha a^\dagger - \alpha^* a} = \mathrm{e}^{\sqrt{\frac{m\omega}{\hbar}}\frac{\alpha-\alpha^*}{\sqrt{2}}X}\,\mathrm{e}^{-\frac{i}{\sqrt{m\hbar\omega}}\frac{\alpha+\alpha^*}{\sqrt{2}}P}\,\mathrm{e}^{\frac{\alpha^{*2}-\alpha^2}{4}} \tag{70}$$

Substituting this result into (68), we find:

$$\begin{aligned} \psi_\alpha(x) &= \mathrm{e}^{\frac{\alpha^{*2}-\alpha^2}{4}}\langle x|\mathrm{e}^{\sqrt{\frac{m\omega}{\hbar}}\frac{\alpha-\alpha^*}{\sqrt{2}}X}\,\mathrm{e}^{-\frac{i}{\sqrt{m\hbar\omega}}\frac{\alpha+\alpha^*}{\sqrt{2}}P}|\varphi_0\rangle \\ &= \mathrm{e}^{\frac{\alpha^{*2}-\alpha^2}{4}}\,\mathrm{e}^{\sqrt{\frac{m\omega}{\hbar}}\frac{\alpha-\alpha^*}{\sqrt{2}}x}\langle x|\mathrm{e}^{-\frac{i}{\sqrt{m\hbar\omega}}\frac{\alpha+\alpha^*}{\sqrt{2}}P}|\varphi_0\rangle \end{aligned} \tag{71}$$

Now, the operator $\mathrm{e}^{-i\lambda P/\hbar}$ is the translation operator of λ along Ox (*cf.* Complement E_{II}):

$$\langle x|\mathrm{e}^{-\frac{i}{\sqrt{m\hbar\omega}}\frac{\alpha+\alpha^*}{\sqrt{2}}P} = \langle x - \sqrt{\tfrac{\hbar}{2m\omega}}(\alpha+\alpha^*)| \tag{72}$$

Relation (71) therefore yields:

$$\psi_\alpha(x) = \mathrm{e}^{\frac{\alpha^{*2}-\alpha^2}{4}}\,\mathrm{e}^{\sqrt{\frac{m\omega}{\hbar}}\frac{\alpha-\alpha^*}{\sqrt{2}}x}\,\varphi_0\left(x - \sqrt{\tfrac{\hbar}{2m\omega}}(\alpha+\alpha^*)\right) \tag{73}$$

If we write α and α^* in terms of $\langle X\rangle_\alpha$ and $\langle P\rangle_\alpha$ [formulas (56)], $\psi_\alpha(x)$ becomes:

$$\psi_\alpha(x) = \mathrm{e}^{i\theta_\alpha}\,\mathrm{e}^{i\langle P\rangle_\alpha x/\hbar}\,\varphi_0\left(x - \langle X\rangle_\alpha\right) \tag{74}$$

where the global phase factor $\mathrm{e}^{i\theta_\alpha}$ is defined by:

$$\mathrm{e}^{i\theta_\alpha} = \mathrm{e}^{\frac{\alpha^{*2}-\alpha^2}{4}} \tag{75}$$

Relation (74) shows that $\psi_\alpha(x)$ can easily be obtained from the wave function $\varphi_0(x)$ of the ground state of the oscillator: it is sufficient to translate this function along Ox by the quantity $\langle X\rangle_\alpha$ and to multiply it by the oscillating exponential $\mathrm{e}^{i\langle P\rangle_\alpha x/\hbar}$ (since the factor $\mathrm{e}^{i\theta_\alpha}$ plays no physical role, it can be omitted)[1].

If we replace φ_0 in (74) by its explicit expression, we obtain, finally:

$$\psi_\alpha(x) = \mathrm{e}^{i\theta_\alpha}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left\{-\left[\frac{x-\langle X\rangle_\alpha}{2\,\Delta X_\alpha}\right]^2 + i\langle P\rangle_\alpha\frac{x}{\hbar}\right\} \tag{76}$$

The form of the wave packet associated with the $|\alpha\rangle$ state is therefore given by:

$$|\psi_\alpha(x)|^2 = \sqrt{\frac{m\omega}{\pi\hbar}}\exp\left\{-\frac{1}{2}\left[\frac{x-\langle X\rangle_\alpha}{\Delta X_\alpha}\right]^2\right\} \tag{77}$$

For any $|\alpha\rangle$ state, we obtain a Gaussian wave packet. This result is consitent with the fact that the product $\Delta X_\alpha.\Delta P_\alpha$ is always minimal (*cf.* Complement $\mathrm{C_{III}}$).

2-e. The scalar product of two $|\alpha\rangle$ states. Closure relation

The $|\alpha\rangle$ states are eigenvectors of the non-Hermitian operator a. There is therefore no obvious reason for these states to satisfy orthogonality and closure relations. In this section, we shall investigate this question.

First, we shall consider two eigenkets $|\alpha\rangle$ and $|\alpha'\rangle$ of the operator a. Relation (41) gives their scalar product, since:

$$\langle\alpha|\alpha'\rangle = \sum_n c_n^*(\alpha)c_n(\alpha') \tag{78}$$

We therefore have:

$$\begin{aligned}\langle\alpha|\alpha'\rangle &= \mathrm{e}^{-|\alpha|^2/2}\,\mathrm{e}^{-|\alpha'|^2/2}\sum_n\frac{(\alpha^*\alpha')^n}{n!}\\ &= \mathrm{e}^{-|\alpha|^2/2}\,\mathrm{e}^{-|\alpha'|^2/2}\,\mathrm{e}^{\alpha^*\alpha'}\end{aligned} \tag{79}$$

from which we conclude:

$$|\langle\alpha|\alpha'\rangle|^2 = \mathrm{e}^{-|\alpha-\alpha'|^2} \tag{80}$$

[1] The exponential $\mathrm{e}^{i\langle P\rangle_\alpha x/\hbar}$ is obviously not a global phase factor since its value depends on x. The presence of this exponential in (74) insures that the mean value of P in the state described by $\psi_\alpha(x)$ be equal to $\langle P\rangle_\alpha$.

This scalar product is therefore never zero.

However, we shall show that the $|\alpha\rangle$ states do satisfy a closure relation, which is written:

$$\frac{1}{\pi}\iint |\alpha\rangle\langle\alpha| \mathrm{d}\{\mathrm{Re}\,\alpha\}\mathrm{d}\{\mathrm{Im}\,\alpha\} = 1 \tag{81}$$

To do so, we replace $|\alpha\rangle$, on the left-hand side of (81), by its expression (41). This yields:

$$\frac{1}{\pi}\iint \mathrm{e}^{-|\alpha|^2} \sum_n \frac{\alpha^n}{\sqrt{n!}}|\varphi_n\rangle \sum_m \frac{\alpha^{*m}}{\sqrt{m!}}\langle\varphi_m| \mathrm{d}\{\mathrm{Re}\,\alpha\}\mathrm{d}\{\mathrm{Im}\,\alpha\} \tag{82}$$

that is, going into polar coordinates in the complex α plane (setting $\alpha = \rho\,\mathrm{e}^{i\varphi}$):

$$\frac{1}{\pi}\int_0^\infty \rho\,\mathrm{d}\rho \int_0^{2\pi} \mathrm{d}\varphi\,\mathrm{e}^{-\rho^2} \sum_{nm} \mathrm{e}^{i(n-m)\varphi}\frac{\rho^{n+m}}{\sqrt{n!m!}}|\varphi_n\rangle\langle\varphi_m| \tag{83}$$

The integral over φ is easily calculated:

$$\int_0^{2\pi} \mathrm{e}^{i(n-m)\varphi}\,\mathrm{d}\varphi = 2\pi\,\delta_{nm} \tag{84}$$

which yields for (83):

$$\sum_n I_n \frac{1}{n!}|\varphi_n\rangle\langle\varphi_n| \tag{85}$$

with:

$$I_n = 2\int_0^\infty \rho\,\mathrm{d}\rho\,\mathrm{e}^{-\rho^2}\rho^{2n} = \int_0^\infty \mathrm{d}u\,\mathrm{e}^{-u}\,u^n \tag{86}$$

Integrating by parts, we find a recurrence relation for the I_n:

$$I_n = n\,I_{n-1} \tag{87}$$

whose solution is:

$$I_n = n!\,I_0 = n! \tag{88}$$

Substituting this result into (85), we see that the left-hand side of formula (81) can finally be written:

$$\sum_n |\varphi_n\rangle\langle\varphi_n| \tag{89}$$

which proves that formula.

3. Time evolution of a quasi-classical state

Consider a harmonic oscillator which, at the instant $t = 0$, is in a particular $|\alpha\rangle$ state:

$$|\psi(0)\rangle = |\alpha_0\rangle \tag{90}$$

How do its physical properties evolve over time? We already know (*cf.* § 1-b) that the mean values $\langle X\rangle(t)$ and $\langle P\rangle(t)$ always remain equal to the corresponding classical values. We shall now study other interesting properties of the state vector $|\psi(t)\rangle$.

3-a. A quasi-classical state always remains an eigenvector of a

Starting with (41), we can use the general rule to obtain $|\psi(t)\rangle$ when the Hamiltonian is not time-dependent (*cf.* Chap. III, § D-2-a):

$$|\psi(t)\rangle = \mathrm{e}^{-|\alpha|^2/2} \sum_n \frac{\alpha_0^n}{\sqrt{n!}}\, \mathrm{e}^{-iE_n t/\hbar}\, |\varphi_n\rangle$$

$$= \mathrm{e}^{-i\omega t/2}\, \mathrm{e}^{-|\alpha_0|^2/2} \sum_n \frac{\alpha_0^n\, \mathrm{e}^{-in\omega t}}{\sqrt{n!}}\, |\varphi_n\rangle \tag{91}$$

If we compare this result with (41), we see that, to go from $|\psi(0)\rangle = |\alpha_0\rangle$ to $|\psi(t)\rangle$, all we must do is change α_0 to $\alpha_0\, \mathrm{e}^{-i\omega t}$ and multiply the ket obtained by $\mathrm{e}^{-i\omega t/2}$ (which is a global phase factor with no physical consequences):

$$|\psi(t)\rangle = \mathrm{e}^{-i\omega t/2}\, |\alpha = \alpha_0\, \mathrm{e}^{-i\omega t}\rangle \tag{92}$$

In other words, we see that a quasi-classical state remains an eigenvector of a for all time, with an eigenvalue $\alpha_0\, \mathrm{e}^{-i\omega t}$ which is nothing more than the parameter $\alpha(t)$ of Figure 1 (corresponding to the point M), which characterizes at all times the classical oscillator whose motion is reproduced by the state $|\psi(t)\rangle$.

3-b. Evolution of physical properties

Using (92) and changing α to $\alpha_0\, \mathrm{e}^{-i\omega t}$ in (56), we immediately obtain:

$$\begin{cases} \langle X\rangle(t) = \sqrt{\dfrac{2\hbar}{m\omega}}\ \mathrm{Re}[\alpha_0\, \mathrm{e}^{-i\omega t}] \\ \langle P\rangle(t) = \sqrt{2m\hbar\omega}\ \mathrm{Im}[\alpha_0\, \mathrm{e}^{-i\omega t}] \end{cases} \tag{93}$$

As predicted, these equations are similar to the classical equations (9).

The average energy of the oscillator is time-independent:

$$\langle H\rangle = \hbar\omega \left[|\alpha_0|^2 + \frac{1}{2}\right] \tag{94}$$

According to (51) and (58), the root mean square deviations ΔH, ΔX and ΔP are equal to:

$$\Delta H = \hbar\omega|\alpha_0| \tag{95}$$

and:

$$\begin{cases} \Delta X = \sqrt{\dfrac{\hbar}{2m\omega}} \\ \Delta P = \sqrt{\dfrac{m\hbar\omega}{2}} \end{cases} \tag{96}$$

ΔX and ΔP are not time-dependent, the wave packet remains a minimum wave packet for all time.

3-c. Motion of the wave packet

Let us calculate the wave function at time t:

$$\psi(x,t) = \langle x|\psi(t)\rangle \tag{97}$$

where $|\psi(t)\rangle$ is given by (92). From (76), we obtain:

$$\psi(x,t) = \mathrm{e}^{i\theta_\alpha} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-i\omega t/2}\, \mathrm{e}^{i\frac{x\langle P\rangle(t)}{\hbar}}\, \mathrm{e}^{-\left[\frac{x-\langle X\rangle(t)}{2\Delta X}\right]^2} \tag{98}$$

At t, the wave packet is still Gaussian. Its form does not vary with time, since:

$$|\psi(x,t)|^2 = |\varphi_0\,[x - \langle X\rangle(t)]|^2 \tag{99}$$

Thus, it remains "minimum" for all time [*cf.* (96)].

Figure 2 shows the motion of the wave packet, which performs a periodic oscillation ($T = 2\pi/\omega$) along the Ox axis, without becoming distorted. We saw in Complement G_I that a Gaussian wave packet, when it is free, becomes distorted as it propagates, since its width varies ("spreading" of the wave packet). We see here that nothing of the sort occurs when a wave packet is subject to the influence of a parabolic potential. Physically, this result arises from the fact that the tendency of the wave packet to spread is compensated by the potential, whose effect is to push the wave packet towards the origin from regions where $V(x)$ is large.

What happens to these results when $|\alpha|$ is very large? The root mean square deviations ΔX and ΔP do not change, as is shown by (96). On the other hand, the oscillation amplitudes of $\langle X\rangle(t)$ and $\langle P\rangle(t)$ become much larger than ΔX and ΔP. By choosing a sufficiently large value of $|\alpha|$, one can obtain a quantum mechanical motion for which the position and momentum of the oscillator are, in relative value, as well-defined as might be desired. Therefore when $|\alpha| \gg 1$, an $|\alpha\rangle$ state describes very well the motion of a macroscopic harmonic oscillator, for which the position, the momentum and the energy can be considered to be classical quantities.

4. Example: quantum mechanical treatment of a macroscopic oscillator

Let us consider a concrete example: a macroscopic body of mass $m = 1$ kg suspended from a rope of length $l = 0.1$ m and placed in the gravitational field ($g \simeq 10$ m/s^2). We know that, for small oscillations, the period T of the motion is given by:

$$T = 2\pi\sqrt{\frac{l}{g}} \tag{100}$$

In our case, we obtain:

$$\begin{cases} T \simeq 0.63 \text{ s} \\ \omega = 10 \text{ rd/s} \end{cases} \tag{101}$$

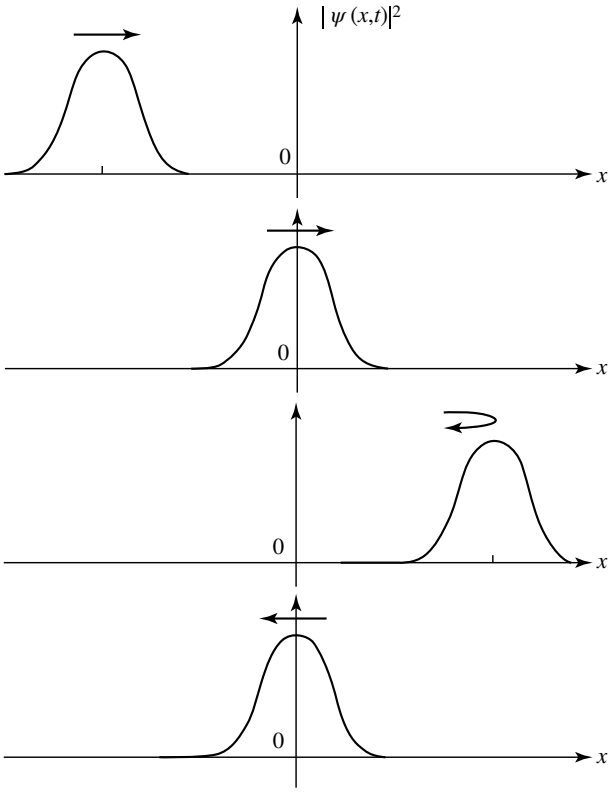


Figure 2: Motion of the Gaussian wave packet associated with an $|\alpha\rangle$ state: under the effect of the parabolic potential $V(x)$, the wave packet oscillates without becoming distorted.

Let us now assume that this oscillator performs a periodic motion of amplitude $x_M = 1$ cm. What is the quantum mechanical state that best represents its oscillation?

We have seen that this state is an $|\alpha\rangle$ state in which, according to (93), α satisfies the relation:

$$|\alpha| = \sqrt{\frac{m\omega}{2\hbar}} x_M \tag{102}$$

that is, in our case:

$$|\alpha| \simeq \sqrt{5} \times 10^{15} \simeq 2.2 \times 10^{15} \gg 1 \tag{103}$$

(the argument of α is determined by the initial phase of the motion).

The root mean square deviations ΔX and ΔP are then:

$$\begin{cases} \Delta X = \sqrt{\dfrac{\hbar}{2m\omega}} \simeq 2.2 \times 10^{-18}\,\text{m} \ll x_M \\ \Delta P = \sqrt{\dfrac{m\hbar\omega}{2}} \simeq 2.2 \times 10^{-17}\,\text{kg m/s} \end{cases} \tag{104}$$

The root mean square deviation ΔV of the velocity is equal to:

$$\Delta V \simeq 2.2 \times 10^{-17}\,\text{m/s} \tag{105}$$

Since the maximum velocity of the oscillator is 0.1 m/s, we see that the uncertainties of its position and its velocity are completely negligible compared to the quantities involved in the problem. For example, ΔX is less than a Fermi (10^{-15} m), that is, the approximate size of a nucleus. Measuring a macroscopic length with this accuracy is obviously out of the question.

Finally, the energy of the oscillator is known with an excellent relative accuracy, since, according to (52):

$$\frac{\Delta H}{\langle H \rangle} \simeq \frac{1}{|\alpha|} \simeq 0.4 \times 10^{-15} \ll 1 \tag{106}$$

The laws of classical mechanics are therefore quite adequate for the study of the evolution of the macroscopic oscillator.

References and suggestions for further reading:

Glauber's lectures in (15.2).

Complement H_V

Normal vibrational modes of two coupled harmonic oscillators

1 Vibration of the two coupled in classical mechanics 599

1-a Equations of motion 599

1-b Solving the equations of motion 600

1-c The physical meaning of each of the modes 601

1-d Motion of the system in the general case 602

2 Vibrational states of the system in quantum mechanics . . 605

2-a Commutation relations 605

2-b Transformation of the Hamiltonian operator 606

2-c Stationary states of the system 607

2-d Evolution of the mean values 609

This complement is devoted to the study of the motion of two coupled (one-dimensional) harmonic oscillators. Such a study is of interest because it permits the introduction, in a very simple case, of a physically important concept: that of *normal vibrational modes*. This concept, encountered in quantum mechanics as well as in classical mechanics, appears in numerous problems: for example, in the study of atomic vibrations in a crystal (*cf.* Complement J_V) and of the vibrations of electromagnetic radiation (*cf.* Complement K_V).

1. Vibration of the two coupled in classical mechanics

1-a. Equations of motion

Let us therefore consider two particles (1) and (2), of the same mass m, moving along the Ox axis, with abscissas x_1 and x_2. To begin, we assume their potential energy to be:

$$U_0(x_1, x_2) = \frac{1}{2}m\omega^2(x_1 - a)^2 + \frac{1}{2}m\omega^2(x_2 + a)^2 \tag{1}$$

When $x_1 = a$ and $x_2 = -a$, the potential energy $U_0(x_1, x_2)$ is minimal, and the two particles are in stable equilibrium. If the particles move from these equilibrium positions, they are subjected to the forces F_1 and F_2, respectively:

$$\begin{cases} F_1 = -\dfrac{\partial}{\partial x_1}U_0(x_1, x_2) = -m\omega^2(x_1 - a) \\ F_2 = -\dfrac{\partial}{\partial x_2}U_0(x_1, x_2) = -m\omega^2(x_2 + a) \end{cases} \tag{2}$$

and their motion is given by:

$$\begin{cases} m\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}x_1(t) = -m\omega^2(x_1 - a) \\ m\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}x_2(t) = -m\omega^2(x_2 + a) \end{cases} \tag{3}$$

Each particle therefore follows an independent sinusoidal motion, centered at its equilibrium position. The amplitude of the motion of each particle is arbitrary[1] and can be fixed by a suitable choice of the initial conditions.

Now let us assume the potential energy $U(x_1, x_2)$ of the two particles to be:

$$U(x_1, x_2) = U_0(x_1, x_2) + V(x_1, x_2) \tag{4}$$

with:

$$V(x_1, x_2) = \lambda m\omega^2(x_1 - x_2)^2 \tag{5}$$

(where λ is a dimensionless positive constant which we shall call the "coupling constant"). To the forces F_1 and F_2 written in (2), we must add, respectively, the forces F_1' and F_2' given by:

$$\begin{cases} F_1' = -\dfrac{\partial}{\partial x_1}V(x_1, x_2) = 2\lambda m\omega^2(x_2 - x_1) \\ F_2' = -\dfrac{\partial}{\partial x_2}V(x_1, x_2) = 2\lambda m\omega^2(x_1 - x_2) \end{cases} \tag{6}$$

We see that the introduction of $V(x_1, x_2)$ takes into account an attractive force between the particles, which is proportional to the distance between them. The two particles (1) and (2) are therefore no longer independent; what is their motion now? Before attacking this problem from a quantum mechanical point of view, we shall recall the results given by classical mechanics.

1-b. Solving the equations of motion

In the presence of the coupling $V(x_1, x_2)$, we must replace (3) by the system of coupled differential equations:

$$\begin{cases} m\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}x_1(t) = -m\omega^2(x_1 - a) + 2\lambda m\omega^2(x_2 - x_1) \\ m\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}x_2(t) = -m\omega^2(x_2 + a) + 2\lambda m\omega^2(x_1 - x_2) \end{cases} \tag{7}$$

We know how to solve such a system (see for example Chapter IV, § C-3-a). We diagonalize the matrix K of the coefficients appearing on the right-hand side of (7):

$$K = -m\omega^2\begin{pmatrix} 1 + 2\lambda & -2\lambda \\ -2\lambda & 1 + 2\lambda \end{pmatrix} \tag{8}$$

[1] Of course, the choice of the potential (1) implies that we are not taking into account the collisions that could occur if sufficiently large amplitudes were chosen.

We then replace $x_1(t)$ and $x_2(t)$ by linear combinations of these two functions (given by the eigenvectors of K) whose time dependence obeys uncoupled linear differential equations (with coefficients which are the eigenvalues of K).

In this case, these linear combinations are:

$$x_G(t) = \frac{1}{2}\left[x_1(t) + x_2(t)\right] \tag{9}$$

(the position of the center of mass of the two particles) and:

$$x_R(t) = x_1(t) - x_2(t) \tag{10}$$

(the abscissa of the "relative particle"). Substituting (9) and (10) into (7), we obtain (taking the sum and the difference):

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}x_G(t) = -\omega^2 x_G(t)$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}x_R(t) = -\omega^2\left[x_R(t) - 2a\right] - 4\lambda\omega^2 x_R(t) \tag{11}$$

These equations can be integrated immediately:

$$\begin{cases} x_G(t) = x_G^0 \cos(\omega_G t + \theta_G) \\ x_R(t) = \dfrac{2a}{1+4\lambda} + x_R^0 \cos(\omega_R t + \theta_R) \end{cases} \tag{12}$$

with:

$$\begin{cases} \omega_G = \omega \\ \omega_R = \omega\sqrt{1+4\lambda} \end{cases} \tag{13}$$

x_G^0, x_R^0, θ_G and θ_R are integration constants fixed by the initial conditions. To obtain the motion of particles (1) and (2), all we must do is invert formulas (9) and (10):

$$\begin{cases} x_1(t) = x_G(t) + \dfrac{1}{2}x_R(t) \\ x_2(t) = x_G(t) - \dfrac{1}{2}x_R(t) \end{cases} \tag{14}$$

and substitute (12) into these equations.

1-c. The physical meaning of each of the modes

Through the change of functions performed in (9) and (10), we have been able to find the motion of the two interacting particles by associating with them two *fictitious particles* (G) and (R), of abscissas $x_G(t)$ and $x_R(t)$. These fictitious particles do not

interact; their motions are independent, so their amplitudes and phases can be fixed arbitrarily by a suitable choice of the initial conditions. For example, it is possible to require one of the two fictitious particles to be motionless without this being the case for the other one: we then say that a *vibrational mode* of the system is excited. It must be understood that, in a vibrational mode, the real particles (1) and (2) are both in motion with the same angular frequency (ω_R or ω_G, depending on the mode). No solution of the equations of motion exists for which one of the two real particles (1) or (2) remains motionless while the other one vibrates. If, at the instant $t = 0$, one were to give an initial velocity to only one of the two particles, (1) or (2), the coupling force would set the other one in motion (*cf.* discussion of § 1-d below).

The simplest case is of course the one in which neither of the two modes is excited. In formulas (12) such a situation corresponds to $x_G^0 = x_R^0 = 0$; $x_G(t)$ and $x_R(t)$ then always remain equal to zero and $2a/(1+4\lambda)$ respectively, which, according to (14), yields:

$$x_1 = -x_2 = \frac{a}{1+4\lambda} \tag{15}$$

The system does not oscillate and the two particles (1) and (2) remain motionless in their new equilibrium positions given by (15) (it can be verified that, for these values of x_1 and x_2, the forces exerted on the particles are zero; the fact that these equilibrium positions are closer in the presence of the coupling than when $\lambda = 0$ is due to their mutual attraction).

To excite only the mode corresponding to $x_G(t)$, one places the two particles (1) and (2) at the initial instant at the same distance $2a/(1+4\lambda)$ as in the preceding case, and one gives them equal velocities. One then finds that $x_R(t)$ remains equal to $2a/(1+4\lambda)$ (the initial conditions require x_R^0 to be zero). The two particles move "in unison", performing the same motion without the distance between them varying. For this mode, the two-particle system can be treated like a single undeformable particle of mass $2m$ on which is exerted the force $F_1 + F_2 = -2m\omega^2 x_G(t)$. We then see why the angular frequency of this mode is $\omega_G = \omega$ [*cf.* formula (A-3) of Chapter V].

To excite only the mode corresponding to $x_R(t)$, one chooses an initial state in which the positions and initial velocities of the two particles are opposite. One then finds that, at every subsequent instant, $x_G(t) = 0$, and the two particles move symmetrically with respect to the origin O. For this mode, the distance $(x_2 - x_1)$ varies and the attractive force between the two particles enters into the equations of motion; this is the reason why the angular frequency of this mode is not ω but $\omega_R = \omega\sqrt{1+4\lambda}$.

The dynamical variables $x_G(t)$ and $x_R(t)$ associated with the independent modes, that is, with the fictitious particles (G) and (R), are called *normal variables.*

1-d. Motion of the system in the general case

In the general case, both modes are excited and the positions $x_1(t)$ and $x_2(t)$ are both given by the superposition of two oscillations of different frequencies ω_G and ω_R [*cf.* formulas (14)]. The motion of the system is not periodic, except in the case in which the ratio ω_G/ω_R is rational[2].

[2] If $\omega_G/\omega_R = 1/\sqrt{1+4\lambda}$ is equal to an irreducible rational fraction r_1/r_2, the period of the motion is: $T = 2\pi r_1/\omega_G = 2\pi r_2/\omega_R$.

Let us investigate, for example, what happens if, at the initial time t_0, particle (1) is motionless at its equilibrium position $x_1 = a/(1 + 4\lambda)$, while particle (2) has a non-zero velocity (this is, in classical mechanics, the analogue of the problem studied in § C-3-b of Chapter IV). In the absence of coupling, particle (2) would oscillate alone and particle (1) would remain motionless. We shall show that the coupling sets particle (1) in motion. Two different angular frequencies ω_G and ω_R appear in the time evolution of $x_1(t)$ and $x_2(t)$. The two corresponding oscillations give rise to a beat phenomenon (Fig. 1), whose frequency is:

$$\nu = \frac{\omega_R - \omega_G}{2\pi} = \frac{\omega}{2\pi}[\sqrt{1 + 4\lambda} - 1] \tag{16}$$

If the coupling is weak ($\lambda \ll 1$), this frequency $\nu \simeq \lambda\omega/\pi$ is negligible with respect to ω_R and ω_G. In this case, as long as $(t - t_0) \ll 1/\nu$, particle (2) is practically the only one to oscillate; the vibrational energy is then slowly transferred to particle (1), whose amplitude of oscillation increases, while that of (2) decreases. After a certain time, the original situation is inverted: particle (1) oscillates strongly while particle (2) is practically motionless. Then the amplitude of (1) slowly decreases and that of (2) increases until the energy is again almost entirely localized in oscillator (2). The same process is repeated indefinitely. The effect of a weak coupling is to cause the energy of the oscillator associated with particle (1) to be constantly transferred to the one associated with particle (2) and vice versa, with a frequency proportional to the intensity of the coupling.

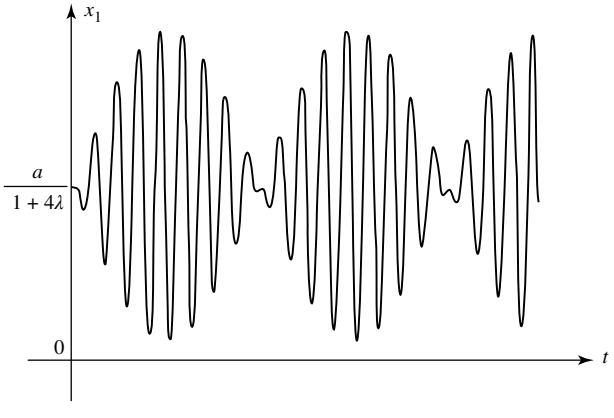


Figure 1: Oscillations of the position of particle (1), assumed to be motionless at its equilibrium position at $t = 0$, particle (2) having an initial velocity. A beat phenomenon is produced between the two modes, and the amplitude of the oscillations of particle (1) varies over time.

Comments:

(*i*) If p_1 and p_2 are the respective momenta of particles (1) and (2), the classical Hamiltonian of the system under study can be written:

$$\mathcal{H}(x_1, x_2, p_1, p_2) = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + U_0(x_1, x_2) + V(x_1, x_2) \tag{17}$$

If we set:

$$\begin{cases} p_G(t) = p_1(t) + p_2(t) \\ p_R(t) = \dfrac{1}{2}[p_1(t) - p_2(t)] \end{cases} \tag{18}$$

and:

$$\begin{cases} \mu_G = 2m \\ \mu_R = \dfrac{m}{2} \end{cases} \tag{19}$$

it can be verified that $\mathcal{H}$ becomes:

$$\mathcal{H} = \frac{p_G^2}{2\mu_G} + \frac{1}{2}\mu_G\omega_G^2 x_G^2 + \frac{p_R^2}{2\mu_R} + \frac{1}{2}\mu_R\omega_R^2\left(x_R - \frac{2a}{1+4\lambda}\right)^2 + m\omega^2 a^2 \frac{4\lambda}{1+4\lambda} \tag{20}$$

By a suitable change in the energy origin, one can eliminate the last term of this expression, which is constant. $\mathcal{H}$ can then be seen to be the sum of two energies, each of which corresponds to one of the modes. Unlike the situation in (17), in which the terms in x_1x_2 of $V(x_1, x_2)$ are responsible for a coupling between the particles, there is no coupling term in (20) between the modes: they are indeed independent.

(*ii*) We have assumed, for simplicity, the masses m_1 and m_2 of particles (1) and (2) to be equal. It is easy to eliminate this hypothesis by replacing (9), (10), (18) and (19) by:

$$\begin{cases} x_G(t) = \dfrac{m_1x_1(t) + m_2x_2(t)}{m_1 + m_2} \\ p_G(t) = p_1(t) + p_2(t) \\ \mu_G = m_1 + m_2 \end{cases} \tag{21}$$

(the position, momentum and mass associated with the center of mass) and:

$$\begin{cases} x_R(t) = x_1(t) - x_2(t) \\ p_R(t) = \dfrac{m_2p_1(t) - m_1p_2(t)}{m_1 + m_2} \\ \mu_R = \dfrac{m_1m_2}{m_1 + m_2} \end{cases} \tag{22}$$

(the position, momentum and mass of the "relative particle"). One then finds a result analogous to (20).

(*iii*) In the absence of coupling, the two modes have the same angular frequency ω; in the presence of coupling, two different angular frequencies ω_G and ω_R appear. This is an example of a result which is often found in physics: the effect of a coupling between two oscillations is, in most cases, to separate their normal frequencies (the same phenomenon would occur here if the two oscillators originally had different angular frequencies). If, instead of two, we have an infinite number of oscillators (which, if isolated, would have the same frequency), we shall see in Complement J_V that the effect of the coupling is to create an infinite number of different frequencies for the modes.

2. Vibrational states of the system in quantum mechanics

Let us now reconsider the problem from a quantum mechanical point of view. We must now replace the positions $x_1(t)$, $x_2(t)$ and the momenta $p_1(t)$, $p_2(t)$ of the particles by operators, which we shall denote, respectively, by X_1, X_2, P_1, P_2. We then introduce, by analogy with (9), (10) and (18), the observables:

$$\begin{cases} X_G = \dfrac{1}{2}(X_1 + X_2) \\ P_G = P_1 + P_2 \end{cases} \tag{23}$$

$$\begin{cases} X_R = X_1 - X_2 \\ P_R = \dfrac{1}{2}(P_1 - P_2) \end{cases} \tag{24}$$

To see if the operator H, the Hamiltonian of the system, can be put into a form analogous to (20), we shall begin by examining the commutation relations of X_G, P_G, X_R et P_R.

2-a. Commutation relations

Since all the observables concerning only particle (1) commute with those concerning particle (2), the only non-zero commutators involving X_1, X_2, P_1 and P_2 are:

$$\begin{aligned} [X_1, P_1] &= i\hbar \\ [X_2, P_2] &= i\hbar \end{aligned} \tag{25}$$

In particular, X_1 commutes with X_2, and we see immediately that:

$$[X_G, X_R] = 0 \tag{26}$$

Similarly:

$$[P_G, P_R] = 0 \tag{27}$$

Calculating the commutator $[X_G, P_G]$, we obtain:

$$[X_G, P_G] = \frac{1}{2}\{[X_1, P_1] + [X_1, P_2] + [X_2, P_1] + [X_2, P_2]\}$$

$$= \frac{1}{2}\{i\hbar + i\hbar\} = i\hbar \tag{28}$$

Similarly, one finds:

$$[X_R, P_R] = i\hbar \tag{29}$$

The two commutators $[X_G, P_R]$ and $[X_R, P_G]$ remain to be examined; they are equal to:

$$[X_G, P_R] = \frac{1}{4}\{[X_1, P_1] - [X_1, P_2] + [X_2, P_1] - [X_2, P_2]\}$$

$$= \frac{1}{4}\{i\hbar - i\hbar\} = 0 \tag{30}$$

and, similarly:

$$[X_R, P_G] = 0 \tag{31}$$

We can thus consider X_G, P_G and X_R, P_R to be the position and momentum operators of two distinct particles. Formulas (28) and (29) are the canonical commutation relations for each of these particles. Moreover, relations (26), (27), (30) and (31) express the fact that all the observables concerning one of them commute with all those which concern the other one.

2-b. Transformation of the Hamiltonian operator

In the presence of the coupling $V(X_1, X_2)$, we have:

$$H = T + U \tag{32}$$

with:

$$T = \frac{1}{2m}\left(P_1^2 + P_2^2\right) \tag{33}$$

(the kinetic energy operator) and:

$$U = \frac{1}{2}m\omega^2\left[(X_1 - a)^2 + (X_2 + a)^2 + 2\lambda(X_1 - X_2)^2\right] \tag{34}$$

(the potential energy operator). Since P_1 and P_2 commute, (33) can be transformed as if these operators were numbers; we find:

$$T = \frac{1}{2\mu_G}P_G^2 + \frac{1}{2\mu_R}P_R^2 \tag{35}$$

where μ_G and μ_R are defined in (19). Similarly, since X_1 and X_2 commute, we have, as above [formula (20)]:

$$U = \frac{1}{2}\mu_G\omega_G^2 X_G^2 + \frac{1}{2}\mu_R\omega_R^2\left(X_R - \frac{2a}{1+4\lambda}\right)^2 + m\omega^2 a^2\frac{4\lambda}{1+4\lambda} \tag{36}$$

where ω_G and ω_R are given by (13).

Thus we see that H can be put into a form which is analogous to (20), in which there is no coupling term:

$$H = H_G + H_R + m\omega^2 a^2\frac{4\lambda}{1+4\lambda} \tag{37}$$

with:

$$\begin{cases} H_G = \dfrac{P_G^2}{2\mu_G} + \dfrac{1}{2}\mu_G\ \omega_G^2\ X_G^2 \\ H_R = \dfrac{P_R^2}{2\mu_R} + \dfrac{1}{2}\mu_R\ \omega_R^2\left[X_R - \dfrac{2a}{1+4\lambda}\right]^2 \end{cases} \tag{38}$$

2-c. Stationary states of the system

The state space of the system is the tensor product $\mathcal{E}(1)\otimes\mathcal{E}(2)$ of the state spaces of particles (1) and (2); it is also the tensor product $\mathcal{E}(G)\otimes\mathcal{E}(R)$ of the state spaces of the fictitious particles, the "center of mass" and the "relative particle" associated with each of the two modes. Since H is the sum of two operators H_G and H_R which act only in $\mathcal{E}(G)$ and $\mathcal{E}(R)$ respectively (the constant $m\omega^2a^2\dfrac{4\lambda}{1+4\lambda}$ merely introduces a shift in the energy origin), we know (Chap. II, § F) that we can find a basis of eigenvectors of H in the form:

$$|\varphi\rangle = |\varphi^G\rangle|\varphi^R\rangle \tag{39}$$

where $|\varphi^G\rangle$ and $|\varphi^R\rangle$ are, respectively, eigenvectors of H_G and H_R in $\mathcal{E}(G)$ and $\mathcal{E}(R)$. Now, H_G and H_R are Hamiltonians of one-dimensional harmonic oscillators, whose eigenvectors and eigenvalues we know. If the operators $a_G^\dagger$ and $a_R^\dagger$ are defined by:

$$\begin{cases} a_G^\dagger = \dfrac{1}{\sqrt{2}}\left[\sqrt{\dfrac{\mu_G\omega_G}{\hbar}}X_G - i\dfrac{P_G}{\sqrt{\mu_G\hbar\omega_G}}\right] \\ a_R^\dagger = \dfrac{1}{\sqrt{2}}\left[\sqrt{\dfrac{\mu_R\omega_R}{\hbar}}X'_R - i\dfrac{P_R}{\sqrt{\mu_R\hbar\omega_R}}\right] \end{cases} \tag{40a}$$

with:

$$X'_R = X_R - \frac{2a}{1+4\lambda} \tag{40b}$$

and if $|\varphi_0^G\rangle$ and $|\varphi_0^R\rangle$ denote respectively the ground states of H_G and H_R, the eigenvectors of H_G are the vectors:

$$|\varphi_n^G\rangle = \frac{1}{\sqrt{n!}}(a_G^\dagger)^n|\varphi_0^G\rangle \tag{41}$$

whose eigenvalues are:

$$E_n^G = \left(n + \frac{1}{2}\right)\hbar\omega_G \tag{42}$$

those of H_R being:

$$|\varphi_p^R\rangle = \frac{1}{\sqrt{p!}}(a_R^\dagger)^p|\varphi_0^R\rangle \tag{43}$$

with the eigenvalues:

$$E_p^R = \left(p + \frac{1}{2}\right)\hbar\omega_R \tag{44}$$

Thus we have here a situation which is analogous to the one encountered in the study of a two-dimensional anisotropic (since $\omega_G \neq \omega_R$) harmonic oscillator. The stationary states of the system are given by:

$$|\varphi_{n,p}\rangle = |\varphi_n^G\rangle|\varphi_p^R\rangle = \frac{(a_G^\dagger)^n(a_R^\dagger)^p}{\sqrt{n!p!}}|\varphi_{0,0}\rangle \tag{45}$$

and their energies are:

$$\begin{aligned} E_{n,p} &= E_n^G + E_p^R + m\omega^2 a^2 \frac{4\lambda}{1+4\lambda} \\ &= \left(n + \frac{1}{2}\right)\hbar\omega_G + \left(p + \frac{1}{2}\right)\hbar\omega_R + m\omega^2 a^2 \frac{4\lambda}{1+4\lambda} \end{aligned} \tag{46}$$

The operators a_G and $a_G^\dagger$ or (a_R and $a_R^\dagger$) can thus be seen to be annihilation or creation operators of an energy quantum in the mode corresponding to (G) [or (R)]. We see from (45) that, through the repeated action of $a_G^\dagger$ and $a_R^\dagger$, we can obtain stationary states of the system in which the number of quanta in each mode is arbitrary. The action of $a_G^\dagger$, a_G, $a_R^\dagger$ or a_R on the stationary states $|\varphi_{n,p}\rangle$ is very simple:

$$\begin{aligned} a_G^\dagger|\varphi_{n,p}\rangle &= \sqrt{n+1}\ |\varphi_{n+1,p}\rangle \\ a_G|\varphi_{n,p}\rangle &= \sqrt{n}\ |\varphi_{n-1,p}\rangle \\ a_R^\dagger|\varphi_{n,p}\rangle &= \sqrt{p+1}\ |\varphi_{n,p+1}\rangle \\ a_R|\varphi_{n,p}\rangle &= \sqrt{p}\ |\varphi_{n,p-1}\rangle \end{aligned} \tag{47}$$

In general, there are no degenerate levels since there do not exist two different pairs of integers $\{n, p\}$ and $\{n', p'\}$ such that:

$$n\omega_G + p\omega_R = n'\omega_G + p'\omega_R \tag{48}$$

(except when the ratio $\omega_R/\omega_G = \sqrt{1+4\lambda}$ is rational).

2-d. Evolution of the mean values

The most general state of the system is a linear superposition of stationary states $|\varphi_{n,p}\rangle$:

$$|\varphi(t)\rangle = \sum_{n,p} c_{n,p}(t)|\varphi_{n,p}\rangle \tag{49}$$

with:

$$c_{n,p}(t) = c_{n,p}(0)\,\mathrm{e}^{-iE_{n,p}t/\hbar} \tag{50}$$

According to relations (40) and their adjoints, X_G (X_R) is a linear combination of a_G and $a_G^\dagger$ (of a_R and $a_R^\dagger$). We then see, by using (47), that X_G has non-zero matrix elements between two states $|\varphi_{n,p}\rangle$ and $|\varphi_{n',p'}\rangle$ only when $n - n' = \pm 1$, $p = p'$ (for X'_R, we would have $n = n'$, $p - p' = \pm 1$). From this we deduce that the only Bohr frequencies which can appear in the time evolution of $\langle X_G\rangle(t)$ and $\langle X_R\rangle(t)$ are, respectively[3]:

$$\frac{E_{n\pm1,p} - E_{n,p}}{\hbar} = \pm\omega_G$$

$$\frac{E_{n,p\pm1} - E_{n,p}}{\hbar} = \pm\omega_R \tag{51}$$

Thus we again find that $\langle X_G\rangle(t)$ and $X_R(t)$ oscillate at angular frequencies of ω_G and ω_R, which is consistent with the classical result obtained in § 1-a.

References and suggestions for further reading:

Coupling between two classical oscillators: Berkeley 3 (7.1), §§ 1.4 and 3.3; Alonso and Finn (6.1), Vol. I, § 12.10.

[3]For these frequencies actually to appear, at least one of the products $c^*_{n\pm1,p}\,c_{n,p}$, or $c^*_{n,p\pm1}\,c_{n,p}$ must be different from zero.

Complement J_V

Vibrational modes of an infinite linear chain of coupled harmonic oscillators; phonons

1 Classical treatment 612

1-a Equations of motion 612

1-b Simple solutions of the equations of motion 613

1-c Normal variables 616

1-d Total energy and energy of each of the modes 620

2 Quantum mechanical treatment 622

2-a Stationary states in the absence of coupling 622

2-b Effects of the coupling 623

2-c Normal operators. Commutation relations 624

2-d Stationary states in the presence of coupling 626

3 Application to the study of crystal vibrations: phonons . . 626

3-a Outline of the problem 626

3-b Normal modes. Speed of sound in the crystal 627

In Complement H_V, we studied the motion of a system of two coupled harmonic oscillators. We concluded, in essence, that, while the individual dynamical variables of each oscillator do not evolve independently, it is possible to introduce linear combinations of them (normal variables) which possess the important property of being uncoupled. Such variables describe vibrational normal modes of well-defined frequencies. Expressed in terms of these normal variables, the Hamiltonian of the system appears in the form of a sum of Hamiltonians of independent harmonic oscillators, thus making quantization simple.

In this complement we shall show that these ideas are also applicable to a system formed by an infinite series of identical harmonic oscillators, regularly spaced along an axis, each one coupled to its neighbors.

To do this, we shall determine the various vibrational normal modes of the system and show that each one corresponds to a collective vibration of the system of particles, characterized by an angular frequency Ω and a wave vector k. The process of finding the eigenstates and eigenvalues of the quantum mechanical Hamiltonian is then greatly simplified by the fact that the total energy of the system is the sum of the energies associated with each vibrational normal mode.

The results obtained will enable us to indicate how vibrations propagate in a crystal and to introduce the concept of a phonon, a central idea in solid state physics. Of course, in this complement, we shall emphasize the introduction and quantization of the normal modes, and not the detailed properties of phonons, which would be treated in a solid state physics course.

1. Classical treatment

1-a. Equations of motion

Let us consider an infinite chain of identical one-dimensional harmonic oscillators, each one labeled by an integer q (positive, negative or zero). The particle M_q of mass m which constitutes oscillator (q) has its equilibrium position at the point whose abscissa is ql (Fig. 1), where l is the unit distance of the oscillator chain. We denote by x_q the (algebraic) displacement of oscillator (q) with respect to its equilibrium position. The state of the system at the instant t is defined by specifying the dynamical variables $x_q(t)$ and their time derivatives $\dot{x}_q(t)$ at this instant.

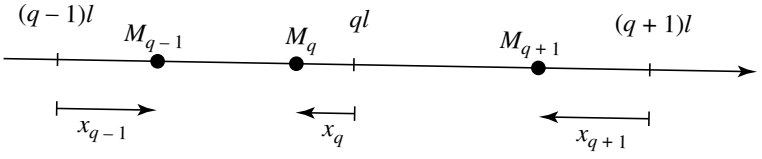


Figure 1: Infinite chain of oscillators; the displacement of the qth particle with respect to its equilibrium position ql is denoted by x_q.

In the absence of interactions between the various particles, the potential energy of the system is:

$$U(..., x_{-1}, x_0, x_{+1}, ...) = \sum_{q=-\infty}^{+\infty} \frac{1}{2} m\omega^2 x_q^2 \tag{1}$$

where ω is the angular frequency of each oscillator. The evolution of the system is then given by the equations:

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2} x_q(t) = -m\omega^2 x_q(t) \tag{2}$$

whose solutions are:

$$x_q(t) = x_q^M \cos(\omega t - \varphi_q) \tag{3}$$

where the integration constants x_q^M and φ_q are fixed by the initial conditions of the motion. The oscillators therefore vibrate independently.

Now imagine that these particles are interacting. For simplicity, we shall assume that one need take into account only the forces exerted on a particle by its two nearest neighbors and that these forces are attractive and proportional to the distance. Thus, particle (q) is subjected to two new attractive forces exerted by particles $(q + 1)$ and $(q - 1)$. These forces are proportional to $|M_q M_{q+1}|$ and $|M_q M_{q-1}|$ (the coefficient of proportionality being the same in both cases). The total force F_q to which particle (q)

is subjected can therefore be written:

$$F_q = -m\omega^2 x_q - m\omega_1^2 \left[ql + x_q - (q+1)l - x_{q+1}\right] - m\omega_1^2 \left[ql + x_q - (q-1)l - x_{q-1}\right]$$
$$= -m\omega^2 x_q - m\omega_1^2 \left(x_q - x_{q+1}\right) - m\omega_1^2 \left(x_q - x_{q-1}\right) \tag{4}$$

where ω_1 is a constant [having the dimensions of an inverse time] which characterizes the intensity of the coupling. Equations (2) must now be replaced by:

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}x_q(t) = -m\omega^2 x_q(t) - m\omega_1^2 \left[2x_q(t) - x_{q+1}(t) - x_{q-1}(t)\right] \tag{5}$$

It can easily be verified that the interaction forces [terms in ω_1^2 of (4)] are derived from the potential energy of the coupling V given by:

$$V(..., x_{-1}, x_0, x_{+1}, ...) = \frac{1}{2}m\omega_1^2 \sum_{q=-\infty}^{+\infty} \left(x_q - x_{q+1}\right)^2 \tag{6}$$

According to (5), the evolution of x_q depends on x_{q+1} and x_{q-1}. Therefore, we must solve an infinite system of coupled differential equations. Before we introduce new variables that allow these equations to be uncoupled, it is interesting to try to find simple solutions of equations (5) and investigate their physical significance.

1-b. Simple solutions of the equations of motion

α. Existence of simple solutions

The infinite chain of coupled oscillators we are studying is analogous to an infinite macroscopic spring. Now, we know that progressive longitudinal waves (corresponding to expansions and compressions) can propagate along this spring. Under the influence of a sinusoidal wave of this type, of wave vector k and angular frequency Ω, the point of the spring whose abscissa is x at equilibrium is found at time t at $x + u(x,t)$, with:

$$u(x,t) = \mu\, \mathrm{e}^{i(kx-\Omega t)} + \mu^*\, \mathrm{e}^{-i(kx-\Omega t)} \tag{7}$$

Such solutions of the equations of motion (5) do indeed exist. However, since the oscillator chain is not a continuous medium, the effects of the wave are observed only at a series of points, corresponding to the abscissas $x = ql$; $u(ql,t)$ thus represents the displacement of oscillator (q) at time t:

$$x_q(t) = u(ql,t) = \mu\, \mathrm{e}^{i(kql-\Omega t)} + \mu^*\, \mathrm{e}^{-i(kql-\Omega t)} \tag{8}$$

It is easy to verify that this expression is a solution of equations (5) if Ω and k satisfy:

$$-m\Omega^2 = -m\omega^2 - m\omega_1^2 \left[2 - \mathrm{e}^{ikl} - \mathrm{e}^{-ikl}\right] \tag{9}$$

Ω is therefore related to k by the "dispersion relation":

$$\Omega(k) = \sqrt{\omega^2 + 4\omega_1^2 \sin^2\left(\frac{kl}{2}\right)} \tag{10}$$

which we shall discuss in detail later (§ 1-b-δ).

β. Physical interpretation

In the solution (8) of the equations of motion, all the oscillators are vibrating at the same frequency $\Omega/2\pi$, with the same amplitude $|2\mu|$, but with a phase that depends periodically on their rest positions. It is as if the displacements of the various oscillators were determined by a progressive sinusoidal wave of wave vector k and phase velocity:

$$v_\varphi(k) = \frac{\Omega(k)}{k} \tag{11}$$

This is easy to show. Using (8), we see that:

$$x_{q_1+q_2}(t) = x_{q_1}\left(t - \frac{q_2 l}{v_\varphi}\right) \tag{12}$$

Thus, oscillator $(q_1 + q_2)$ performs the same motion as oscillator (q_1), shifted by the time taken by the wave to travel, at a velocity v_φ, the distance $q_2 l$ separating the two oscillators. Since all the oscillators are then in motion, solutions (8) are called "collective modes" of vibration of the system.

γ. Possible values of the wave vector k

Consider two values of the wave vector, k and k', which differ by an integral number of $2\pi/l$:

$$k' = k + \frac{2n\pi}{l} \quad \text{with } n \text{ an integer (positive or negative);} \tag{13}$$

We have, obviously:

$$\begin{cases} \mathrm{e}^{ik'ql} = \mathrm{e}^{ikql} \\ \\ \Omega(k') = \Omega(k) \end{cases} \tag{14}$$

where the second relation follows directly from (10).

We see from (8) that the two progressive waves k and k' lead to the same motion for the oscillators and are, consequently, *physically indistinguishable*. Therefore, in the problem we are studying here, it suffices to let k vary over an interval of $2\pi/l$. For reasons of symmetry, we choose:

$$-\frac{\pi}{l} < k \leq \frac{\pi}{l} \tag{15}$$

The corresponding interval is often called the "first Brillouin zone".

δ. Dispersion relation

The dispersion relation (10) which gives the angular frequency $\Omega(k)$ associated with each value of k enables us to study the propagation of vibrations in the system. If, for example, we form a "wave packet" by superposing waves with different wave vectors, we know that it has a group velocity given by:

$$v_G = \frac{\mathrm{d}\Omega(k)}{\mathrm{d}k} \tag{16}$$

which is different from v_φ.

Figure 2 shows the form of the variation of $\Omega(k)$ with respect to k, where k varies within the first Brillouin zone.

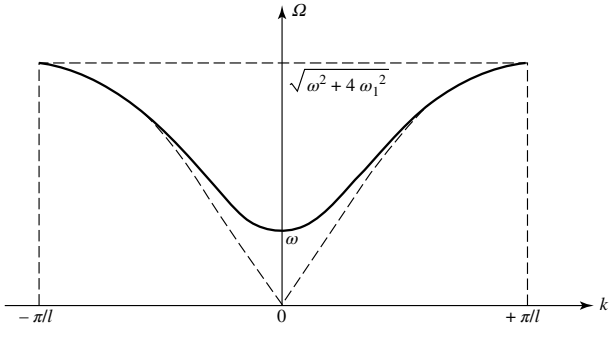


Figure 2: Dispersion relation giving the variation of the angular frequency of the vibrational normal modes with respect to the wave number k in the first Brillouin zone $[-\pi/l, +\pi/l]$. *The dashed line corresponds to the case* $\omega = 0$.

It is clear from this figure that $\Omega(k)$ cannot take on arbitrary values: a vibration of frequency ν can propagate freely in the medium only if ν falls within the "allowed band":

$$\frac{\omega}{2\pi} \leq \nu \leq \frac{\sqrt{\omega^2 + 4\omega_1^2}}{2\pi} \tag{17}$$

The other values of ν correspond to "forbidden bands". The two limiting frequencies of interval (17) are often called "cut-off frequencies".

The mode of lowest angular frequency, $\Omega(0) = \omega$, has a zero wave vector k; it corresponds to an in-phase vibration of all the oscillators, whose particles are moving "together" without changing their relative distances (Fig. 3). This explains why the angular frequency of this mode is the same as in the absence of coupling (*cf.* Complement H_V, § 1-b).

As for the mode of highest angular frequency, $\Omega(\pm\pi/l) = \sqrt{\omega^2 + 4\omega_1^2}$, in the

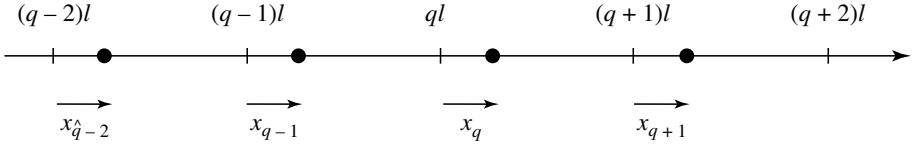


Figure 3: The lowest frequency mode ($k = 0$; $\Omega = \omega$*) corresponds to a displacement of the system of oscillators "as a whole". This is why its frequency does not depend on the coupling V.*

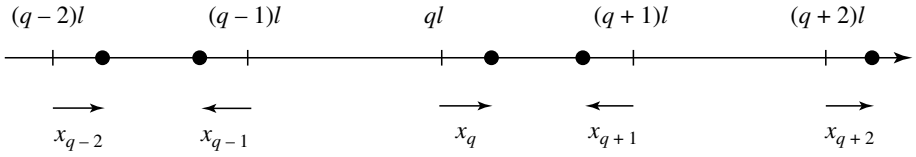


Figure 4: The modes $k = \pm\pi/l$ are those in which two neighboring oscillators are completely out of phase; the coupling V strongly modifies their frequency.

corresponding vibration of the system, two adjacent oscillators are completely out of phase (Fig. 4); the effect of the attractive forces due to the coupling V is then maximal.

1-c. Normal variables

α. Obtaining uncoupled equations

Returning to the equations of motion (5), let us introduce new dynamical variables (linear combinations of the x_q) that evolve independently.

To do so, we multiply both sides of equation (5) by the quantity e^{-iqkl} and sum over q. If we notice that:

$$\sum_{q=-\infty}^{+\infty} x_{q\pm1}\,\mathrm{e}^{-iqkl} = \mathrm{e}^{\pm ikl}\sum_{q=-\infty}^{+\infty} x_{q\pm1}\,\mathrm{e}^{-i(q\pm1)kl}$$

$$= \mathrm{e}^{\pm ikl}\sum_{q=-\infty}^{+\infty} x_q\,\mathrm{e}^{-iqkl} \tag{18}$$

and if we set:

$$\sum_{q=-\infty}^{+\infty} x_q(t)\,\mathrm{e}^{-iqkl} = \xi(k,t) \tag{19}$$

we see that (5) becomes:

$$\frac{\partial^2}{\partial t^2}\xi(k,t) = -\left[\omega^2 + \omega_1^2\left(2 - \mathrm{e}^{ikl} - \mathrm{e}^{-ikl}\right)\right]\xi(k,t) \tag{20}$$

that is, taking (9) into account:

$$\frac{\partial^2}{\partial t^2}\xi(k,t) = -\Omega^2(k)\,\xi(k,t) \tag{21}$$

This relation shows that the time evolution of $\xi(k,t)$ is independent of that of $\xi(k',t)$ for k' different from k. The quantities $\xi(k,t)$ introduced in (19) are therefore completely uncoupled and have a remarkably simple equation of motion.

Comments:

(*i*) Equations (5) are easy to uncouple because the problem is invariant under a translation of the system of oscillators by a quantity $\pm l$ (replacement of q by $q \pm 1$). This invariance is itself due to the fact that the chain is regular and infinite.

(*ii*) In reality, every chain is of course finite, even if it contains a very large number $\mathcal{N}$ of oscillators. To find its vibrational normal modes, one must thus take into account the boundary conditions at the two ends of the chain, and the problem becomes much more complicated (edge effects). Instead of obtaining, as we do here, a continuous infinity of vibrational normal modes corresponding to the various values of k in the first Brillouin zone, one finds a finite number of eigenmodes, equal to the number $\mathcal{N}$ of oscillators. When one is concerned only with the behavior of the chain far from the ends, one often introduces artificial boundary conditions which are different from the real boundary conditions but which have the advantage of simplifying the calculations while conserving the essential physical properties. Thus, one requires the two end oscillators to have the same motion ("periodic" boundary conditions, also called "Born-Von Karman conditions"). We shall have the opportunity to return to this question in connection with the study of other periodic structures (*cf.* Complement $\mathrm{F_{XI}}$; see also § 1-c of Complement $\mathrm{C_{XIV}}$). Therefore we shall not dwell any further on periodic boundary conditions, but shall continue our discussion, confining ourselves to the simple case of an infinite chain.

The function $\xi(k,t)$ introduced in (19) is, by definition, the sum of a Fourier series whose coefficients are the displacements $x_q(t)$. It is a periodic function, of period $2\pi/l$, which is therefore perfectly well-defined when its values in the interval $-\pi/l \leq k \leq \pi/l$ are specified [this is the first Brillouin zone, defined in (15)]. $\xi(k,t)$ depends on the positions of all the oscillators at time t. Conversely, these positions are unambiguously defined when the values of ξ in the interval (15) are given for time t. This is true because it is possible to invert relation (19) since, using:

$$\int_{-\pi/l}^{+\pi/l} \mathrm{d}k\, \mathrm{e}^{i(q-q')kl} = \frac{2\pi}{l}\delta_{qq'} \tag{22}$$

we obtain:

$$x_q(t) = \frac{l}{2\pi}\int_{-\pi/l}^{+\pi/l} \mathrm{d}k\, \xi(k,t)\, \mathrm{e}^{iqkl} \tag{23}$$

Note also that, since the displacements $x_q(t)$ are real, the function $\xi(k,t)$ satisfies:

$$\xi(-k,t) = \xi^*(k,t) \tag{24}$$

Similarly, one can, using the momenta $p_q(t) = m\dot{x}_q(t)$, define the function:

$$\pi(k,t) = \sum_q p_q(t)\, \mathrm{e}^{-ikql} \tag{25}$$

and the $p_q(t)$ can be expressed as:

$$p_q(t) = \frac{l}{2\pi} \int_{-\pi/l}^{+\pi/l} \pi(k,t)\, \mathrm{e}^{ikql}\, \mathrm{d}k \tag{26}$$

The fact that the $p_q(t)$ are real implies that:

$$\pi(-k,t) = \pi^*(k,t) \tag{27}$$

Differentiating both sides of (19) term by term, and using (25) and then (21), we finally obtain:

$$\begin{cases} m\dfrac{\partial}{\partial t}\xi(k,t) = \pi(k,t) & \text{(28a)} \\ \dfrac{\partial}{\partial t}\pi(k,t) = -m\Omega^2(k)\,\xi(k,t) & \text{(28b)} \end{cases}$$

At time t, the dynamical state of the system can be characterized just as well by the specification of the $x_q(t)$ and $p_q(t)$ for all integers (positive, negative or zero), as by that of the "normal variables" $\xi(k,t)$ and $\pi(k,t)$ (where k can take on any value in the first Brillouin zone). The equations of motion (28) of the normal variables corresponding to each value of k describe the evolution of the position and momentum of a harmonic oscillator of mass m and angular frequency $\Omega(k)$; however, ξ and π are complex. *Thus we have reduced the study of an infinite but discrete chain of coupled harmonic oscillators to that of a continuous system of fictitious independent oscillators* (labeled by the index k).

Comments:

Rigorously, these fictitious oscillators are not completely independent since, according to conditions (24) and (27), the initial values $\xi(k,0)$ and $\pi(k,0)$ must satisfy:

$$\begin{aligned} \xi(k,0) &= \xi^*(-k,0) \\ \pi(k,0) &= \pi^*(-k,0) \end{aligned} \tag{29}$$

β. Normal variables $\alpha(k,t)$ associated with the progressive waves

It is convenient (see also § 1-a of Complement G_V) to condense the two normal variables $\xi(k,t)$ and $\pi(k,t)$ into one, $\alpha(k,t)$, defined by:

$$\alpha(k,t) = \frac{1}{\sqrt{2}}\left[\hat{\xi}(k,t) + i\hat{\pi}(k,t)\right] \tag{30}$$

where $\hat{\xi}(k,t)$ and $\hat{\pi}(k,t)$ are dimensionless quantities proportional to $\xi(k,t)$ and $\pi(k,t)$:

$$\begin{cases} \hat{\xi}(k,t) = \beta(k)\,\xi(k,t) \\ \hat{\pi}(k,t) = \dfrac{1}{\hbar\beta(k)}\,\pi(k,t) \end{cases} \tag{31}$$

To simplify the quantum mechanical calculations presented later, we shall set:

$$\beta(k) = \sqrt{\frac{m\Omega(k)}{\hbar}} \tag{32}$$

It is easy to show, using (30), that the two equations (28) are equivalent to the single equation:

$$i\frac{\partial}{\partial t}\alpha(k,t) = \Omega(k)\,\alpha(k,t) \tag{33}$$

which is of first order in t [$\alpha(k,t)$ is completely defined by the specification of $\alpha(k,0)$, while $\xi(k,t)$ depends on $\xi(k,0)$ and $\pi(k,0)$]. The general solution of (33) is:

$$\alpha(k,t) = \alpha(k,0)\,\mathrm{e}^{-i\Omega(k)t} \tag{34}$$

Using (19) and (25), we easily obtain the expression for $\alpha(k,t)$ in terms of the $x_q(t)$ and $p_q(t)$:

$$\alpha(k,t) = \frac{1}{\sqrt{2}}\beta(k)\sum_q \mathrm{e}^{-iqkl}\left[x_q(t) + i\frac{p_q(t)}{m\Omega(k)}\right] \tag{35}$$

Let us show, conversely, that the $x_q(t)$ and $p_q(t)$ can be simply expressed in terms of the $\alpha(k,t)$. According to (24) and (27):

$$\begin{aligned}\alpha^*(-k,t) &= \frac{1}{\sqrt{2}}\left[\hat{\xi}^*(-k,t) - i\,\hat{\pi}^*(-k,t)\right] \\ &= \frac{1}{\sqrt{2}}\left[\hat{\xi}(k,t) - i\,\hat{\pi}(k,t)\right]\end{aligned} \tag{36}$$

From this we deduce:

$$\begin{cases} \hat{\xi}(k,t) = \dfrac{1}{\sqrt{2}}\left[\alpha(k,t) + \alpha^*(-k,t)\right] & \text{(37a)} \\ \hat{\pi}(k,t) = -\dfrac{i}{\sqrt{2}}\left[\alpha(k,t) - \alpha^*(-k,t)\right] & \text{(37b)} \end{cases}$$

which allows us to write formula (23) in the form:

$$x_q(t) = \frac{l}{2\pi\sqrt{2}}\left\{\int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \frac{\alpha(k,t)}{\beta(k)}\,\mathrm{e}^{iqkl}\mathrm{d}k + \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \frac{\alpha^*(-k,t)}{\beta(k)}\,\mathrm{e}^{iqkl}\mathrm{d}k\right\} \tag{38}$$

Changing k to $-k$ in the second integral, we finally obtain [$\beta(k)$ is an even function of k]:

$$x_q(t) = \frac{l}{2\pi\sqrt{2}}\left\{\int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \frac{\alpha(k,t)}{\beta(k)}\,\mathrm{e}^{iqkl}\mathrm{d}k + \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \frac{\alpha^*(k,t)}{\beta(k)}\,\mathrm{e}^{-iqkl}\mathrm{d}k\right\} \tag{39}$$

An analogous calculation, starting with (26), yields:

$$p_q(t) = \frac{l}{2\pi\sqrt{2}} \frac{\hbar}{i} \left\{ \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \beta(k)\, \alpha(k,t)\, \mathrm{e}^{iqkl} \mathrm{d}k - \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \beta(k)\, \alpha^*(k,t)\, \mathrm{e}^{-iqkl} \mathrm{d}k \right\} \tag{40}$$

The state of the system is therefore described just as well by the $\alpha(k,t)$ as by the set of $x_q(t)$ and $p_q(t)$.

If we replace, in (39), $\alpha(k,t)$ by its general expression (34), $x_q(t)$ takes on the form:

$$x_q(t) = \frac{l}{2\pi\sqrt{2}} \left\{ \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \mathrm{d}k \frac{\alpha(k,0)}{\beta(k)}\, \mathrm{e}^{i[qkl - \Omega(k)t]} + c.c. \right\} \tag{41}$$

The most general solution of the problem of the chain of coupled oscillators is therefore a linear superposition of the progressive waves introduced in § 1-b (where the coefficients of this linear combination are $\frac{l}{2\pi\sqrt{2}} \frac{\alpha(k,0)}{\beta(k)}$). These progressive waves constitute the vibrational normal modes of the system[1].

Comment:

For each value of k, the two terms appearing on the right-hand sides of (39) and (40) are complex conjugates of each other. This insures the reality of the $x_q(t)$ and $p_q(t)$ without the necessity of imposing an arbitrary condition on the $\alpha(k,t)$. The $\alpha(k,t)$, consequently, are truly independent variables.

1-d. Total energy and energy of each of the modes

The total energy of the system under consideration is the sum of the kinetic energies of each particle (q) and the potential energies (1) and (6):

$$\mathcal{H}(\ldots x_{-1},\, x_0,\, x_{+1},\, \ldots p_{-1},\, p_0,\, p_{+1} \ldots) = \sum_{q=-\infty}^{+\infty} \left[\frac{1}{2m} p_q^2 + \frac{1}{2} m\omega^2 x_q^2 + \frac{1}{2} m\omega_1^2 \left(x_q - x_{q+1}\right)^2 \right] \tag{42}$$

We shall see in this section that this energy can be expressed very simply in terms of the energies that can be associated with each of the modes.

Let us therefore calculate the various sums involved in (42). Since the displacements x_q are the coefficients of the Fourier series defining the function $\xi(k,t)$, Parseval's

[1] We could also have introduced the modes corresponding to stationary waves in the system (sum of two progressive waves of the same frequency and opposite velocities). We would then have obtained equivalent results but the motion of the system would have been expanded on a different "basis". An expansion of this type is used in Complement K_V.

relation [Appendix I, relation (18)] immediately yields:

$$\sum_{q=-\infty}^{+\infty} (x_q)^2 = \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} |\xi(k,t)|^2 \, \mathrm{d}k \tag{43}$$

$$\sum_{q=-\infty}^{+\infty} (p_q)^2 = \frac{l}{2\pi} \int_{-\pi/l}^{+\pi/l} |\pi(k,t)|^2 \, \mathrm{d}k \tag{44}$$

The sum which, in (42), corresponds to the coupling, now remains to be calculated. To do so, notice as in (18) that, if the displacements x_q are the coefficients of the Fourier series of $\xi(k,t)$, the x_{q+1} are those of $\mathrm{e}^{ikl}\ \xi(k,t)$. The quantities $(x_q - x_{q+1})$ are therefore the coefficients of the Fourier series of $[1 - \mathrm{e}^{ikl}]\xi(k,t)$, and Parseval's relation yields:

$$\begin{aligned} \sum_{q=-\infty}^{+\infty} (x_q - x_{q+1})^2 &= \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} |(1 - \mathrm{e}^{ikl}\xi(k,t)|^2 \, \mathrm{d}k \\ &= \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} 4 \sin^2\left(\frac{kl}{2}\right) |\xi(k,t)|^2 \, \mathrm{d}k \end{aligned} \tag{45}$$

Substituting (43), (44) and (45) into (42), we finally obtain:

$$\mathcal{H} = \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \left\{ \frac{m}{2} \left[\omega^2 + 4\omega_1^2 \sin^2\left(\frac{kl}{2}\right) \right] |\xi(k,t)|^2 + \frac{1}{2m} |\pi(k,t)|^2 \right\} \mathrm{d}k \tag{46}$$

We shall write this result in the form:

$$\mathcal{H} = \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} h(k) \, \mathrm{d}k \tag{47}$$

with:

$$h(k) = \frac{1}{2}\ m\ \Omega^2(k)\ |\xi(k,t)|^2 + \frac{1}{2m}\ |\pi(k,t)|^2 \tag{48}$$

where $\Omega(k)$ is given by (10). $\mathcal{H}$ is thus the sum (in fact, the integral) of the energies associated with the fictitious uncoupled harmonic oscillators for which $\xi(k,t)$ gives the position and $\pi(k,t)$, the momentum.

We can also express $h(k)$ in terms of the variables $\alpha(k,t)$ associated with each normal mode. Using (37), we transform expression (48) into:

$$h(k) = \frac{1}{2}\ \hbar\Omega(k)\ \Big[\alpha(k,t)\alpha^*(k,t) + \alpha^*(-k,t)\alpha(-k,t)\Big] \tag{49}$$

that is, taking (34) into account:

$$h(k) = \frac{1}{2}\ \hbar\Omega(k)\ \Big[\alpha(k,0)\alpha^*(k,0) + \alpha^*(-k,0)\alpha(-k,0)\Big] \tag{50}$$

$h(k)$ is therefore time-independent, which is not surprising since $h(k)$ is the energy of a harmonic oscillator. Furthermore, we again find in (47) that the fictitious oscillators are

independent, since the total energy $\mathcal{H}$ is simply the sum of the energies associated with each of them.

Substituting expression (49) into (47), we obtain:

$$\mathcal{H} = \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \mathrm{d}k \, \frac{1}{2} \hbar\Omega(k) \left[\alpha(k,t)\alpha^*(k,t) + \alpha(-k,t)\alpha^*(-k,t) \right] \tag{51}$$

We can then change k to $-k$ in the integral of the second term and consider $\mathcal{H}$ to be the sum of the energies $h'(k)$ associated with the normal modes characterized by the $\alpha(k,t)$:

$$\mathcal{H} = \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \mathrm{d}k \, h'(k) \tag{52}$$

with:

$$\begin{aligned} h'(k) &= \hbar\Omega(k) \; \alpha^*(k,t) \; \alpha(k,t) \\ &= \hbar\Omega(k) \; \alpha^*(k,0) \; \alpha(k,0) \end{aligned} \tag{53}$$

2. Quantum mechanical treatment

The quantum mechanical treatment of the problem of the infinite chain of coupled oscillators is based, in accordance with the general quantization rules, on the replacement of the classical quantities $x_q(t)$ and $p_q(t)$ by the observables X_q and P_q which satisfy the canonical commutation relations:

$$[X_{q_1}, P_{q_2}] = i\hbar \, \delta_{q_1 q_2} \tag{54}$$

2-a. Stationary states in the absence of coupling

In the absence of coupling ($\omega_1 = 0$), the Hamiltonian H of the system can be written:

$$\begin{aligned} H(\omega_1 = 0) &= \sum_q \left[\frac{1}{2} m\omega^2 X_q^2 + \frac{1}{2m} P_q^2 \right] \\ &= \sum_q H_q \end{aligned} \tag{55}$$

where H_q is the Hamiltonian of a one-dimensional harmonic oscillator acting in the state space of particle (q).

We introduce the operator a_q defined by:

$$a_q = \frac{1}{\sqrt{2}} \left[\sqrt{\frac{m\omega}{\hbar}} X_q + \frac{i}{\sqrt{m\hbar\omega}} P_q \right] \tag{56}$$

H_q can then be written:

$$H_q = \frac{1}{2} \left(a_q a_q^\dagger + a_q^\dagger a_q \right) \hbar\omega = \left(a_q^\dagger a_q + \frac{1}{2} \right) \hbar\omega \tag{57}$$

$a_q^\dagger$ and a_q are the creation and annihilation operators of an energy quantum for oscillator (q). We know (Chap. V, § C-1-a) that the eigenstates of H_q are given by:

$$|\varphi_{n_q}^q\rangle = \frac{1}{\sqrt{(n_q)!}} \left(a_q^\dagger\right)^{n_q} |\varphi_0^q\rangle \tag{58}$$

where $|\varphi_0^q\rangle$ is the ground state of oscillator (q) and n_q is a positive integer or zero. If we choose as the energy origin the energy of the ground state [which amounts to omitting, in (57), the term 1/2], we obtain for the energy $E_{n_q}^q$ of the state $|\varphi_{n_q}^q\rangle$):

$$E_{n_q}^q = n_q\,\hbar\omega \tag{59}$$

In the absence of coupling, the stationary states of the global system are tensor products of the form:

$$... \otimes |\varphi_{n_{-1}}^{-1}\rangle \otimes |\varphi_{n_0}^0\rangle \otimes |\varphi_{n_1}^1\rangle \otimes ... \tag{60}$$

and their energies are[2]:

$$E = \sum_q E_{n_q}^q = [... + n_{-1} + n_0 + n_1 + ...]\,\hbar\omega \tag{61}$$

The ground state, whose energy was chosen for the origin, is not degenerate since $E = 0$ is obtained, in (61), only for:

$$n_q = 0 \quad \text{for all } q \tag{62}$$

The corresponding state (60) is therefore unique. On the other hand, all the other levels are infinitely degenerate. For example, to the first level, of energy $\hbar\omega$, correspond all states (60) for which all the numbers n_q are zero except for one, which is itself equal to one. All but one of the oscillators are then in their ground states. It is because the excitation can be localized in any one of the oscillators that the level $E = \hbar\omega$ is infinitely degenerate.

2-b. Effects of the coupling

When the coupling is not zero, the Hamiltonian operator becomes:

$$H = H(\omega_1 = 0) + V \tag{63}$$

with:

$$V = \frac{1}{2} m\omega_1^2 \sum_q (X_q - X_{q+1})^2 \tag{64}$$

[2]If we had not changed the energy origin of each oscillator by omitting the term 1/2 in (57), we would have found an infinite energy for the system, whatever the quantum number n_q. This difficulty does not arise if, instead of an infinite chain, one considers a chain formed by a very large but finite number of oscillators. However, problems related to "edge effects" then appear.

The states (60) are no longer, in this case, the stationary states of the system as they are eigenstates of $H(\omega_1 = 0)$, but not of V. To see this, we write V in terms of the operators a_q and $a_q^\dagger$:

$$V = \frac{1}{4}\hbar\omega_1 \frac{\omega_1}{\omega} \sum_q \left(a_q + a_q^\dagger - a_{q+1} - a_{q+1}^\dagger\right)^2 \tag{65}$$

Now, it is clear that the action of V on a state of type (60) changes the state: the numbers n_q are no longer "good quantum numbers" since, for example, V can transfer an excitation from site (q) to site $(q+1)$ (term in $a_{q+1}^\dagger a_q$).

To find the stationary states of the system in the presence of the coupling, it is useful, as in classical mechanics, to introduce "normal variables", that is, operators associated with the normal modes of the system.

2-c. Normal operators. Commutation relations

To the normal variables $\xi(k,t)$ and $\pi(k,t)$ correspond the operators $\Xi(k)$ and $\Pi(k)$ defined by:

$$\Xi(k) = \sum_q X_q \, \mathrm{e}^{-iqkl} \tag{66a}$$

$$\Pi(k) = \sum_q P_q \, \mathrm{e}^{-iqkl} \tag{66b}$$

The domain of variation of the continuous parameter k is again limited to the first Brillouin zone (15). Note that, since the normal variables $\xi(k,t)$ and $\pi(k,t)$ are complex, the associated operators $\Xi(k)$ and $\Pi(k)$ are not Hermitian, unlike X_q and P_q. The relations corresponding to (24) and (27) are here:

$$\Xi(-k) = \Xi^\dagger(k) \tag{67a}$$

$$\Pi(-k) = \Pi^\dagger(k) \tag{67b}$$

The canonical commutation relations (54) enable us to calculate the commutators of $\Xi(k)$ and $\Pi(k)$. We immediately see that $\Xi(k)$ and $\Xi(k')$ commute, as do $\Pi(k)$ and $\Pi(k')$. As for the commutator $[\Xi(k), \Pi^\dagger(k')]$, it can be written:

$$\begin{aligned}[\Xi(k), \Pi^\dagger(k')] &= \sum_q \sum_{q'} [X_q, P_{q'}] \, \mathrm{e}^{-iqkl} \mathrm{e}^{+iq'k'l} \\ &= i\hbar \sum_q \mathrm{e}^{-iq(k-k')l}\end{aligned} \tag{68}$$

Using formula (31) of Appendix II and the fact that k and k' both belong to interval (15), we obtain:

$$[\Xi(k), \Pi^\dagger(k')] = i\hbar \frac{2\pi}{l} \delta(k - k') \tag{69}$$

We saw in § 1-c-β that it is convenient to condense the two normal variables $\xi(k,t)$ and $\pi(k,t)$ into one, $\alpha(k,t)$ [formula (30)]. The operator associated with $\alpha(k,t)$ will be:

$$a(k) = \frac{1}{\sqrt{2}} \left[\beta(k)\, \Xi(k) + \frac{i}{\hbar\beta(k)} \Pi(k) \right] \tag{70}$$

where $\beta(k)$ is defined in (32). Note that the adjoint of $a(k)$ can be written:

$$a^\dagger(k) = \frac{1}{\sqrt{2}} \left[\beta(k)\, \Xi^\dagger(k) - \frac{i}{\hbar\beta(k)} \Pi^\dagger(k) \right] \tag{71}$$

Using (69) and (67), we find without difficulty that:

$$[a(k), a(k')] = \left[a^\dagger(k), a^\dagger(k')\right] = 0 \tag{72a}$$

$$\left[a(k), a^\dagger(k')\right] = \frac{2\pi}{l} \delta(k - k') \tag{72b}$$

To the classical quantity $h(k)$ defined in (48) corresponds the operator:

$$H(k) = \frac{1}{2m}\ \Pi(k)\, \Pi^\dagger(k) + \frac{1}{2}\ m\, \Omega^2(k)\ \Xi(k)\, \Xi^\dagger(k) \tag{73}$$

since $\Xi(k)$ and $\Xi^\dagger(k)$ commute, as do $\Pi(k)$ and $\Pi^\dagger(k)$. To obtain the equivalent of the classical formula (49), one must take into consideration the fact that $a(k)$ and $a^\dagger(k)$ do not commute; the order in which these operators appear must therefore be retained throughout the calculation. Relations (37), taking (67) into account, can be written here:

$$\beta(k)\, \Xi(k) = \frac{1}{\sqrt{2}} \left[a(k) + a^\dagger(-k)\right] \tag{74a}$$

$$\frac{1}{\hbar\beta(k)}\, \Pi(k) = -\frac{i}{\sqrt{2}} \left[a(k) - a^\dagger(-k)\right] \tag{74b}$$

Substituting these expressions into (73), we find:

$$H(k) = \frac{1}{2} \hbar\Omega(k) \left[a(k) a^\dagger(k) + a^\dagger(-k) a(-k)\right] \tag{75}$$

As in (52), we can put the total Hamiltonian H of the system in the form:

$$H = \frac{l}{2\pi} \int_{-\frac{\pi}{l}}^{+\frac{\pi}{l}} \mathrm{d}k\, H'(k) \tag{76}$$

with:

$$H'(k) = \frac{1}{2} \hbar\Omega(k) \left[a(k) a^\dagger(k) + a^\dagger(k) a(k)\right] \tag{77}$$

$a(k)$ and $a^\dagger(k)$ thus can be seen to be annihilation and creation operators analogous to those of a harmonic oscillator. However, since k is a continuous index, the commutation relations (72) involve $\delta(k-k')$ instead of a Kronecker delta, so $H'(k)$ must remain in the symmetrical form (77). It can easily be shown that the various operators $H'(k)$ commute:

$$[H'(k), H'(k')] = 0 \tag{78}$$

2-d. Stationary states in the presence of coupling

According to formulas (76) and (77), the ground state $|0\rangle$ of the system of coupled oscillators is defined by the condition:

$$a(k)|0\rangle = 0 \tag{79}$$

for all values of k. The other stationary states can be obtained from the state $|0\rangle$ by the action of the operators $a^\dagger(k)$; their corresponding energy is the integral of the energies associated with each of the modes. A certain number of difficulties arise because of the continuous infinity of normal modes; in particular, the energy of the ground state that can be deduced from (76) and (77) is infinite. We shall not discuss these difficulties here; in any case, they do not arise for a real chain, that is, a finite one (*cf.* footnote 2).

Formula (10) gives the value of the energy quantum $\hbar\Omega(k)$ associated with each of the modes. It therefore indicates what energy quanta the system can absorb or emit: they must correspond to frequencies situated within the allowed band (17).

3. Application to the study of crystal vibrations: phonons

3-a. Outline of the problem

Consider a solid body, composed of a large number of atoms (or ions) whose equilibrium positions are regularly arranged in space, forming a crystalline lattice. For simplicity, we shall assume that this lattice is one-dimensional and can be treated like an infinite chain of atoms. We intend to use here the results of the preceding section to study the motion of the nuclei of these atoms about their equilibrium positions.

With this object in view, we shall make use of the same approximation as in the study of molecular vibrations (Born-Oppenheimer approximation; *cf.* Complement A_V, comment of § 1-a). We shall assume that the motion of the electrons can be calculated as if the positions of the nuclei were fixed parameters x_q. Thus, we shall solve the corresponding Schrödinger equation (actually, this equation is too complex to be solved exactly; in practice, one must again settle for approximations). We shall then denote the energy of the electronic system in its ground state by $E_{el}(..., x_{-1}, x_0, x_1, ...)$ where x_q is the displacement of nucleus (q) from its equilibrium position. It can be shown that it is then possible to calculate the motion of the nuclei, to a good approximation, by assuming that they possess a total potential energy $U_N(... x_{-1}, x_0, x_1, ...)$ equal to the sum of their electrostatic interaction energy and $E_{el}(..., x_{-1}, x_0, x_1, ...)$.

In fact, we shall further simplify the problem by making some reasonable hypotheses about U_N (this is indispensable, since we do not know E_{el}). We shall assume that U_N describes essentially the interactions of each of the nuclei with its nearest neighbors (in an infinite linear chain, each nucleus has two such neighbors), that is, that the forces exerted between non-adjacent nuclei can be neglected. In addition, we shall grant that, in the range of values that the displacements x_q can attain, U_N is well represented by an expression of the form:

$$U_N \simeq \frac{1}{2} m\omega_1^2 \sum_q (x_q - x_{q+1})^2 \tag{80}$$

where m is the mass of a nucleus and ω_1 characterizes the intensity of its interaction

with its neighbors. We shall not, therefore, take into account terms of higher order in $(x_q - x_{q+1})$, that is, the anharmonicity of the potential.

Since expression (80) is identical to (6), we can apply the results of the preceding sections to the simple model of a solid body that we have just defined. Note, nevertheless, that we must choose $\omega = 0$ because U_N is the total potential energy of the system of the nuclei, which interact with their neighbors, but are not elastically bound to their equilibrium positions[3].

3-b. Normal modes. Speed of sound in the crystal

Each of the vibrational normal modes of the crystal is characterized by a wave vector k and an angular frequency $\Omega(k)$. In solid state physics, the energy quantum associated with a mode is called a "phonon". The phonons can be considered to be particles of energy $\hbar\Omega(k)$ and momentum $\hbar k$. Actually, a phonon is not a true particle, since its existence involves a state of collective vibration of the real particles which constitute the crystal. The phonons are sometimes said to be "quasi-particles": they are entirely analogous to the fictitious particles, of position $x_G(t)$ and $x_R(t)$, introduced in Complement H_V. In addition, a phonon can be created or destroyed by giving to or taking from the crystal the corresponding vibrational energy, while (at least in the non-relativistic domain to which we are restricting ourselves) a particle such as an electron cannot be created or annihilated. In this connection, note that, since the number of phonons in a given mode is arbitrary, phonons are bosons (Chap. XIV).

The dispersion relation giving the function $\Omega(k)$ differs, for phonons, from the one discussed in § 1.b.δ, since the angular frequency ω is zero here. In this case, choosing $\omega = 0$ in (10), we obtain:

$$\Omega(k) = 2\omega_1 \left|\sin\left(\frac{kl}{2}\right)\right| \tag{81}$$

The curve representing $\Omega(k)$ is given in Figure 5; it is composed of two half-arcs of a sinusoid. Unlike what happens for ω not equal to zero, $\Omega(k)$ now goes to zero for $k = 0$ and varies linearly when k is very small, since, as long as:

$$|k| \ll \frac{1}{l} \tag{82}$$

we have:

$$\Omega(k) \simeq \omega_1 |kl| = v_s |k| \tag{83}$$

where:

$$v_s = \omega_1 l \tag{84}$$

[3]The Einstein model, which we described in Complement A_V, is based on a different hypothesis: each nucleus is assumed to "see" an average potential, due to its interactions with the other nuclei, but practically independent of the exact positions of these other nuclei. To a first approximation, this average potential is assumed to be parabolic, and one has a system of independent harmonic oscillators. Here, on the other hand, we are studying a somewhat more elaborate model, in which we explicitly (although approximately) take into account interactions between the nuclei.

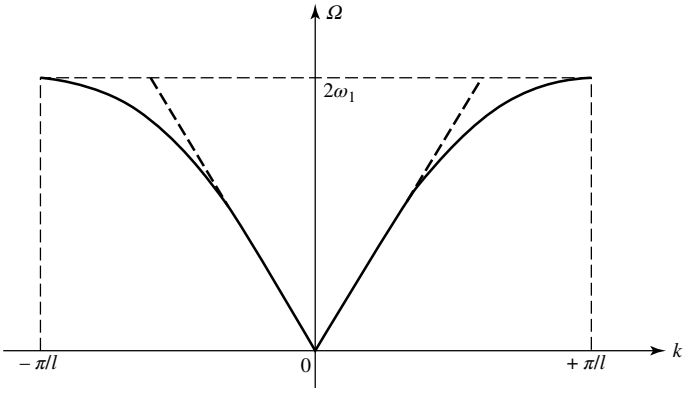


Figure 5: Dispersion relation for phonons (curve of Figure 2 for $\omega = 0$); the slope of the curve at the origin gives the speed of sound in the crystal.

Condition (82) means that the wavelength $2\pi/|k|$ associated with the mode being considered must be much greater than the separation between nuclei. For such wavelengths, the discontinuous structure of the chain is negligible, and the medium is not dispersive: the phase velocity $\Omega(k)/|k| \simeq v_s$ is independent of k, which implies that a wave packet involving only small values of k (of the same sign) propagates without being deformed at the velocity v_s. Since acoustical wavelengths satisfy (82), v_s is the speed of sound in the crystal.

When $|k|$ is of the order of $1/l$, the discontinuous structure of the chain becomes important, and the angular frequency $\Omega(k)$ increases less rapidly with $|k|$ than formula (83) would indicate (in Figure 5, the curve deviates from the straight dashed lines which are its tangents at the origin). The medium is then dispersive, and a wave packet moves with a group velocity:

$$v_G = \frac{\mathrm{d}\Omega(k)}{\mathrm{d}k} \neq \frac{\Omega(k)}{k} \tag{85}$$

Finally, when k approaches the edges of the first Brillouin zone ($k \longrightarrow \pm\pi/l$), we see from Figure 5 that the group velocity approaches zero. As in an electromagnetic waveguide, the propagation velocity goes to zero when the cut-off frequency (here $\omega_1/2\pi$) is attained.

Figure 5 can also be seen as giving the spectrum of possible phonon energies $\hbar\Omega(k)$ in terms of their momentum $\hbar k$. Knowledge of such a spectrum, for a real crystal, is very important. It gives the precise energies and momenta that the crystal can supply or absorb when it interacts with another system. For example, the inelastic scattering of light by a crystal (the Brillouin effect) can be interpreted as the result of the annihilation or creation of a phonon, with a change in the energy and momentum of the incident photon (total energy and momentum being conserved throughout the process).

Comment:

The simple one-dimensional model described here has allowed us to present some important physical concepts, which remain valid for a real crystal: energy quanta associated with the normal modes, dispersion of the medium, allowed and forbidden frequency (and, therefore, energy) bands. In reality, the crystalline lattice is three-dimensional, and a normal mode is characterized by a true wave vector $\mathbf{k}$; Ω then depends, in general, not only on the absolute value of $\mathbf{k}$, but also on its direction. Also, the situation may arise (as is the case for an ionic crystal) in which the vertices of the lattice are not all occupied by identical particles, but rather, for example, by two different types of particles in regular alternation[4]. Then, for each wave vector $\mathbf{k}$, several angular frequencies $\Omega(\mathbf{k})$ appear. Some of them, which go to zero when $|\mathbf{k}| \longrightarrow 0$, constitute "acoustic branches" like the one encountered above; the others belong to what are called "optical branches"[5], in which a phonon of zero momentum has a non-zero energy. It would be out of the question to study all these problems here, although they are of primordial importance in solid state physics.

References and suggestions for further reading:

Chains of coupled classical oscillators: Berkeley 3 (7.1), §§ 2.4 and 3.5.
See section 13 of the bibliography, particularly Kittel (13.2), Chap. 5.
Other examples of collective oscillations: Feynman III (1.2), Chap. 15.

[4]A real crystal also contains impurities and imperfections distributed at random. Here we are speaking only about perfect crystals.

[5]This name arises from the fact that, in an ionic crystal, the "optical" phonons are coupled to electromagnetic waves like those in the visible domain, whose wavelengths are much larger than the atomic separation.

Complement K_V

Vibrational modes of a continuous physical system. Application to radiation; photons

1 Outline of the problem 631

2 Vibrational modes of a continuous mechanical system: example of a vibrating string 632

2-a Notation. Dynamical variables of the system 632

2-b Classical equations of motion 633

2-c Introduction of the normal variables 633

2-d Classical Hamiltonian 634

2-e Quantization 637

3 Vibrational modes of radiation: photons 639

3-a Notation. Equations of motion 639

3-b Introduction of the normal variables 640

3-c Classical Hamiltonian 641

3-d Quantization 642

1. Outline of the problem

In Complements H_V and J_V, we introduced the idea of normal variables for a system of two or a countable infinity of coupled harmonic oscillators. The aim of this complement is to show that the same ideas can also be applied to the electromagnetic field, which is a *continuous* physical system (there is no natural lower bound for the wavelength of radiation). The quantization of the electromagnetic field will be studied in much more detail in Chapter XIX.

This study raises a certain number of delicate problems. Therefore, before beginning, and in order to make a smooth transition between this complement and the preceding ones, H_V and J_V, we shall start by studying in § 2 the vibrational modes of a continuous *mechanical* system: a vibrating string. It is obvious that, on the atomic scale, such a system is not continuous: the string is composed of a very large number of atoms. However, we shall ignore this atomic structure and treat the string as if it were really continuous, since the fundamental aim of the calculation is to show how normal variables can be introduced for a continuous system. Since, moreover, we are dealing with a mechanical system, we can, without difficulty, define the conjugate momenta of the normal variables, calculate the Hamiltonian of the system, and show that it indeed appears in the form of a sum of Hamiltonians of independent one-dimensional harmonic oscillators. We shall also discuss in detail the quantization of such a system.

The results obtained in § 2 will enable us to attack, in § 3, the problem of the vibrational modes of radiation. We shall show that the study of radiation confined to a parallelepiped cavity leads to equations very similar to those of the vibrating string. The

same transformations allow the introduction of completely uncoupled normal variables for the radiation (associated with the standing waves which can exist inside the cavity). Then we shall generalize the results obtained in § 2 to derive simply the concept of a photon (as it is not feasible here to show rigorously how one can introduce, for a non-mechanical system like the electromagnetic field, conjugate momenta, a Lagrangian and a Hamiltonian).

2. Vibrational modes of a continuous mechanical system: example of a vibrating string

2-a. Notation. Dynamical variables of the system

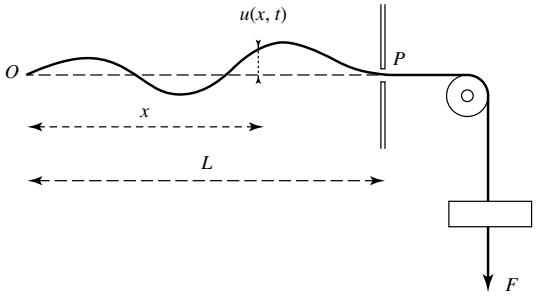


Figure 1: Vibrating string passing through two fixed points O and P and subjected to a tension F; $u(x,t)$ denotes the deviation with respect to the equilibrium position of the point of the string situated at a distance x from O.

The string is fixed at a point O (Fig. 1). It passes through a very small hole P pierced in a plate, and a weight exerts a tension F on it. For simplicity, we assume that the string is inextensible (it has a constant length) and that it always remains in the same plane, passing through O and P. Its state is defined at time t when we know at this time the displacement $u(x,t)$ of the various points (labeled by their abscissas x on OP), as well as the corresponding velocities $\dfrac{\partial u(x,t)}{\partial t}$. The constraints imposed at O and P are expressed by the boundary conditions:

$$u(0,t) = u(L,t) = 0 \tag{1}$$

where 0 and L are the abscissas of the points O and P.

It is important to remember that, in this problem, the dynamical variables are the displacements $u(x,t)$ at each point of abscissa x: there is a continuous infinity of dynamical variables. Consequently, x is not a dynamical variable, but rather a continuous index which labels the dynamical variable with which we are concerned (x plays the same role as indices 1 and 2 of Complement H_V or as the index q of Complement J_V).

2-b. Classical equations of motion

Let μ be the mass per unit length of the string. If we assume the string to be perfectly flexible and if we confine ourselves to small displacements, a classical calculation enables us to obtain the partial differential equation satisfied by u. We find:

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right) u(x,t) = 0 \tag{2}$$

where:

$$v = \sqrt{F/\mu} \tag{3}$$

is the propagation velocity of a perturbation along the string.

Such an equation expresses the fact that the evolution of the variable u corresponding to the point x depends on the variables u at infinitely near points (via $\partial^2 u/\partial x^2$). Thus the variables $u(x,t)$ are all coupled to each other. We can then ask the following question: is it possible, as in Complements $\mathrm{H_V}$ and $\mathrm{J_V}$, to introduce new, uncoupled variables that are linear combinations of the variables $u(x,t)$ associated with the various points x?

2-c. Introduction of the normal variables

Consider the set of functions of x:

$$f_k(x) = \sqrt{\frac{2}{L}} \sin\left(k\frac{\pi x}{L}\right) \tag{4}$$

where k is a positive integer: $k = 1, 2, 3...$ The $f_k(x)$ satisfy the same boundary conditions as $u(x,t)$:

$$f_k(0) = f_k(L) = 0 \tag{5}$$

In addition, it is easy to verify the relations:

$$\int_0^L f_k(x)\, f_{k'}(x)\, \mathrm{d}x = \delta_{kk'} \tag{6}$$

(orthonormalization relation) and:

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{k^2\pi^2}{L^2}\right) f_k(x) = 0 \tag{7}$$

It can be shown that any function which goes to zero at $x = 0$ and $x = L$ [as is the case for $u(x,t)$] can be expanded in one and only one way in terms of the $f_k(x)$. We can therefore write:

$$u(x,t) = \sum_{k=1}^{\infty} q_k(t)\, f_k(x) \tag{8}$$

$q_k(t)$ can be obtained easily, using (6):

$$q_k(t) = \int_0^L u(x,t)\, f_k(x)\, \mathrm{d}x \tag{9}$$

The state of the string at the instant t can be defined either by the set of values $\left\{u(x,t), \dfrac{\partial}{\partial t}u(x,t)\right\}$ corresponding to the various points x or by the set of numbers $\{q_k(t), \dot{q}_k(t)\}$. The new variables $q_k(t)$ just introduced are linear combinations of the old $u(x,t)$, as can be seen from (9). The converse is obviously also true [*cf.* formula (8)].

To obtain the equation satisfied by the $q_k(t)$, we substitute expansion (8) into the equation of motion (2). Using (7), we obtain, after a simple calculation:

$$\sum_{k=1}^{\infty} f_k(x) \left[\frac{1}{v^2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}q_k(t) + \frac{k^2\pi^2}{L^2}q_k(t)\right] = 0 \tag{10}$$

that is, since the $f_k(x)$ are linearly independent:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_k^2\right] q_k(t) = 0 \tag{11}$$

with:

$$\omega_k = \frac{k\pi v}{L} \tag{12}$$

Thus we see that the new variables $q_k(t)$, also called "normal variables", evolve independently: they are *uncoupled*. Moreover, equation (11) is identical to that of a one-dimensional harmonic oscillator of angular frequency ω_k, so that:

$$q_k(t) = A_k \cos(\omega_k t - \varphi_k) \tag{13}$$

Each of the terms $q_k(t)\ f_k(x)$ appearing on the right-hand side of (8) consequently represents a standing wave of frequency $\omega_k/2\pi$ and half-wavelength L/k. Each normal variable q_k is therefore associated with a vibrational normal mode of the string, the most general motion of the string being a linear superposition of these normal modes.

Comment:

In Complement J_V, we started with an infinite discrete set of harmonic oscillators and introduced a continuous infinity of normal variables. Here we find ourselves in the opposite situation: the $u(x,t)$ form a continuous set with respect to the index x, while, because of the boundary conditions, the normal variables $q_k(t)$ are labeled by a discrete index k.

2-d. Classical Hamiltonian

α. Kinetic energy

The kinetic energy of the segment of string included between x and $x + \mathrm{d}x$ is $\dfrac{1}{2}\mu\,\mathrm{d}x\left[\dfrac{\partial u(x,t)}{\partial t}\right]^2$. It follows that the total kinetic energy E_{kin} of the string is equal to:

$$E_{\text{kin}} = \frac{\mu}{2}\int_0^L \left[\frac{\partial u(x,t)}{\partial t}\right]^2 \mathrm{d}x \tag{14}$$

E_{kin} can be expressed simply in terms of the q_k, using (8):

$$E_{\text{kin}} = \frac{\mu}{2} \sum_k \sum_{k'} \frac{\mathrm{d}q_k(t)}{\mathrm{d}t} \frac{\mathrm{d}q_{k'}(t)}{\mathrm{d}t} \int_0^L f_k(x)\, f_{k'}(x)\, \mathrm{d}x \tag{15}$$

which can also be written, taking (6) into account:

$$E_{\text{kin}} = \frac{\mu}{2} \sum_k \left(\frac{\mathrm{d}q_k}{\mathrm{d}t} \right)^2 \tag{16}$$

β. Potential energy

Consider the segment of string included between the abscissas x and $x + \mathrm{d}x$. It makes an angle θ with the Ox axis such that:

$$\tan\theta = \frac{\partial u(x,t)}{\partial x} \tag{17}$$

Its length is therefore equal to:

$$\frac{\mathrm{d}x}{\cos\theta} = \mathrm{d}x \left[1 + \tan^2\theta\right]^{\frac{1}{2}} \tag{18}$$

Since the displacements are small, θ is very small, and we can write:

$$\frac{\mathrm{d}x}{\cos\theta} = \mathrm{d}x \left[1 + \frac{1}{2} \left(\frac{\partial u(x,t)}{\partial x} \right)^2 \right] \tag{19}$$

From this we deduce that the total increase in the length of the string with respect to its equilibrium position (which corresponds to $u \equiv 0$ for all x) is equal to:

$$\Delta L = \frac{1}{2} \int_0^L \left(\frac{\partial u(x,t)}{\partial x} \right)^2 \mathrm{d}x \tag{20}$$

Now, ΔL represents the distance over which the end of the inextensible rope have risen, which means that the weight has the same motion. The potential energy E_{pot} of the string, with respect to the value corresponding to the equilibrium position, is therefore equal to:

$$E_{\text{pot}} = F\Delta L = \frac{1}{2} F \int_0^L \left(\frac{\partial u(x,t)}{\partial x} \right)^2 \mathrm{d}x \tag{21}$$

E_{pot} can also be expressed in terms of the normal variables q_k. A simple calculation, using (8) and (4), yields:

$$E_{\text{pot}} = \frac{F}{2} \sum_k \frac{k^2\pi^2}{L^2}\, q_k^2 \tag{22}$$

γ. Conjugate momenta of the q_k; classical Hamiltonian

The Lagrangian $\mathcal{L}$ of the system (*cf.* Appendix III) can be written:

$$\mathcal{L} = E_{\text{kin}} - E_{\text{pot}} = \frac{\mu}{2} \sum_k \left[\dot{q}_k^2 - \omega_k^2 q_k^2 \right] \tag{23}$$

From this, we deduce the expression for the conjugate momentum p_k of q_k:

$$p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k} = \mu \dot{q}_k \tag{24}$$

so that we finally obtain, for the Hamiltonian $\mathcal{H}(q_k, p_k)$ of the system, the expression:

$$\mathcal{H} = E_{\text{kin}} + E_{\text{pot}} = \sum_k \left[\frac{p_k^2}{2\mu} + \frac{1}{2} \mu \omega_k^2 \ q_k^2 \right] \tag{25}$$

that is:

$$\mathcal{H} = \sum_k h_k \tag{26}$$

with:

$$h_k = \frac{p_k^2}{2\mu} + \frac{1}{2} \mu \omega_k^2 q_k^2 \tag{27}$$

Since p_k and q_k are conjugate variables, we recognize h_k to be the Hamiltonian of a one-dimensional harmonic oscillator of angular frequency ω_k. $\mathcal{H}$ is therefore a sum of Hamiltonians of independent one-dimensional harmonic oscillators (independent, since the normal variables are uncoupled).

It is useful to introduce, as in Complements B_V and C_V, the dimensionless variables:

$$\hat{q}_k = \beta_k q_k \tag{28a}$$

$$\hat{p}_k = \frac{1}{\beta_k \hbar} \ p_k \tag{28b}$$

where:

$$\beta_k = \sqrt{\frac{\mu \omega_k}{\hbar}} \tag{29}$$

is a (dimensional) constant. $\mathcal{H}$ can then be written:

$$\mathcal{H} = \sum_k \frac{1}{2} \hbar \omega_k \left[\hat{q}_k^2 + \hat{p}_k^2 \right] \tag{30}$$

2-e. Quantization

α. Preliminary comment

The calculations performed in this section are, of course, not intended to reveal quantum effects in the motion of a macroscopic vibrating string. The vibrational frequencies $\omega_k/2\pi$ which can be excited on such a string are so low (of the order of a kilohertz at most) and the elementary energies $\hbar\omega_k$ so much smaller than the macroscopic energy of the string, that a classical description is quite sufficient. It might be thought that ω_k could be as large as we like because k has no upper bound in formula (12). In fact, for sufficiently small wavelengths $2L/k$, the rigidity of the string can no longer be neglected, and equation (2) is no longer valid. Furthermore, as we pointed out in the introduction, the string is not really a continuous system, and it would make no sense to consider wavelengths smaller than the interatomic distance.

The calculations which will be presented here must be considered as a simple first approach to problems posed by the quantum mechanical description of radiation. Now, radiation constitutes a truly continuous system (no natural lower bound exists for the wavelength) and satisfies an equation which is analogous to (2), whatever frequencies and wavelengths may be involved[1].

β. Eigenstates and eigenvalues of the quantum mechanical Hamiltonian H

We quantize each oscillator by associating with $\hat{q}_k$ and $\hat{p}_k$ [see formula (28)] observables $\hat{Q}_k$ and $\hat{P}_k$ such that:

$$\left[\hat{Q}_k, \hat{P}_k\right] = i \tag{31}$$

Since the normal variables are uncoupled, we also assume that the operators relating to two different oscillators commute. We therefore have:

$$\left[\hat{Q}_k, \hat{P}_{k'}\right] = i\delta_{kk'} \tag{32}$$

Let:

$$H_k = \frac{1}{2}\hbar\omega_k \left(\hat{Q}_k^2 + \hat{P}_k^2\right) \tag{33}$$

be the quantum mechanical Hamiltonian of oscillator k. From the results of Chapter V, we know its eigenstates and eigenvalues:

$$H_k|n_k\rangle = \left(n_k + \frac{1}{2}\right)\hbar\omega_k|n_k\rangle \tag{34}$$

where n_k is a non-negative integer (to simplify the notation, we shall writen $|n_k\rangle$ instead of $|\varphi_{nk}\rangle$).

Since the H_k commute, we can choose the eigenstates of H to be in the form of tensor products of the $|n_k\rangle$:

$$|n_1\rangle|n_2\rangle...|n_k\rangle... = |n_1, n_2, ..., n_k, ...\rangle \tag{35}$$

[1] If we were really interested in a microscopic "vibrating string" (for example, a linear macromolecule), it would be more realistic to consider, as in Complement J_V, a chain of atoms and to study not only their longitudinal displacements but also their transversal ones (transverse phonons).

The ground or "vacuum" state corresponds to all the n_k equal to zero:

$$|0, 0, \ldots 0 \ldots\rangle = |0\rangle \tag{36}$$

When we choose the energy origin to be the energy of the state $|0\rangle$, the energy of state (35) is equal to:

$$E_{n_1\, n_2 \ldots n_k \ldots} = \sum_k n_k\, \hbar\omega_k \tag{37}$$

A state such as (35) can be considered to represent a set of n_1 energy quanta $\hbar\omega_1$, ..., n_k energy quanta $\hbar\omega_k$, ... These vibrational quanta are analogous to the phonons studied in Complement J_V.

Finally, using $\hat{P}_k$ and $\hat{Q}_k$, we can introduce, as in § B of Chapter V, creation and annihilation operators for an energy quantum $\hbar\omega_k$:

$$a_k = \frac{1}{\sqrt{2}}\left(\hat{Q}_k + i\hat{P}_k\right) \tag{38}$$

where $a_k^\dagger$ is the adjoint of a_k. We then have:

$$\left[a_k, a_{k'}^\dagger\right] = \delta_{kk'} \tag{39}$$

and:

$$\begin{cases} a_k|n_1,\, n_2,\, \ldots,\, n_k,\, \ldots\rangle = \sqrt{n_k}\ |n_1,\, n_2,\, \ldots,\, n_k - 1,\, \ldots\rangle \\ a_k^\dagger|n_1,\, n_2,\, \ldots,\, n_k,\, \ldots\rangle = \sqrt{n_k + 1}\ |n_1,\, n_2,\, \ldots,\, n_k + 1,\, \ldots\rangle \end{cases} \tag{40}$$

All the states (35) can be expressed in terms of the vacuum state $|0\rangle$:

$$|n_1,\, n_2,\, \ldots,\, n_k,\, \ldots\rangle = \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}}\ \frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}}\ \ldots\ \frac{(a_k^\dagger)^{n_k}}{\sqrt{n_k!}} \ldots |0\rangle \tag{41}$$

γ. Quantum mechanical state of the system

The most general quantum mechanical state of the system is a linear superposition of the states $|n_1,\, n_2,\, \ldots,\, n_k,\, \ldots\rangle$:

$$|\psi(t)\rangle = \sum_{n_1, n_2, \ldots, n_k, \ldots} c_{n_1, n_2, \ldots, n_k \ldots}(t)\ |n_1,\, n_2,\, \ldots,\, n_k \ldots\rangle \tag{42}$$

The equation of motion of $|\psi(t)\rangle$ is the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{43}$$

Using (37) and (43), we easily obtain:

$$c_{n_1, n_2, \ldots, n_k \ldots}(t) = c_{n_1, n_2, \ldots, n_k, \ldots}(0)\ \mathrm{e}^{-i\sum_k n_k \omega_k t} \tag{44}$$

δ. Observables associated with the dynamical variables $u(x,t)$

When the system is quantized, $u(x,t)$ becomes an observable $U(x)$ which does not depend[2] on t, and which is obtained by replacing in (8) $q_k(t)$ by the observable Q_k:

$$U(x) = \sum_k f_k(x)\ Q_k$$

$$= \sum_k \frac{1}{\beta_k\sqrt{2}}\ f_k(x)\left[a_k + a_k^\dagger\right] \tag{45}$$

We see that a displacement observable $U(x)$ can be defined for each value of x, and that it depends linearly on the creation and annihilation operators $a_k^\dagger$ and a_k.

It is interesting to compare the mean value of $U(x)$, $\langle\psi(t)|U(x)|\psi(t)\rangle$, with the classical quantity $u(x,t)$. Since, according to (40), a_k and $a_k^\dagger$ can link only states whose energy difference is $\mp\hbar\omega_k$, we deduce from (45) that the only Bohr frequencies that can appear in the evolution of $\langle U(x)\rangle(t)$ are the frequencies $\omega_1/2\pi$, $\omega_2/2\pi$, ..., $\omega_k/2\pi$, ... associated, respectively, with the spatial functions $f_1(x)$, $f_2(x)$, ... $f_k(x)$ ... Thus we find for $\langle U(x)\rangle(t)$ a linear superposition of the standing waves that can exist on the string. This analogy can, moreover, be pursued further. Let us calculate the derivative $\dfrac{\partial^2}{\partial t^2}\langle U(x)\rangle(t)$; using the fact that [*cf.* Complement $\mathrm{G_V}$, equation (17)]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle a_k\rangle = -i\omega_k\langle a_k\rangle \tag{46}$$

and relations (7) and (12), we easily find that the mean value of $U(x)$ given by (45) satisfies the differential equation:

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right)\langle U(x)\rangle(t) = 0 \tag{47}$$

which is identical to (2).

Finally, note that since Q_k does not commute with H_k, $U(x)$ does not commute with H. The displacement and the total energy are therefore, in quantum mechanics, incompatible physical quantities.

3. Vibrational modes of radiation: photons

3-a. Notation. Equations of motion

The classical state of the electromagnetic field at a given t is defined when we know, for this time, the value of the components of the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ at each point $\mathbf{r}$ of space. As in § 2 above, we therefore have a continuous infinity of dynamical variables: the six components E_x, E_y, E_z and B_x, B_y, B_z at each point $\mathbf{r}$.

In order to stress the importance of the idea of normal variables (or normal modes) of a field, we shall introduce a simplification which consists of forgetting the vector nature

[2] Recall that, in quantum mechanics, the time dependence is usually contained in the state vector and not in the observables (*cf.* discussion of § D-1-d of Chapter III).

of the fields **E** and **B**: we shall base our arguments on a scalar field $\mathcal{S}(\mathbf{r},t)$ which obeys (like each of the components of **E** and **B**) the equation:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2}-\Delta\right)\mathcal{S}(\mathbf{r},t)=0 \tag{48}$$

where c is the speed of light.

We shall assume the field to be confined to a parallelepiped cavity whose inside walls are perfectly conducting and whose edges, parallel to Ox, Oy, Oz, have, respectively, the lengths L_1, L_2, L_3. As boundary conditions, we require $\mathcal{S}(\mathbf{r},t)$ to be zero on the walls of the cavity (in the real problem, it is, for example, the tangential components of the electric field **E** that must go to zero on these walls). We can therefore write:

$$\begin{aligned}\mathcal{S}(x=0,y,z,t)=\mathcal{S}(x=L_1,y,z,t)=\mathcal{S}(x,y=0,z,t)=...\\ =\mathcal{S}(x,y,z=L_3,t)=0\end{aligned} \tag{49}$$

3-b. Introduction of the normal variables

Consider the set of functions of x, y, z:

$$f_{klm}(x,y,z)=\sqrt{\frac{8}{L_1L_2L_3}}\sin\left(\frac{k\pi x}{L_1}\right)\sin\left(\frac{l\pi y}{L_2}\right)\sin\left(\frac{m\pi z}{L_3}\right) \tag{50}$$

where k, l, m are positive integers ($k,l,m=1,2,3,...$). The $f_{klm}(x,y,z)$ go to zero on the walls of the cavity and therefore satisfy the same boundary conditions as $\mathcal{S}(x,y,z,t)$:

$$f_{klm}(x=0,y,z)=f_{klm}(x=L_1,y,z)=...=f_{klm}(x,y,z=L_3)=0 \tag{51}$$

In addition, the following relations are simple to verify:

$$\int_0^{L_1}\mathrm{d}x\int_0^{L_2}\mathrm{d}y\int_0^{L_3}\mathrm{d}z\ f_{klm}(x,y,z)\,f_{k'l'm'}(x,y,z)=\delta_{kk'}\ \delta_{ll'}\ \delta_{mm'} \tag{52}$$

and:

$$\left[\Delta+\left(\frac{k^2}{L_1^2}+\frac{l^2}{L_2^2}+\frac{m^2}{L_3^2}\right)\pi^2\right]f_{klm}(x,y,z)=0 \tag{53}$$

Any function that goes to zero on the walls of the cavity, $\mathcal{S}(\mathbf{r},t)$ in particular, can be expanded in one and only one way in terms of the $f_{klm}(x,y,z)$. We therefore have:

$$\mathcal{S}(x,y,z,t)=\sum_{k,l,m}q_{klm}(t)\,f_{klm}(x,y,z) \tag{54}$$

Formula (54) can easily be inverted, with the help of (52):

$$q_{klm}(t)=\int_0^{L_1}\mathrm{d}x\int_0^{L_2}\mathrm{d}y\int_0^{L_3}\mathrm{d}z\ f_{klm}(x,y,z)\ \mathcal{S}(x,y,z,t) \tag{55}$$

Thus we see that the field at time t is described, either by the set of variables $q_{klm}(t)$, or by the set of variables $\mathcal{S}(x, y, z, t)$. Formulas (54) and (55) enable us to go from one set to the other one.

Substituting (54) into (48) and using (53), we obtain, after a simple calculation:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega_{klm}^2\right] q_{klm}(t) = 0 \tag{56}$$

where:

$$\omega_{klm}^2 = c^2\pi^2 \left[\frac{k^2}{L_1^2} + \frac{l^2}{L_2^2} + \frac{m^2}{L_3^2}\right] \tag{57}$$

The normal variables $q_{klm}(t)$ are therefore *uncoupled*. According to (56), $q_{klm}(t)$ varies like $A\cos(\omega_{klm}t - \varphi)$. Each of the terms $q_{klm}(t)\, f_{klm}(x, y, z)$ of the sum (54) therefore represents a standing wave (a vibrational normal mode of the field in the cavity) characterized by its frequency $\omega_{klm}/2\pi$ and its spatial dependence in the three directions Ox, Oy, Oz (half-wavelengths L_1/k, L_2/l and L_3/m respectively).

Thus we have been able to generalize the results of § 2-c without difficulty. Note, nevertheless, that when the vectorial nature of the electromagnetic field is taken into account, the structure of the modes is more complex, However, the general idea is the same, and one reaches similar conclusions.

3-c. Classical Hamiltonian

Basing our discussion on the very close analogy between the results of § 2-c and those of 3-b, we shall assume without proof that one can associate with the field $\mathcal{S}(\mathbf{r}, t)$ a Lagrangian $\mathcal{L}$, from which can be deduced the equation of motion (48), the conjugate momenta $p_{klm}(t)$ of the normal variables, and finally, the expression for the Hamiltonian $\mathcal{H}$ of the system. The only point which is important here is that this Hamiltonian is analogous to (30):

$$\mathcal{H} = \sum_{k,l,m} \frac{1}{2}\hbar\omega_{klm}\left[(\hat{q}_{klm})^2 + (\hat{p}_{klm})^2\right] \tag{58}$$

where $\hat{q}_{klm}$ and $\hat{p}_{klm}$ are dimensionless variables proportional to the q_{klm} and p_{klm}:

$$\hat{q}_{klm} = \beta_{klm} q_{klm} \qquad \hat{p}_{klm} = \frac{1}{\hbar\beta_{klm}} p_{klm} \tag{59}$$

β_{klm} is a (dimensional) constant which is analogous to the one introduced in (29).

Comments:

(i) The equation of motion of each normal variable q_{klm} [established in (56)] is analogous to that of a one-dimensional harmonic oscillator of angular frequency ω_{klm}. Thus we see why we obtain for $\mathcal{H}$ a sum of Hamiltonians of independent one-dimensional harmonic oscillators. It is possible, moreover, to obtain (56) from (58).

The Hamilton-Jacobi equations (*cf.* Appendix III) can, in fact, be written, taking (59) into account:

$$\begin{cases} \dfrac{\mathrm{d}\hat{q}_{klm}}{\mathrm{d}t} = \dfrac{1}{\hbar}\dfrac{\partial \mathcal{H}}{\partial \hat{p}_{klm}} \\ \dfrac{\mathrm{d}\hat{p}_{klm}}{\mathrm{d}t} = -\dfrac{1}{\hbar}\dfrac{\partial \mathcal{H}}{\partial \hat{q}_{klm}} \end{cases} \tag{60}$$

that is, with the form (58) of $\mathcal{H}$:

$$\frac{\mathrm{d}\hat{q}_{klm}}{\mathrm{d}t} = \omega_{klm}\hat{p}_{klm} \tag{61a}$$

$$\frac{\mathrm{d}\hat{p}_{klm}}{\mathrm{d}t} = -\omega_{klm}\hat{q}_{klm} \tag{61b}$$

Eliminating $\hat{p}_{klm}$ between these two equations, we indeed find (56).

(*ii*) For the real electromagnetic field, composed of two fields **E** and **B**, one can also obtain expression (58) for $\mathcal{H}$ directly without using the Lagrangian. One simply takes the total energy $\mathcal{H}$ of the field as the sum of the electrical and magnetic energies contained in the cavity:

$$\mathcal{H} = \frac{\varepsilon_0}{2}\int_0^{L_1}\mathrm{d}x\int_0^{L_2}\mathrm{d}y\int_0^{L_3}\mathrm{d}z\left[\mathbf{E}^2 + c^2\mathbf{B}^2\right] \tag{62}$$

and uses for **E** and **B** expansions analogous to (54). Thus one finds that the terms in $\hat{p}_{klm}^2$ and $\hat{q}_{klm}^2$ of (58) correspond respectively to the electrical and magnetic energies.

3-d. Quantization

Now, starting with equation (58), we can carry out the same operations as in § 2-e.

α. Eigenstates and eigenvalues of H

We associate with $\hat{q}_{klm}$ and $\hat{p}_{klm}$ two observables $\hat{Q}_{klm}$ and $\hat{P}_{klm}$ whose commutator is equal to i. Since observables relating to two different modes commute, we can write:

$$\left[\hat{Q}_{klm}\ ,\ \hat{P}_{k'l'm'}\right] = i\ \delta_{kk'}\ \delta_{ll'}\ \delta_{mm'} \tag{63}$$

Let H_{klm} be the Hamiltonian associated with the mode (klm):

$$H_{klm} = \frac{\hbar\omega_{klm}}{2}\left[\left(\hat{Q}_{klm}\right)^2 + \left(\hat{P}_{klm}\right)^2\right] \tag{64}$$

We know its eigenstates and eigenvalues:

$$H_{klm}|n_{klm}\rangle = \left(n_{klm} + \frac{1}{2}\right)\hbar\omega_{klm}|n_{klm}\rangle \tag{65}$$

where n_{klm} is a non-negative integer.

Since the H_{klm} commute with each other, we can choose the eigenstates of $H = \sum_{klm} H_{klm}$ to be in the form of tensor products of the $|n_{klm}\rangle$:

$$|n_{111},\ n_{211},\ n_{121},\ n_{112},\ ...,\ n_{klm},\ ...\rangle \tag{66}$$

The ground state, called the "vacuum", corresponds to all the n_{klm} equal to zero:

$$|0,\ 0,\ 0,\ 0,\ ...,\ 0,\ ...\rangle = |0\rangle \tag{67}$$

With respect to the vacuum state energy, the energy of state (66) is equal to:

$$E_{n_{111}...n_{klm}...} = \sum_{klm} n_{klm}\,\hbar\omega_{klm} \tag{68}$$

A state such as (66) can be considered to represent a set of n_{111} energy quanta $\hbar\omega_{111}$, ..., n_{klm} energy quanta $\hbar\omega_{klm}$, ... These quanta are none other than the *photons*. Thus we see that a certain type of photon is associated with each normal mode of the cavity.

We can, as in (38), introduce annihilation and creation operators for a photon of type (klm):

$$\begin{cases} a_{klm} = \dfrac{1}{\sqrt{2}}\left(\hat{Q}_{klm} + i\hat{P}_{klm}\right) \\ a^{\dagger}_{klm} = \dfrac{1}{\sqrt{2}}\left(\hat{Q}_{klm} - i\hat{P}_{klm}\right) \end{cases} \tag{69}$$

and establish formulas identical to (39), (40) and (41):

$$\left[a_{klm}\ ,\ a^{\dagger}_{k'l'm'}\right] = \delta_{kk'}\ \delta_{ll'}\ \delta_{mm'} \tag{70}$$

$$\begin{cases} a_{klm}|n_{111},\ ...,\ n_{klm},\ ...\rangle = \sqrt{n_{klm}}\ |n_{111},\ ...,\ n_{klm}-1,...\rangle \\ a^{\dagger}_{klm}|n_{111},\ ...,\ n_{klm},\ ...\rangle = \sqrt{n_{klm}+1}\ |n_{111},\ ...,\ n_{klm}+1,...\rangle \end{cases} \tag{71}$$

$$|n_{111},\ ...,\ n_{klm},\ ...\rangle = \frac{\left(a^{\dagger}_{111}\right)^{n_{111}}}{\sqrt{n_{111!}}}\ ...\ \frac{\left(a^{\dagger}_{klm}\right)^{n_{klm}}}{\sqrt{n_{klm!}}}\ ...|0\rangle \tag{72}$$

β. Quantum state of the field

The most general state of the field is a linear superposition of states (66):

$$|\psi(t)\rangle = \sum_{n_{111},\,...,\,n_{klm},\,...} c_{n_{111},\,...,\,n_{klm},\,...}(t)\ |n_{111},\ ...,\ n_{klm},\ ...\rangle \tag{73}$$

The Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{74}$$

enables us to obtain the coefficients $c_{n_{111},\,...,\,n_{klm},\,...}(t)$ in the form:

$$c_{n_{111},\,...,\,n_{klm},\,...}(t) = c_{n_{111},\,...,\,n_{klm},\,...}(0)\ \mathrm{e}^{-i\sum_{klm} n_{klm}\,\omega_{klm}t} \tag{75}$$

γ. Field operator

Upon quantization, $\mathcal{S}(\mathbf{r},t)$ becomes an observable $S(\mathbf{r})$ which no longer depends on t and is obtained by replacing in (54) $q_{klm}(t)$ by Q_{klm}:

$$S(\mathbf{r}) = \sum_{klm} \frac{1}{\beta_{klm}} f_{klm}(\mathbf{r})\, \hat{Q}_{klm} \tag{76}$$

One can also, with the help of (69), express $S(\mathbf{r})$ in terms of the creation and annihilation operators:

$$S(\mathbf{r}) = \frac{1}{\sqrt{2}} \sum_{klm} \frac{1}{\beta_{klm}} f_{klm}(\mathbf{r}) \left[a_{klm} + a^{\dagger}_{klm} \right] \tag{77}$$

The same arguments as in § 2-e-δ enable us to show, using formulas (71) and (75), that the only Bohr frequencies that can appear in the time evolution of the mean value of the field:

$$\langle S(\mathbf{r}) \rangle(t) = \langle \psi(t) | S(\mathbf{r}) | \psi(t) \rangle$$

are the frequencies $\omega_{111}/2\pi$, $\omega_{211}/2\pi$, ..., $\omega_{klm}/2\pi$, ... associated, respectively, with the spatial functions $f_{111}(\mathbf{r})$, $f_{211}(\mathbf{r})$, ... $f_{klm}(\mathbf{r})$... Thus we find for $\langle S(\mathbf{r}) \rangle(t)$ a linear superposition of the classical standing waves that can exist in the cavity. A calculation identical to the one in § 2-e-δ would enable us to show that $\langle S(\mathbf{r}) \rangle$ satisfies equation (48).

Finally, we find that $S(\mathbf{r})$ and H do not commute. It is therefore impossible, in quantum theory, to know simultaneously and with certainty both the number of photons and the value of the electromagnetic field at a point in space.

Comment:

For the electromagnetic field, coherent states can be constructed which are analogous to the ones introduced in Complement G_V and which represent the best possible compromise between the incompatible quantities, field and energy.

δ. Vacuum fluctuations

We saw in § D-1 of Chapter V that, in the ground state of a harmonic oscillator, $\langle X \rangle$ is zero while $\langle X^2 \rangle$ is not, and we discussed the physical meaning of this typically quantum mechanical effect.

In the problem we are studying here, $S(\mathbf{r})$ presents many analogies with the X operator of Chapter V; we see from (77) that $S(\mathbf{r})$ is a *linear* combination of creation and annihilation operators. Consider the mean value of $S(\mathbf{r})$ in the ground state $|0\rangle$ of the field, that is, the "vacuum" state of photons. Since the diagonal elements of a and $a^{\dagger}$ are zero according to (71), we see that:

$$\langle 0 | S(\mathbf{r}) | 0 \rangle = 0 \tag{78}$$

On the other hand, the corresponding matrix element of $[S(\mathbf{r})]^2$ is not zero. According to (71):

$$\begin{cases} a_{klm}|0\rangle = 0 \\ \langle 0 | a^{\dagger}_{k'l'm'} = 0 \\ \langle 0 | a_{k'l'm'} a^{\dagger}_{klm} | 0 \rangle = \delta_{kk'}\ \delta_{ll'}\ \delta_{mm'} \end{cases} \tag{79}$$

Therefore, a simple calculation enables us to establish, using (77), that:

$$\langle 0[S(\mathbf{r})]^2|0\rangle = \frac{1}{2}\sum_{klm}\frac{1}{\beta_{klm}^2}\left[f_{klm}(\mathbf{r})\right]^2 \tag{80}$$

From this we see that in the vacuum, that is, in the absence of photons, the electromagnetic field $S(\mathbf{r})$ at a point of space has a zero mean value but a *non-zero root mean square deviation*. This means, for example, that if we perform *one* measurement of $S(\mathbf{r})$, we can find a non-zero result (varying, of course, from one measurement to another), even if there is no photon present in space. This effect has no equivalent in classical theory, in which, when the energy is zero, the field is rigorously zero. The preceding result is often expressed by saying that the "vacuum state" of photons is subjected to fluctuations of the field, characterized by (78) and (80) and called *vacuum fluctuations*.

The existence of these fluctuations has interesting physical consequences for the interaction of an atomic system with the electromagnetic field. Consider, for example, an atom in a state $|E\rangle$ of energy E, interacting with a *classically represented* electromagnetic wave. We shall see in Complement A_{XIII}, using time-dependent perturbation theory (*cf.* Chap. XIII), that under the effect of such an excitation, the atom can move to a higher energy state (*absorption*) or to a lower energy state (*induced emission*). But in this semi-classical treatment, if no field is present in space, the atom must remain indefinitely in the state $|E\rangle$. In fact, we have just established that, even in the absence of incident photons, the atom "sees" the "vacuum fluctuations" related to the quantum mechanical nature of the electromagnetic field. Under the effect of these fluctuations, it can emit a photon and fall back into a lower energy state (the energy of the global system being conserved during this process). This is the phenomenon of *spontaneous emission*, which can thus be considered to be, as it were, an "emission induced by the vacuum fluctuations". (No spontaneous absorption is possible, since this would cause the atom to move to a higher energy state, and no electromagnetic energy can be extracted from the field, which is in its ground state.)

It can also be shown that another effect of "vacuum fluctuations" is to impart to the atomic electrons an erratic motion which slightly modifies the energies of the levels. The observation of this effect in the hydrogen atom spectrum (the "Lamb shift") constituted the point of departure for the development of modern quantum electrodynamics.

Comment:

In the preceding discussions, we have always chosen the energy of the vacuum state as the energy origin. In fact, harmonic oscillator theory gives us the absolute value of the energy of the vacuum state:

$$E_0 = \sum_{klm}\frac{1}{2}\hbar\omega_{klm} \tag{81}$$

There is obviously a close relationship between E_0 and the electrical and magnetic energy associated with "vacuum fluctuations". One of the difficulties of quantum electrodynamics, of which we have just given a brief overview, is that the sum (81) is in fact infinite, as is, moreover, (80)! Nevertheless, it is possible to surmount this difficulty: using the procedure called "renormalization", one manages to bypass infinite quantities and calculate the physical effects that are actually observable, such as the "Lamb shift", with remarkable accuracy. It is obviously out of the question to consider these vast problems here.

References and suggestions for further reading:

Vibration modes of a continuous string in classical mechanics: Berkeley 3 (1.1), §§ 2.1, 2.2 and 2.3.

Quantization of the electromagnetic field: Mandl (2.9); Schiff (1.18), Chap. 14; Messiah (1.17), Chap. XXI; Bjorken and Drell (2.10), Chap. 11; Power (2.11); Heitler (2.13).

The Lamb shift: Lamb and Retherford (3.11); Frisch (3.13); Kuhn (11.1), Chap. III, § A 5 e; Series (11.7), Chaps. VIII, IX and X.

Complement L_V

One-dimensional harmonic oscillator in thermodynamic equilibrium at a temperature T

1 Mean value of the energy 648
1-a Partition function 648
1-b Calculation of $\langle H \rangle$ 648
2 Discussion 650
2-a Comparison with the classical oscillator 650
2-b Comparison with a two-level system 651
3 Applications 651
3-a Blackbody radiation 651
3-b Bose-Einstein distribution law 652
3-c Specific heats of solids at constant volume 653
4 Probability distribution of the observable X 655
4-a Definition of the probability density $\rho(x)$ 655
4-b Calculation of $\rho(x)$ 656
4-c Discussion 658
4-d Bloch's theorem 659

This complement is devoted to the study of the physical properties of a one-dimensional harmonic oscillator in thermodynamic equilibrium with a reservoir at temperature T. We know (*cf.* Complement E_{III}) that such an oscillator is not in a pure state (it is impossible to describe its state by a ket $|\psi\rangle$). The partial information which we possess about it and the results of statistical mechanics enable us to characterize it by a statistical mixture of stationary states $|\varphi_n\rangle$ with weights proportional to e^{-E_n/k_BT} (k_B: Boltzmann constant; E_n: energy of the state $|\varphi_n\rangle$). We saw in Complement E_{III} (§ 5-a) that the corresponding density operator is then written:

$$\rho = Z^{-1} e^{-H/k_BT} \tag{1}$$

where H is the Hamiltonian operator, and:

$$Z = \mathrm{Tr}\, e^{-H/k_BT} \tag{2}$$

is a normalization factor which insures that:

$$\mathrm{Tr}\, \rho = 1 \tag{3}$$

(Z is the partition function, *cf.* Appendix VI).

We shall calculate the mean value $\langle H \rangle$ of the oscillator's energy, interpret the result obtained, and show how it enters into numerous problems in physics (blackbody radiation, specific heat of solids, ...). Finally, we shall establish and discuss the expression for the probability density of the particle's position (the observable X).

1. Mean value of the energy

1-a. Partition function

The energies E_n of the states $|\varphi_n\rangle$ are, according to the results of § B of Chapter V, equal to $(n + 1/2)\hbar\omega$. Since the energy levels are not degenerate, we have, according to (2):

$$\begin{aligned} Z &= \sum_{n=0}^{\infty} \langle\varphi_n|\mathrm{e}^{-H/k_BT}|\varphi_n\rangle \\ &= \sum_{n=0}^{\infty} \mathrm{e}^{-(n+1/2)\hbar\omega/k_BT} \\ &= \mathrm{e}^{-\hbar\omega/2k_BT}\left[1 + \mathrm{e}^{-\hbar\omega/k_BT} + \mathrm{e}^{-2\hbar\omega/k_BT} + ...\right] \end{aligned} \tag{4}$$

Inside the brackets, we recognize a geometric progression of common ratio $\mathrm{e}^{-\hbar\omega/k_BT}$. Therefore:

$$Z = \frac{\mathrm{e}^{-\hbar\omega/2k_BT}}{1 - \mathrm{e}^{-\hbar\omega/k_BT}} \tag{5}$$

1-b. Calculation of $\langle H\rangle$

According to formula (31) of Complement E_{III} and expression (1) for ρ:

$$\langle H\rangle = \mathrm{Tr}(H\rho) = Z^{-1}\,\mathrm{Tr}(H\,\mathrm{e}^{-H/k_BT}) \tag{6}$$

Writing the trace explicitly in the $\{|\varphi_n\rangle\}$ basis, we obtain:

$$\langle H\rangle = Z^{-1}\sum_{n=0}^{\infty}(n + 1/2)\hbar\omega\,\mathrm{e}^{-(n+1/2)\hbar\omega/k_BT} \tag{7}$$

To calculate this quantity, we differentiate both sides of (4) with respect to T:

$$\frac{\mathrm{d}Z}{\mathrm{d}T} = \frac{1}{k_BT^2}\sum_{n=0}^{\infty}(n + 1/2)\hbar\omega\ \mathrm{e}^{-(n+1/2)\hbar\omega/k_BT} \tag{8}$$

We see that:

$$\langle H\rangle = k_BT^2\frac{1}{Z}\frac{\mathrm{d}Z}{\mathrm{d}T} \tag{9}$$

Using (5), we then find, after a simple calculation:

$$\boxed{\langle H\rangle = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{\mathrm{e}^{\hbar\omega/k_BT} - 1}} \tag{10}$$

Comments:

(i) Isotropic three-dimensional oscillator

Using the results and notation of Complement $\mathrm{E_V}$, we can write:

$$\langle H \rangle = \langle H_x \rangle + \langle H_y \rangle + \langle H_z \rangle \tag{11}$$

where $\langle H_x \rangle$ is given by:

$$\begin{aligned}\langle H_x \rangle &= Z^{-1} \operatorname{Tr}(H_x \, \mathrm{e}^{-H/k_B T}) \\ &= \frac{\displaystyle\sum_{n_x=0}^{\infty} \sum_{n_y=0}^{\infty} \sum_{n_z=0}^{\infty} (n_x + 1/2)\hbar\omega \; \mathrm{e}^{-\left[(n_x+1/2)+(n_y+1/2)+(n_z+1/2)\right]\hbar\omega/k_B T}}{\displaystyle\sum_{n_x=0}^{\infty} \sum_{n_y=0}^{\infty} \sum_{n_z=0}^{\infty} \mathrm{e}^{-\left[(n_x+1/2)+(n_y+1/2)+(n_z+1/2)\right]\hbar\omega/k_B T}}\end{aligned} \tag{12}$$

The sums over n_y and n_z can be separated out and are identical in the numerator and denominator, so that:

$$\langle H_x \rangle = \frac{\displaystyle\sum_{n_x=0}^{\infty} (n_x + 1/2)\hbar\omega \, \mathrm{e}^{-(n_x+1/2)\hbar\omega/k_B T}}{\displaystyle\sum_{n_x=0}^{\infty} \mathrm{e}^{-(n_x+1/2)\hbar\omega/k_B T}} \tag{13}$$

Aside from the replacement of n by n_x, the final expression is identical to the one calculated in the preceding section; $\langle H_x \rangle$ is therefore equal to the value given in (10). It is easy to show that the same is true for $\langle H_y \rangle$ and $\langle H_z \rangle$. Therefore, we have established the following result: at thermodynamic equilibrium, the mean energy of an isotropic three-dimensional oscillator is equal to three times that of a one-dimensional oscillator of the same angular frequency.

(ii) Classical oscillator

The energy $\mathcal{H}(x, p)$ of a classical one-dimensional oscillator is equal to:

$$\mathcal{H}(x, p) = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2 \tag{14}$$

In expression (14), x and p can take on any values between $-\infty$ and $+\infty$. According to the results of classical statistical mechanics, the mean energy of this classical oscillator is given by:

$$\langle \mathcal{H} \rangle = \frac{\displaystyle\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{H}(x, p) \; \mathrm{e}^{-\mathcal{H}(x,p)/k_B T} \; \mathrm{d}x \, \mathrm{d}p}{\displaystyle\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathcal{H}(x,p)/k_B T} \; \mathrm{d}x \, \mathrm{d}p} \tag{15}$$

Substituting (14) into (15), we find, after a simple calculation:

$$\langle \mathcal{H} \rangle = k_B T \tag{16}$$

An argument analogous to that of comment (i) shows that result (16) must be multiplied by 3 when we go from one to three dimensions.

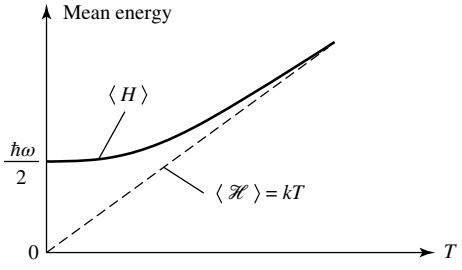


Figure 1: As a function of the temperature, variation of the mean energy of a quantum mechanical oscillator (solid line) compared with that of a classical oscillator (straight dashed line).

2. Discussion

2-a. Comparison with the classical oscillator

In Figure 1, the solid line gives the mean energy $\langle H \rangle$ of the one-dimensional quantum mechanical oscillator as a function of T. The dashed line corresponds to the mean energy $\langle \mathcal{H} \rangle$ of the classical oscillator.

For $T = 0$, $\langle H \rangle = \hbar\omega/2$. This result corresponds to the fact that at absolute zero, one is sure that the oscillator is in the ground state $|\varphi_0\rangle$, with energy $\hbar\omega/2$ ($\hbar\omega/2$ is, for this reason, sometimes called the "zero point energy"). As for the classical oscillator, it is motionless ($p = 0$) at its stable equilibrium position ($x = 0$), and its energy is zero: $\langle \mathcal{H} \rangle = 0$.

As long as T remains small – more precisely, as long as $k_B T \ll \hbar\omega$ – only the population of the ground state is appreciable, and $\langle H \rangle$ remains practically equal to $\hbar\omega/2$: in this region, the solid-line curve of Figure 1 has a horizontal tangent. We can see this directly from expression (10), which can be written, for small T:

$$\langle H \rangle = \frac{\hbar\omega}{2} + \hbar\omega\, \mathrm{e}^{-\hbar\omega/k_B T} + \ldots \tag{17}$$

On the other hand, for large T, that is, for $k_B T \gg \hbar\omega$, the same formula yields:

$$\langle H \rangle = \frac{\hbar\omega}{2} + k_B T \left(1 - \frac{1}{2}\frac{\hbar\omega}{k_B T} + \ldots\right) \tag{18}$$

or:

$$\langle H \rangle \simeq k_B T \tag{19}$$

to within an infinitesimal of the order of $k_B T(\hbar\omega/k_B T)^2$. The asymptote of the curve giving $\langle H \rangle$ as a function of T is therefore the straight line $\langle \mathcal{H} \rangle = k_B T$.

In conclusion, the quantum mechanical and classical oscillators have the same mean energy, $k_B T$, at high temperatures ($k_B T \gg \hbar\omega$). Striking differences appear at low temperatures ($k_B T \lesssim \hbar\omega$): it is no longer possible to ignore the quantization of the oscillator's energy once the energy $k_B T$ that characterizes the reservoir becomes of the order of (or smaller than) the energy difference $\hbar\omega$ separating two adjacent levels of the oscillator.

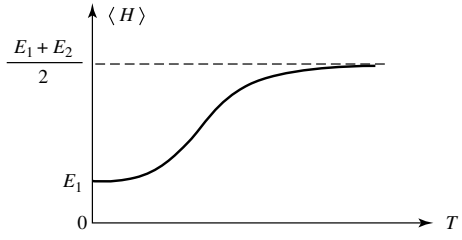


Figure 2: Mean energy of a quantum mechanical system with two energy levels E_1 and E_2, in thermodynamic equilibrium at a temperature T.

2-b. Comparison with a two-level system

It is interesting to compare the preceding results with those obtained for a two-level system. Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be the corresponding states, with energies E_1 and E_2 (with $E_1 < E_2$). For such a system, the general equation (6) yields:

$$\langle H\rangle = \frac{E_1\, \mathrm{e}^{-E_1/k_BT} + E_2\, \mathrm{e}^{-E_2/k_BT}}{\mathrm{e}^{-E_1/k_BT} + \mathrm{e}^{-E_2/k_BT}} \tag{20}$$

The mean energy of a two-level system, given by (20), is shown in Figure 2. For small T ($k_BT \ll E_2 - E_1$), the terms in e^{-E_1/k_BT} are preponderant in both the numerator and the denominator of (20) (since $E_1 < E_2$) and we obtain:

$$\langle H\rangle \underset{T\to 0}{\longrightarrow} E_1 \tag{21}$$

It can be verified that the curve starts with a horizontal tangent. For large T ($k_BT \gg E_2 - E_1$), the asymptote of the curve is the straight line parallel to the T-axis of ordinate $(E_1 + E_2)/2$. The preceding results are easy to understand: for $T = 0$, the system is in its ground state $|\psi_1\rangle$, of energy E_1; at high temperatures, the populations of the two levels are practically equal, and $\langle H\rangle$ approaches half the sum of the two energies E_1 and E_2.

Although the solid-line curves of Figures 1 and 2 have the same shape at low temperatures, we see that this is not at all true at high temperatures. For the harmonic oscillator, $\langle H\rangle$ is not bounded and increases linearly with T, while, for a two-level system, $\langle H\rangle$ cannot exceed a certain value. This difference is due to the fact that the energy spectrum of the harmonic oscillator extends upward indefinitely: when T increases, levels of higher and higher n are occupied, and this causes $\langle H\rangle$ to increase. On the other hand, for a two-level system, once the populations of the two levels are equalized, an additional increase in the temperature does not change the mean energy.

3. Applications

3-a. Blackbody radiation

We have already pointed out, in the introduction to Chapter V (and in Complement $\mathrm{K_V}$, where we justified this result more precisely), that the electromagnetic field in a cavity is equivalent to a set of independent one-dimensional harmonic oscillators. Each of these oscillators is associated with one of the standing waves that can exist inside the cavity (normal modes) and has the same angular frequency as this wave. Let us show that this result, combined with those

obtained above for $\langle \mathcal{H} \rangle$ and $\langle H \rangle$, leads very simply to the Rayleigh-Jeans law and the Planck law for blackbody radiation.

Let $\mathcal{V}$ be the volume of the cavity, whose walls are assumed to be perfectly reflecting. The first modes of the cavity (those of lowest frequency) depend strongly on the form of the cavity. On the other hand, for the high-frequency modes (those whose wavelength $\lambda = c/\nu$ is much smaller than the dimensions of the cavity), a classical electromagnetic calculation shows that, if $N(\nu)\, \mathrm{d}\nu$ denotes the number of modes whose frequency is between ν and $\nu + \mathrm{d}\nu$, $N(\nu)$ is practically independent of the form of the cavity and equal to:

$$N(\nu) = \frac{8\pi\nu^2}{c^3}\mathcal{V} \tag{22}$$

Let $u(\nu)\, \mathrm{d}\nu$ be the electromagnetic energy per unit volume of the cavity contained in the frequency band $(\nu, \nu + \mathrm{d}\nu)$ when the cavity is in thermodynamic equilibrium at a temperature T. To obtain the energy $\mathcal{V}\, u(\nu)\, \mathrm{d}\nu$, one must multiply the number of modes whose frequency is between ν and $\nu + \mathrm{d}\nu$ by the mean energy of the corresponding harmonic oscillators. We calculated this energy above; it is equal[1] to $\langle \mathcal{H} \rangle$ or $\langle H \rangle - \hbar\omega/2$, depending on whether the problem is treated classically or quantum mechanically. We then obtain, using (10), (16) and (22):

$$u_{cl}(\nu) = \frac{8\pi\nu^2}{c^3} k_B T \tag{23}$$

in a classical treatment, and:

$$u_Q(\nu) = \frac{8\pi\nu^2}{c^3} \frac{1}{\mathrm{e}^{h\nu/k_B T} - 1} \tag{24}$$

in a quantum mechanical treatment.

We recognize (23) to be Rayleigh-Jeans' law and (24) to be Planck's law, which reduces to the preceding one in the limit of low frequencies or high temperatures ($h\nu/k_B T \ll 1$). The differences between these two laws reflect those which exist between the two curves of Figure 1. At high frequencies, difficulties arise in Rayleigh-Jeans' law: the quantity $u_{cl}(\nu)$ given in (23) approaches infinity when $\nu \longrightarrow \infty$, which is physically absurd. In order to remedy this defect, Planck was led to postulate that the energy of each oscillator varied discontinuously, by jumps proportional to ν (energy quantization); thus he obtained formula (24), which accounts perfectly for the experimental results.

3-b. Bose-Einstein distribution law

Instead of calculating the mean value $\langle H \rangle$ of the energy, as we did in § 1, let us calculate the mean value of the operator N. Since, according to formula (B-15) of Chapter V:

$$H = \left(N + \frac{1}{2}\right)\hbar\omega \tag{25}$$

we deduce from result (10) that:

$$\langle N \rangle = \frac{1}{\mathrm{e}^{h\nu/k_B T} - 1} \tag{26}$$

[1] We use $\langle H \rangle - \hbar\omega/2$ and not $\langle H \rangle$ for the following reason: $u(\nu)$ represents an electromagnetic energy which can be extracted from the cavity. At absolute zero, all the oscillators are in their ground states and no energy can be radiated outward because the system is in its lowest energy state; $u(\nu)$ must therefore be zero at absolute zero, as experimental observations indeed show it to be. This requires us to define the mean energy of the field in the cavity with respect to the value corresponding to $T = 0$.

The fact that the levels of a one-dimensional harmonic oscillator are equidistant enables us to associate with the oscillator in the state $|\varphi_n\rangle$ a set of n identical particles (quanta) of the same energy $h\nu$. In this interpretation, the operators $a^\dagger$ and a, which take $|\varphi_n\rangle$ into $|\varphi_{n+1}\rangle$ or $|\varphi_{n-1}\rangle$, create or destroy a particle. N is thus the operator associated with the number of particles ($|\varphi_n\rangle$ is the eigenstate of N with the eigenvalue n).

In the special case of the electromagnetic field, the quanta associated with each harmonic oscillator are none other than *photons*. To each mode of the cavity considered in the preceding paragraph correspond photons of a certain type, characterized by the frequency, polarization, and spatial distribution of the mode. Expression (26) gives the mean number of photons associated with a mode of frequency ν at thermodynamic equilibrium. We recognize (26) to be the Bose-Einstein distribution law, which can be derived in a more general way; here, we have established it very simply by studying the harmonic oscillator and interpreting the states $|\varphi_n\rangle$.

Comment:

To be rigorous, we should write the Bose-Einstein distribution law for bosons of energy ε:

$$\langle N\rangle = \frac{1}{\mathrm{e}^{(\varepsilon-\mu)/k_BT}-1} \tag{27}$$

where μ is the chemical potential. In the case of photons, $\mu = 0$. This is due to the fact that the total number of photons in the global system radiation-reservoir is not fixed, because of the possibility of absorption or emission of photons by the walls.

3-c. Specific heats of solids at constant volume

We shall confine ourselves here to the Einstein model (*cf.* Complement $\mathrm{A_V}$), in which a solid is considered to be composed of $\mathcal{N}$ atoms vibrating independently about their equilibrium positions with the same angular frequency ω_E. The internal energy U of the solid at the temperature T is therefore equal to the sum of the mean energies of the $\mathcal{N}$ isotropic three-dimensional oscillators in thermodynamic equilibrium at this temperature. Using comment (i) of § 1, we see that:

$$U = 3\mathcal{N}\langle H\rangle \tag{28}$$

where $\langle H\rangle$ is the mean energy of a one-dimensional harmonic oscillator of angular frequency ω_E. We know that the constant volume specific heat c_V is the derivative of the internal energy U with respect to the temperature:

$$c_V = \frac{\mathrm{d}U}{\mathrm{d}T} = 3\mathcal{N}\frac{\mathrm{d}}{\mathrm{d}T}\langle H\rangle \tag{29}$$

which, taking (10) into account, yields:

$$c_V = 3\mathcal{N}k_B\frac{\left(\frac{\hbar\omega_E}{k_BT}\right)^2\mathrm{e}^{\hbar\omega_E/k_BT}}{\left[\mathrm{e}^{\hbar\omega_E/k_BT}-1\right]^2} \tag{30}$$

The variation of c_V with T is shown in Figure 3. According to (29), c_V is proportional to the derivative of the solid-line curve of Figure 1. It is therefore very simple to describe the behavior of the specific heat c_V as a function of the temperature. In Figure 1, we see that $\langle H\rangle$ has a horizontal tangent at the origin and increases very slowly; c_V is therefore zero for $T = 0$ and also increases very slowly. On the other hand, for large T ($k_BT \gg \hbar\omega_E$), $\langle H\rangle$ approaches k_BT; we deduce that c_V approaches a constant, $3\mathcal{N}k_B$, independent of ω_E. The transition region corresponds to $\hbar\omega_E/k_BT \simeq 1$.

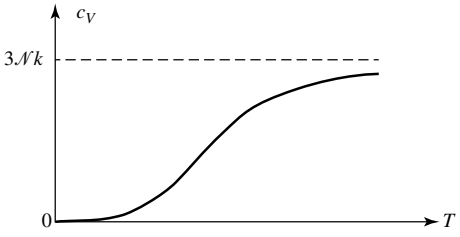


Figure 3: (Constant volume) specific heat c_V of a solid in Einstein's model. The high-temperature limit corresponds to the classical Dulong-Petit law.

The asymptote of Figure 3 corresponds to the Dulong-Petit law: if one takes an gram atom of any solid, $\mathcal{N}$ is equal to Avogadro's number and the limiting value of c_V is equal to $3\,R$ (where R is the ideal gas constant), that is, to about 6 cal. degree^{-1} mole^{-1}.

As we pointed out above, the quantum mechanical nature of crystalline vibrations manifests itself at low temperatures when k_BT becomes of the order of $\hbar\omega_E$ or less. Insofar as c_V is concerned, this means that the specific heat approaches zero when T approaches zero. It is as if the degrees of freedom corresponding to crystalline vibrations were "frozen" beneath a certain temperature and no longer entered into the specific heat. This can be understood physically: at absolute zero, each oscillator is in its ground state $|\varphi_0\rangle$; as long as $k_BT \ll \hbar\omega_E$, it cannot absorb any thermal energy, since its first excited state has an energy far greater than k_BT.

Comments:

(*i*) *Comparison with the specific heat of a two-level system*

We can apply an analogous argument to a sample composed of a set of two-level systems (for example, a paramagnetic sample composed of $\mathcal{N}$ spin 1/2 particles): its specific heat c_V is given, to within a coefficient, by the derivative of the curve of Figure 2. For such a system, the variation of c_V with T is shown in Figure 4.

The behavior for $T \simeq 0$ is the same as in the case of Figure 3. However, we see that c_V approaches zero when $k_BT \gg E_2 - E_1$, since the mean energy then becomes independent of T and is equal to $(E_1 + E_2)/2$ (*cf.* Fig. 2). For a two-level system, c_V therefore has a maximum (Schottky anomaly) whose physical interpretation is the following: like the harmonic oscillator, the two-level system cannot absorb any thermal energy at very low temperatures, as long as $E_2 - E_1 \gg k_BT$; c_V is therefore zero at the origin. Then, as T increases, E_2 becomes populated, and c_V increases. When the temperature is high enough for the two populations to be practically equal, the system cannot absorb any more thermal energy, since the populations can no longer change: c_V therefore approaches zero when $T \longrightarrow \infty$.

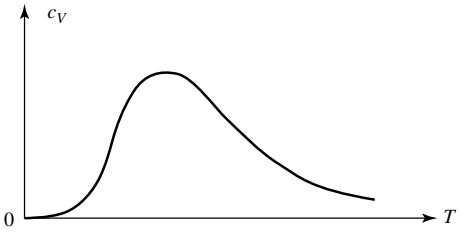


Figure 4: Specific heat c_V for a set of two-level systems. At high temperatures, c_V approaches zero because the energy spectrum has an upper bound.

(*ii*) Einstein's model enables us to understand simply why the specific heat c_V approaches zero when the temperature T approaches zero (a classically inexplicable result). However, it is too schematic to describe the exact dependence of c_V at low temperatures.

In a real crystal, the various oscillators are coupled. This gives rise to a set of vibrational normal modes (phonons) whose frequencies go from zero to a certain cutoff frequency (*cf.* Complement J_V). (30) must then be summed over the different possible frequencies ν (taking into account the fact that the number of modes whose frequencies are included between ν and $\nu + d\nu$ depends on ν). Thus one finds an expression for the specific heat which, at low temperatures, varies like T^3 (this is confirmed experimentally).

4. Probability distribution of the observable X

4-a. Definition of the probability density $\rho(x)$

Let us return to the one-dimensional harmonic oscillator in thermodynamic equilibrium. We seek the probability $\rho(x)\,dx$ of finding, in a measurement of the position X of the particle, a result included between x and $x + dx$. It is clear that $\rho(x)$ plays an important role in a large number of physical problems. For example, for a solid described by Einstein's model, the width of $\rho(x)$ gives an idea of the amplitude of atomic vibrations; the study of the variation of this width with respect to T makes it possible to understand the phenomenon of melting [which occurs when the width of $\rho(x)$ is no longer negligible compared to the interatomic distance].

When the oscillator is in the stationary state $|\varphi_n\rangle$, the corresponding probability density $\rho_n(x)$ is:

$$\rho_n(x) = |\varphi_n(x)|^2 = \langle x|\varphi_n\rangle \langle \varphi_n|x\rangle \tag{31}$$

At thermodynamic equilibrium, the oscillator is described by a statistical mixture of the states $|\varphi_n\rangle$ with the weights: $Z^{-1}\, e^{-E_n/k_BT}$. The probability density $\rho(x)$ is then:

$$\rho(x) = Z^{-1} \sum_n \rho_n(x) e^{-E_n/k_BT} \tag{32}$$

$\rho(x)$ is the weighted sum of the probability densities $\rho_n(x)$ associated with the various states $|\varphi_n\rangle$. Some of the $\rho_n(x)$ are shown in Figures 5 and 6 of Chapter V. We shall see later that the oscillations of the functions $\rho_n(x)$ which are visible in these figures disappear in the summation over n: we shall show that $\rho(x)$ is simply a Gaussian function.

The probability density $\rho(x)$ defined in (32) is related very simply to the density operator ρ of the harmonic oscillator in thermodynamic equilibrium. Using (31) and (32), we obtain:

$$\rho(x) = Z^{-1} \sum_n e^{-E_n/k_BT} \langle x|\varphi_n\rangle \langle \varphi_n|x\rangle \tag{33}$$

On the right-hand side, we can bring in the operator e^{-H/k_BT} which, taking into account the closure relation for the states $|\varphi_n\rangle$, can be written:

$$e^{-H/k_BT} = e^{-H/k_BT} \sum_n |\varphi_n\rangle \langle \varphi_n| = \sum_n e^{-E_n/k_BT} |\varphi_n\rangle \langle \varphi_n| \tag{34}$$

We then see that:

$$\rho(x) = Z^{-1}\langle x|\mathrm{e}^{-H/k_BT}|x\rangle = \langle x|\rho|x\rangle \tag{35}$$

where the density operator ρ is given by formula (1). $\rho(x)$ can then be seen to be the diagonal element of ρ which corresponds to the ket $|x\rangle$.

4-b. Calculation of $\rho(x)$

We know that:

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) \tag{36}$$

so that $\rho(x)$ can be written in the form:

$$\rho(x) = Z^{-1}\,\mathrm{e}^{-\lambda/2}\,F_\lambda(x) \tag{37}$$

with:

$$\lambda = \frac{\hbar\omega}{k_BT} \tag{38}$$

and:

$$F_\lambda(x) = \langle x|\mathrm{e}^{-\lambda a^\dagger a}|x\rangle \tag{39}$$

In order to know $\rho(x)$, all we need to do, therefore, is evaluate this diagonal matrix element.

To do this, let us calculate the variation in $F_\lambda(x)$ when x is changed to $x + \mathrm{d}x$. Since the ket $|x + \mathrm{d}x\rangle$ is given by [*cf.* Complement $\mathrm{E_{II}}$, relation (20)]:

$$|x + \mathrm{d}x\rangle = \left(1 - i\frac{\mathrm{d}x}{\hbar}P\right)|x\rangle \tag{40}$$

we obtain, substituting this relation and the adjoint relation into (39) (neglecting infinitesimals of second order in $\mathrm{d}x$):

$$F_\lambda(x + \mathrm{d}x) = F_\lambda(x) + i\frac{\mathrm{d}x}{\hbar}\langle x|\left[P, \mathrm{e}^{-\lambda a^\dagger a}\right]|x\rangle \tag{41}$$

The matrix element appearing on the right-hand side of (41) involves the operator P, which is proportional to $(a - a^\dagger)$. Now, it is the X operator, proportional to $(a + a^\dagger)$, which acts in a simple way on the kets $|x\rangle$. We shall therefore transform $[P, \mathrm{e}^{-\lambda a^\dagger a}]$ so as to make X appear. We shall begin by seeking the relation between $a\,\mathrm{e}^{-\lambda a^\dagger a}$ and $\mathrm{e}^{-\lambda a^\dagger a}a$. This can be obtained very simply in the $\{|\varphi_n\rangle\}$ representation:

$$a\,\mathrm{e}^{-\lambda a^\dagger a}|\varphi_n\rangle = \sqrt{n}\,\mathrm{e}^{-\lambda n}|\varphi_{n-1}\rangle \tag{42a}$$

$$\mathrm{e}^{-\lambda a^\dagger a}a|\varphi_n\rangle = \sqrt{n}\,\mathrm{e}^{-\lambda(n-1)}|\varphi_{n-1}\rangle \tag{42b}$$

that is:

$$\mathrm{e}^{-\lambda a^\dagger a}a = \mathrm{e}^{\lambda}a\,\mathrm{e}^{-\lambda a^\dagger a} \tag{43}$$

which can also be written:

$$\left(1 - \tanh\frac{\lambda}{2}\right)\mathrm{e}^{-\lambda a^\dagger a}a = \left(1 + \tanh\frac{\lambda}{2}\right)a\,\mathrm{e}^{-\lambda a^\dagger a} \tag{44}$$

Similarly, it can be shown that:

$$\mathrm{e}^{-\lambda a^\dagger a}\, a^\dagger = \mathrm{e}^{-\lambda}\, a^\dagger\, \mathrm{e}^{-\lambda a^\dagger a} \tag{45}$$

that is:

$$\left(1 + \tanh\frac{\lambda}{2}\right) \mathrm{e}^{-\lambda a^\dagger a}\, a^\dagger = \left(1 - \tanh\frac{\lambda}{2}\right) a^\dagger\, \mathrm{e}^{-\lambda a^\dagger a} \tag{46}$$

We now subtract, term by term, relations (44) and (46); we obtain:

$$\left[a - a^\dagger,\, \mathrm{e}^{-\lambda a^\dagger a}\right] = -\tanh\frac{\lambda}{2}\left[a + a^\dagger,\, \mathrm{e}^{-\lambda a^\dagger a}\right]_+ \tag{47}$$

where the symbol $[A, B]_+$ denotes the anticommutator:

$$[A, B]_+ = AB + BA \tag{48}$$

If we take into account the numerical factors which result from formulas (B-1) and (B-7) in Chapter V, (47) finally becomes:

$$\left[P,\, \mathrm{e}^{-\lambda a^\dagger a}\right] = im\omega\, \tanh\frac{\lambda}{2}\, \left[X,\, \mathrm{e}^{-\lambda a^\dagger a}\right]_+ \tag{49}$$

Substituting this result into relation (41):

$$\begin{aligned} F_\lambda(x + \mathrm{d}x) - F_\lambda(x) &= -\frac{m\omega}{\hbar}\, \mathrm{d}x\, \tanh\frac{\lambda}{2}\, \langle x|\left[X, \mathrm{e}^{-\lambda a^\dagger a}\right]_+ |x\rangle \\ &= -2x\frac{m\omega}{\hbar}\, \tanh\frac{\lambda}{2}\, F_\lambda(x)\, \mathrm{d}x \end{aligned} \tag{50}$$

$F_\lambda(x)$ therefore satisfies the differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}x}F_\lambda(x) + \frac{2x}{\xi^2}F_\lambda(x) = 0 \tag{51}$$

where ξ, which has the dimensions of a length, is defined by:

$$\xi = \sqrt{\frac{\hbar}{m\omega}\coth\frac{\lambda}{2}} = \sqrt{\frac{\hbar}{m\omega}\coth\left(\frac{\hbar\omega}{2k_BT}\right)} \tag{52}$$

Equation (51) can be integrated directly:

$$F_\lambda(x) = F_\lambda(0)\, \mathrm{e}^{-x^2/\xi^2} \tag{53}$$

Therefore, we know $\rho(x)$ to within a constant factor, since, according to (37):

$$\rho(x) = Z^{-1}\, \mathrm{e}^{-\lambda/2}\, F_\lambda(0)\, \mathrm{e}^{-x^2/\xi^2} \tag{54}$$

Since we know that the integral of $\rho(x)$ over the whole x-axis must be equal to 1, we obtain finally:

$$\rho(x) = \frac{1}{\xi\sqrt{\pi}}\mathrm{e}^{-x^2/\xi^2} \tag{55}$$

The function $\rho(x)$ is thus a Gaussian, whose width is characterized by the length ξ defined in (52).

4-c. Discussion

Starting from the probability density (55), it is easy to calculate:

$$\langle X \rangle = 0$$

$$\langle X^2 \rangle = (\Delta X)^2 = \frac{\xi^2}{2} \tag{56}$$

Figure 5 shows the variation of $(\Delta X)^2$ with respect to T. We see from (52) that $(\Delta X)^2$ is equal to $\hbar/2m\omega$ when $T = 0$. This result is not surprising: at $T = 0$, the oscillator is in its ground state, and $\rho(x)$ is equal to $|\varphi_0(x)|^2$; ΔX is found to be the root mean square deviation of X in the ground state [*cf.* formula (D-5a) of Chapter V]. Then, when T increases, so does $(\Delta X)^2$; when $k_B T \gg \hbar\omega$, we have:

$$(\Delta X)^2 \underset{T \to \infty}{\sim} \frac{k_B T}{m\omega^2} \tag{57}$$

In this case, $\rho(x)$ becomes identical to the probability density of a classical oscillator in thermodynamic equilibrium at the temperature T:

$$\rho_{cl}(x) = \frac{e^{-V(x)/k_B T}}{\displaystyle\int_{-\infty}^{+\infty} e^{-V(x)/k_B T} dx} = \frac{1}{\sqrt{\dfrac{2\pi k_B T}{m\omega^2}}} \, e^{-\dfrac{m\omega^2 x^2}{2k_B T}} \tag{58}$$

which leads to $(\Delta x_{cl})^2 = k_B T/m\omega^2$ (the straight dashed line in Figure 5). Here again, classical and quantum mechanical predictions meet for $k_B T \gg \hbar\omega$.

Now, let us apply the preceding results to the problem of melting of a solid body (for simplicity, we shall choose the one-dimensional Einstein model; see Complement A_V). Experiments show that the solid melts when ΔX is of the order of an appreciable fraction η of the interatomic distance d. Consequently, the melting point temperature T_f is given

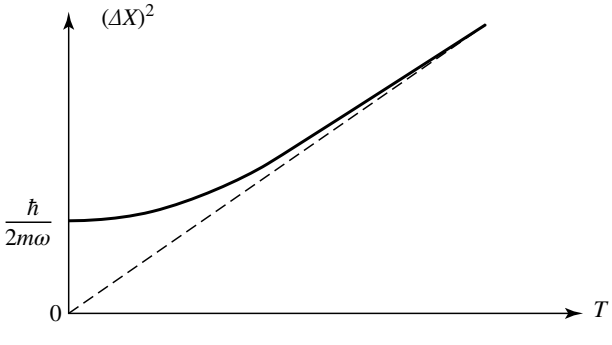


Figure 5: Variation with respect to the temperature T of $(\Delta X)^2$, for a harmonic oscillator in thermodynamic equilibrium. When $T \to \infty$, ΔX is identical to the classical value, shown by the dashed line; at low temperatures, quantum mechanical effects (Heisenberg uncertainty relation) prevent ΔX from approaching zero.

approximately by:

$$\frac{\xi^2}{2d^2} \simeq \eta^2 \tag{59}$$

where ξ can be replaced by its expression (52), with $T = T_f$. Assuming that T_f is large enough that $k_B T_f \gg \hbar\omega_E$, we can use[2] in (59) the asymptotic form (57), and we obtain the law for T_f:

$$\frac{k_B T_f}{m\omega_E^2} \simeq \eta^2 d^2 \tag{60}$$

If we set:

$$\hbar\omega_E = k_B \Theta_E \tag{61}$$

(Θ_E is called the "Einstein temperature") and if we note that d does not vary very much from one substance to another (anyway, much less than ω_E, that is, Θ_E), we find the approximate law:

$$\frac{T_f}{m\Theta_E^2} \simeq C^{te} \tag{62}$$

The melting point temperature of a crystal is therefore approximately proportional to the square of a vibrational frequency which is characteristic of the crystal.

4-d. Bloch's theorem

Consider the operator e^{-iqX}, where q is a real variable. Its mean value:

$$\langle \mathrm{e}^{-iqX} \rangle = \mathrm{Tr}\,[\rho\, \mathrm{e}^{-iqX}] \tag{63}$$

[where ρ is given by (1)] is a function of q, which we shall denote by $f(q)$:

$$f(q) = \langle \mathrm{e}^{-iqX} \rangle \tag{64}$$

In probability theory, $f(q)$ is called the characteristic function of the random variable x.

It is easy to calculate $f(q)$ if we place ourselves in the $\{|x\rangle\}$ representation:

$$\begin{aligned} f(q) &= \int_{-\infty}^{+\infty} \mathrm{d}x\, \langle x|\rho\, \mathrm{e}^{-iqX}|x\rangle \\ &= \int_{-\infty}^{+\infty} \mathrm{d}x\, \langle x|\rho|x\rangle\, \mathrm{e}^{-iqx} \\ &= \int_{-\infty}^{+\infty} \mathrm{d}x\, \rho(x)\, \mathrm{e}^{-iqx} \end{aligned} \tag{65}$$

To within a factor of $\sqrt{2\pi}$, $f(q)$ is therefore the Fourier transform of the function $\rho(x)$ calculated above (§ 4-b). Since $\rho(x)$ is a Gaussian [formula (55)], $f(q)$ is also a Gaussian [*cf.* Appendix I, relation (50)]:

$$f(q) = \mathrm{e}^{-\xi^2 q^2/4} \tag{66}$$

[2] This is not always possible. Recall that helium remains liquid at atmospheric pressure, even at $T = 0$: ξ is never negligible compared to d, whatever T may be (*cf.* Complement $\mathrm{A_V}$).

which, according to formula (56), can be written:

$$\langle e^{-iqX} \rangle = e^{-\frac{q^2}{2}\langle X^2 \rangle} \tag{67}$$

We could perform calculations analogous to those of §§ 4-a and 4-b above for the observable P instead of X. We would then define the probability density $\overline{\rho}(p)$ by:

$$\overline{\rho}(p) = Z^{-1} \sum_n e^{-E_n/k_B T} |\overline{\varphi}_n(p)|^2 \tag{68}$$

Formula (24) of Complement D_V shows that:

$$\overline{\rho}(p) = \frac{1}{m\omega}\ \rho\left(x = \frac{p}{m\omega}\right) \tag{69}$$

Therefore:

$$\overline{\rho}(p) = \frac{1}{m\omega\xi\sqrt{\pi}}\ e^{-\frac{p^2}{m^2\omega^2\xi^2}} \tag{70}$$

Consequently, the study of $\langle e^{-iqP} \rangle$ would lead to the same result as in (67):

$$\langle e^{-iqP} \rangle = e^{-\frac{q^2}{2}\langle P^2 \rangle} \tag{71}$$

The generalization of formulas (67) and (71) is known as Bloch's theorem: if $G(X, P)$ is an arbitrary linear combination of the position X and the momentum P of a one-dimensional harmonic oscillator in thermodynamic equilibrium at the temperature T, then:

$$\langle e^{-iqG} \rangle = e^{-\frac{q^2}{2}\langle G^2 \rangle} \tag{72}$$

This theorem is used in solid state physics, for example in the theory of emission without recoil by the nuclei of a crystalline lattice (the Mössbauer effect).

References and suggestions for further reading:

Specific heats: Kittel (8.2), Chap. 6, p. 91 and 100; Kittel (13.2), Chap. 6; Seitz (13.4), Chap. III; Ziman (13.3), Chap. 2.

Blackbody radiation: Eisberg and Resnick (1.3), Chap. 1; Kittel (8.2), Chap. 15; Reif (8.4) § 9-13 to 9-15; Bruhat (8.3), Chap. XXII.

Bloch's theorem: Messiah (1.17), Chap. XII, § 11-12.

Complement M_V

Exercises

1. Consider a harmonic oscillator of mass m and angular frequency ω. At time $t = 0$, the state of this oscillator is given by:

$$|\psi(0)\rangle = \sum_n c_n |\varphi_n\rangle$$

where the states $|\varphi_n\rangle$ are stationary states with energies $(n + 1/2)\hbar\omega$.

a. What is the probability $\mathcal{P}$ that a measurement of the oscillator's energy performed at an arbitrary time $t > 0$, will yield a result greater than $2\hbar\omega$? When $\mathcal{P} = 0$, what are the non-zero coefficients c_n?

b. From now on, assume that only c_0 and c_1 are different from zero. Write the normalization condition for $|\psi(0)\rangle$ and the mean value $\langle H\rangle$ of the energy in terms of c_0 and c_1. With the additional requirement $\langle H\rangle = \hbar\omega$, calculate $|c_0|^2$ and $|c_1|^2$.

c. As the normalized state vector $|\psi(0)\rangle$ is defined only to within a global phase factor, we fix this factor by choosing c_0 real and positive. We set: $c_1 = |c_1|e^{i\theta_1}$. We assume that $\langle H\rangle = \hbar\omega$ and that:

$$\langle X\rangle = \frac{1}{2}\sqrt{\frac{\hbar}{m\omega}}$$

Calculate θ_1.

d. With $|\psi(0)\rangle$ so determined, write $|\psi(t)\rangle$ for $t > 0$ and calculate the value of θ_1 at t. Deduce the mean value $\langle X\rangle(t)$ of the position at t.

2. Anisotropic three-dimensional harmonic oscillator

In a three-dimensional problem, consider a particle of mass m and of potential energy:

$$V(X,Y,Z) = \frac{m\omega^2}{2}\left[\left(1 + \frac{2\lambda}{3}\right)\left(X^2 + Y^2\right) + \left(1 - \frac{4\lambda}{3}\right)Z^2\right]$$

where ω and λ are constants which satisfy:

$$\omega \geq 0 \quad , \qquad 0 \leq \lambda < \frac{3}{4}$$

a. What are the eigenstates of the Hamiltonian and the corresponding energies?

b. Calculate and discuss, as functions of λ, the variation of the energy, the parity and the degree of degeneracy of the ground state and the first two excited states.

3. Harmonic oscillator: two particles

Two particles of the same mass m, with positions X_1 and X_2 and momenta P_1 and P_2, are subjected to the same potential:

$$V(X) = \frac{1}{2} m\omega^2 X^2$$

The two particles do not interact.

a. Write the operator H, the Hamiltonian of the two-particle system. Show that H can be written:

$$H = H_1 + H_2$$

where H_1 and H_2 act respectively only in the state space of particle (1) and in that of particle (2). Calculate the energies of the two-particle system, their degrees of degeneracy, and the corresponding wave functions.

b. Does H form a C.S.C.O.? Same question for the set $\{H_1, H_2\}$. We denote by $|\Phi_{n_1,n_2}\rangle$ the eigenvectors common to H_1 and H_2. Write the orthonormalization and closure relations for the states $|\Phi_{n_1,n_2}\rangle$.

c. Consider a system which, at $t = 0$, is in the state:

$$|\psi(0)\rangle = \frac{1}{2}\left(|\Phi_{0,0}\rangle + |\Phi_{1,0}\rangle + |\Phi_{0,1}\rangle + |\Phi_{1,1}\rangle\right)$$

What results can be found, and with what probabilities, if at this time one measures:

- the total energy of the system?
- the energy of particle (1)?
- the position or velocity of this particle?

4. (This exercise is a continuation of the preceding one and uses the same notation.)
The two-particle system, at $t = 0$, is in the state $|\psi(0)\rangle$ given in exercise 3.

a. At $t = 0$, one measures the total energy H and one finds the result $2\hbar\omega$.

α Calculate the mean values of the position, the momentum, and the energy of particle (1) at an arbitrary positive t. Same question for particle (2).

β At $t > 0$, one measures the energy of particle (1). What results can be found, and with what probabilities? Same question for a measurement of the position of particle (1); trace the curve for the corresponding probability density.

b. Instead of measuring the total energy H, at $t = 0$, one measures the energy H_2 of particle (2); the result obtained is $\hbar\omega/2$. What happens to the answers to questions α and β of *a*?

5. (This exercise is a continuation of exercise 3 and uses the same notation.)

We denote by $|\Phi_{n_1,n_2}\rangle$ the eigenstates common to H_1 and H_2, of eigenvalues $(n_1 + 1/2)\hbar\omega$ and $(n_2 + 1/2)\hbar\omega$. The "two particle exchange" operator P_e is defined by:

$$P_e|\Phi_{n_1,n_2}\rangle = |\Phi_{n_2,n_1}\rangle$$

a. Prove that $P_e^{-1} = P_e$ and that P_e is unitary. What are the eigenvalues of P_e? Let $B' = P_e B P_e^\dagger$ be the observable resulting from the transformation by P_e of an arbitrary observable B. Show that the condition $B' = B$ (B invariant under exchange of the two particles) is equivalent to $[B, P_e] = 0$.

b. Show that:

$$P_e H_1 P_e^\dagger = H_2$$

$$P_e H_2 P_e^\dagger = H_1$$

Does H commute with P_e? Calculate the action of P_e on the observables X_1, P_1, X_2, P_2.

c. Construct a basis of eigenvectors common to H and P_e. Do these two operators form a C.S.C.O.? What happens to the spectrum of H and the degeneracy of its eigenvalues if one retains only the eigenvectors $|\Phi\rangle$ of H for which $P_e|\Phi\rangle = -|\Phi\rangle$?

6. Charged harmonic oscillator in a variable electric field

A one-dimensional harmonic oscillator is composed of a particle of mass m, charge q and potential energy $V(X) = \frac{1}{2}m\omega^2 X^2$. We assume in this exercise that the particle is placed in an electric field $\mathscr{E}(t)$ parallel to Ox and time-dependent, so that to $V(x)$ must be added the potential energy:

$$W(t) = -q\mathscr{E}(t)\, X$$

a. Write the Hamiltonian $H(t)$ of the particle in terms of the operators a and $a^\dagger$. Calculate the commutators of a and $a^\dagger$ with $H(t)$.

b. Let $\alpha(t)$ be the number defined by:

$$\alpha(t) = \langle\psi(t)|a|\psi(t)\rangle$$

where $|\psi(t)\rangle$ is the normalized state vector of the particle under study. Show from the results of the preceding question that $\alpha(t)$ satisfies the differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha(t) = -i\omega\,\alpha(t) + i\lambda(t)$$

where $\lambda(t)$ is defined by:

$$\lambda(t) = \frac{q}{\sqrt{2m\hbar\omega}}\mathscr{E}(t)$$

Integrate this differential equation. At time t, what are the mean values of the position and momentum of the particle?

c. The ket $|\varphi(t)\rangle$ is defined by:

$$|\varphi(t)\rangle = [a - \alpha(t)]|\psi(t)\rangle$$

where $\alpha(t)$ has the value calculated in *b*. Using the results of questions *a* and *b*, show that the evolution of $|\varphi(t)\rangle$ is given by:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\varphi(t)\rangle = [H(t) + \hbar\omega]|\varphi(t)\rangle$$

How does the norm of $|\varphi(t)\rangle$ vary with time?

d. Assuming that $|\psi(0)\rangle$ is an eigenvector of a with the eigenvalue $\alpha(0)$, show that $|\psi(t)\rangle$ is also an eigenvector of a, and calculate its eigenvalue.

Find at time t the mean value of the unperturbed Hamiltonian $H_0 = H(t) - W(t)$ as a function of $\alpha(0)$. Give the root mean square deviations ΔX, ΔP and ΔH_0; how do they vary with time?

e. Assume that at $t = 0$, the oscillator is in the ground state $|\varphi_0\rangle$. The electric field acts between times 0 and T and then falls to zero. When $t > T$, what is the evolution of the mean values $\langle X\rangle(t)$ and $\langle P\rangle(t)$? Application: assume that between 0 and T, the field $\mathscr{E}(t)$ is given by $\mathscr{E}(t) = \mathscr{E}_0\cos(\omega' t)$; discuss the phenomena observed (resonance) in terms of $\Delta\omega = \omega' - \omega$. If, at $t > T$, the energy is measured, what results can be found, and with what probabilities?

7. Consider a one-dimensional harmonic oscillator of Hamiltonian H and stationary states $|\varphi_n\rangle$:

$$H|\varphi_n\rangle = (n + 1/2)\hbar\omega|\varphi_n\rangle$$

The operator $U(k)$ is defined by:

$$U(k) = \mathrm{e}^{ikX}$$

where k is real.

a. Is $U(k)$ unitary? Show that, for all n, its matrix elements satisfy the relation:

$$\sum_{n'} |\langle\varphi_n|U(k)|\varphi_{n'}\rangle|^2 = 1$$

b. Express $U(k)$ in terms of the operators a and $a^\dagger$. Use Glauber's formula [formula (63) of complement B_{II}] to put $U(k)$ in the form of a product of exponential operators.

c. Establish the relations:

$$\mathrm{e}^{\lambda a}|\varphi_0\rangle = |\varphi_0\rangle$$

$$\langle\varphi_n|\mathrm{e}^{\lambda a^\dagger}|\varphi_0\rangle = \frac{\lambda^n}{\sqrt{n!}}$$

where λ is an arbitrary complex parameter.

d. Find the expression, in terms of $E_k = \hbar^2 k^2/2m$ and $E_\omega = \hbar\omega$, for the matrix element:

$$\langle\varphi_0|U(k)|\varphi_n\rangle$$

What happens when k approaches zero? Could this result have been predicted directly?

8. The evolution operator $U(t,0)$ of a one-dimensional harmonic oscillator is written:

$$U(t,0) = \mathrm{e}^{-iHt/\hbar}$$

with:

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right)$$

a. Consider the operators:

$$\tilde{a}(t) = U^\dagger(t,0)\, a\, U(t,0)$$

$$\tilde{a}^\dagger(t) = U^\dagger(t,0)\, a^\dagger\, U(t,0)$$

By calculating their action on the eigenkets $|\varphi_n\rangle$ of H, find the expression for $\tilde{a}(t)$ and $\tilde{a}^\dagger(t)$ in terms of a and $a^\dagger$.

b. Calculate the operators $\tilde{X}(t)$ and $\tilde{P}(t)$ obtained from X and P by the unitary transformation:

$$\tilde{X}(t) = U^\dagger(t,0)\, X\, U(t,0)$$

$$\tilde{P}(t) = U^\dagger(t,0)\, P\, U(t,0)$$

How can the relations so obtained be interpreted?

c. Show that $U^\dagger\left(\frac{\pi}{2\omega}, 0\right)|x\rangle$ is an eigenvector of P and specify its eigenvalue. Similarly, establish that $U^\dagger\left(\frac{\pi}{2\omega}, 0\right)|p\rangle$ is an eigenvector of X.

d. At $t = 0$, the wave function of the oscillator is $\psi(x,0)$. How can one obtain from $\psi(x,0)$ the wave function of the oscillator at all subsequent times $t_q = q\pi/2\omega$ (where q is a positive integer)?

e. Choose for $\psi(x,0)$ the wave function $\varphi_n(x)$ associated with a stationary state. From the preceding question derive the relation which must exist between $\varphi_n(x)$ and its Fourier transform $\bar{\varphi}_n(p)$.

f. Describe qualitatively the evolution of the wave function in the following cases:

(*i*) $\psi(x,0) = \mathrm{e}^{ikx}$ where k, real, is given.

(*ii*) $\psi(x,0) = \mathrm{e}^{-\rho x}$ where ρ is real and positive.

(*iii*)

$$\psi(x,0) = \begin{cases} = \dfrac{1}{\sqrt{a}} \text{ if } \quad -\dfrac{a}{2} \leq x \leq \dfrac{a}{2} \\ = 0 \qquad \text{everywhere else} \end{cases}$$

(*iv*) $\psi(x,0) = \mathrm{e}^{-\rho^2 x^2}$ where ρ is real.

Chapter VI

General properties of angular momentum in quantum mechanics

A Introduction: the importance of angular momentum 667
B Commutation relations characteristic of angular momentum669
B-1 Orbital angular momentum 669
B-2 Generalization: definition of an angular momentum 670
B-3 Statement of the problem 671
C General theory of angular momentum 671
C-1 Definitions and notation 671
C-2 Eigenvalues of $\mathbf{J}^2$ and J_z 673
C-3 "Standard" $\{|k, j, m\rangle\}$ representations 677
D Application to orbital angular momentum 685
D-1 Eigenvalues and eigenfunctions of $\mathbf{L}^2$ and L_z 685
D-2 Physical considerations 693

A. Introduction: the importance of angular momentum

The present chapter is the first in a series of four Chapters (VI, VII, IX and X) devoted to the study of angular momenta in quantum mechanics. This is an extremely important problem, and the results we are going to establish are used in many domains of physics: the classification of atomic, molecular and nuclear spectra, the spin of elementary particles, magnetism, etc...

We already know the important role played by angular momentum in classical mechanics; the total angular momentum of an isolated physical system is a *constant of*

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

the motion. Furthermore, this is also true in certain cases in which the system is not isolated. For example, if a point particle P, of mass m, is moving in a central potential (one which depends only on the distance between P and a fixed point O), the force to which P is subjected is always directed towards O. Its moment with respect to O is consequently zero, and the angular momentum theorem implies that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{L}} = \mathbf{0} \tag{A-1}$$

where $\boldsymbol{\mathcal{L}}$ is the angular momentum of P with respect to O. This fact has important consequences: the motion of the particle P is limited to a fixed plane (the plane passing through O and perpendicular to the angular momentum $\boldsymbol{\mathcal{L}}$); moreover, this motion obeys the law of constant areas (Kepler's second law).

All these properties have their equivalents in quantum mechanics. With the angular momentum $\boldsymbol{\mathcal{L}}$ of a classical system is associated an observable $\mathbf{L}$, actually a set of three observables, L_x, L_y and L_z, which correspond to the three components of $\boldsymbol{\mathcal{L}}$ in a Cartesian frame. If the physical system under study is a point moving in a central potential, we shall see in Chapter VII that L_x, L_y and L_z are constants of the motion in a quantum mechanical sense, that is, they commute with the Hamiltonian H describing the particle in the central potential $V(r)$. This important property considerably simplifies the determination and classification of eigenstates of H.

Also, we described the Stern-Gerlach experiment in Chapter IV and this revealed the *quantization of angular momentum*: the component, along a fixed axis, of the intrinsic angular momentum of an atom can take on only certain discrete values. We shall see that all angular momenta are quantized in this way. This enables us to understand atomic magnetism, the Zeeman effect, etc... Furthermore, to analyze all these phenomena, we must introduce *typically quantum mechanical angular momenta, which have no classical equivalents* (intrinsic angular momenta of elementary particles, Chap. IX).

From now on, we shall denote by *orbital angular momentum* any angular momentum that has a classical equivalent (and by $\mathbf{L}$, the corresponding observables), and by *spin angular momentum* any intrinsic angular momentum of an elementary particle (for which we reserve the letter $\mathbf{S}$). In a complex system, such a nucleus, an atom, or a molecule, the orbital angular momenta $\mathbf{L}_i$ of the various elementary particles which constitute the system (electrons, protons, neutrons, ...) combine with each other and with the spin angular momenta $\mathbf{S}_i$ of these same particles to form the *total angular momentum* $\mathbf{J}$ of the system. The way in which angular momenta are combined in quantum mechanics (coupling of angular momenta) will be studied in Chapter X. Finally, let us add that $\mathbf{J}$ will also be used to denote an arbitrary angular momentum when it is not necessary to specify whether we are dealing with an orbital angular momentum, a spin, or a combination of several angular momenta.

Before beginning the study of the physical problems just mentioned (energy levels of a particle in a central potential, spin, the Zeeman effect, addition of angular momenta,...), we shall establish, in this chapter, the *general* quantum mechanical properties associated with all angular momenta, whatever their nature.

These properties follow from commutation relations satisfied by the three observables J_x, J_y and J_z, the components of an arbitrary angular momentum $\mathbf{J}$. The origin of these commutation relations is discussed in § B: for an orbital angular momentum, they are simply consequences of the quantization rules (§ B-5 of Chapter III) and the canonical commutation relations [formulas (E-30) of Chapter II]; for spin angular momenta,

which have no classical equivalents, they actually serve as definitions of the corresponding observables[1]. In § C, we study the consequences of these commutation relations which are characteristic of angular momenta. In particular, we discuss space quantization, that is, the fact that any component of an angular momentum possesses a discrete spectrum. Finally, the general results so obtained are applied, in § D, to the orbital angular momentum of a particle.

B. Commutation relations characteristic of angular momentum

B-1. Orbital angular momentum

To obtain the observables L_x, L_y, L_z associated in quantum mechanics with the three components of the angular momentum $\boldsymbol{\mathcal{L}}$ of a spinless particle, we simply apply the quantization rules stated in § B-5 of Chapter III. Consider for instance the component $\mathcal{L}_x$ of the classical angular momentum:

$$\mathcal{L}_x = yp_z - zp_y \tag{B-1}$$

We associate with the position variables y and z, the observables Y and Z, and with the momentum variables p_y and p_z, the observables P_y and P_z. Although formula (B-1) involves products of two classical variables, no precaution needs to be taken in replacing them by the corresponding observables, since Y and P_z commute, as do Z and P_y [see the canonical commutation relations (E-30) of Chapter II]. We therefore do not need to symmetrize expression (B-1) in order to obtain the operator L_x:

$$L_x = YP_z - ZP_y \tag{B-2}$$

For the same reason (Y and P_z commute, as do Z and P_y), the operator thus obtained is Hermitian.

In the same way, we find the operators L_x and L_y corresponding to the components $\mathcal{L}_y$ and $\mathcal{L}_z$ of the classical angular momentum. This allows us to write:

$$\mathbf{L} = \mathbf{R} \times \mathbf{P} \tag{B-3}$$

Since we know the canonical commutation relations of the position $\mathbf{R}$ and momentum $\mathbf{P}$ observables, we can easily calculate the commutators of the operators L_x, L_y and L_z.

For example, let us evaluate $[L_x, L_y]$:

$$\begin{aligned}[L_x, L_y] &= [YP_z - ZP_y, ZP_x - XP_z] \\ &= [YP_z, ZP_x] + [ZP_y, XP_z]\end{aligned} \tag{B-4}$$

since YP_z commutes with XP_z, and ZP_y, with ZP_x. We then have:

$$\begin{aligned}[L_x, L_y] &= Y[P_z, Z]P_x + X[Z, P_z]P_y \\ &= -i\hbar YP_x + i\hbar XP_y \\ &= i\hbar L_z\end{aligned} \tag{B-5}$$

[1] The fundamental origin of these commutation relations is purely geometrical. We shall discuss this point in detail in Complement B_{VI}, in which we demonstrate the intimate relation between the angular momentum of a system with respect to a point O and the *geometrical rotations* of this system about O.

Analogous calculations yield the other two commutators, and we obtain, finally:

$$\begin{aligned}[L_x, L_y] &= i\hbar L_z \\ [L_y, L_z] &= i\hbar L_x \\ [L_z, L_x] &= i\hbar L_y\end{aligned} \tag{B-6}$$

Thus we have established the commutation relations for the components of the angular momentum of a spinless particle.

This result can be generalized to a system of N spinless particles. The total angular momentum of such a system is, in quantum mechanics:

$$\mathbf{L} = \sum_{i=1}^{N} \mathbf{L}_i \tag{B-7}$$

with:

$$\mathbf{L}_i = \mathbf{R}_i \times \mathbf{P}_i \tag{B-8}$$

Now, each of the individual angular momenta $\mathbf{L}_i$ satisfies the commutation relations (B-6) and commutes with $\mathbf{L}_j$ when j is not equal to i (operators acting in state spaces of different particles). Thus we see that relations (B-6) remain valid for the total angular momentum $\mathbf{L}$.

B-2. Generalization: definition of an angular momentum

The three operators associated with the components of an arbitrary classical angular momentum therefore satisfy the commutation relations (B-6). It can be shown, moreover (*cf.* Complement $\mathrm{B_{VI}}$), that the origin of these relations lies in the geometric properties of rotations in three-dimensional space. This is why we shall adopt a more general point of view and define an angular momentum $\mathbf{J}$ as any set of three observables J_x, J_y, J_z that satisfies:

$$\boxed{\begin{aligned}[J_x, J_y] &= i\hbar J_z \\ [J_y, J_z] &= i\hbar J_x \\ [J_z, J_x] &= i\hbar J_y\end{aligned}} \tag{B-9}$$

We then introduce the operator:

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2 \tag{B-10}$$

the (scalar) square of the angular momentum $\mathbf{J}$. This operator is Hermitian, since J_x, J_y and J_z are Hermitian. We shall assume that it is an observable. Let us show that $\mathbf{J}^2$ commutes with the three components of $\mathbf{J}$:

$$\boxed{[\mathbf{J}^2, \mathbf{J}] = \mathbf{0}} \tag{B-11}$$

We perform the calculation for J_x, for example:

$$\begin{aligned}\left[\mathbf{J}^2, J_x\right] &= \left[J_x^2 + J_y^2 + J_z^2, J_x\right] \\ &= \left[J_y^2, J_x\right] + \left[J_z^2, J_x\right]\end{aligned} \tag{B-12}$$

since J_x obviously commutes with itself and, therefore, with its square. The other two commutators can be obtained from (B-9):

$$\begin{aligned}\left[J_y^2, J_x\right] &= J_y\left[J_y, J_x\right] + \left[J_y, J_x\right] J_y \\ &= -i\hbar J_y J_z - i\hbar J_z J_y \qquad \text{(B-13a)}\\ \left[J_z^2, J_x\right] &= J_z\left[J_z, J_x\right] + \left[J_z, J_x\right] J_z \\ &= i\hbar J_z J_y + i\hbar J_y J_z \qquad \text{(B-13b)}\end{aligned}$$

The sum of these two commutators, which enters into (B-12), is indeed zero.

Angular momentum theory in quantum mechanics is founded entirely on the commutation relations (B-9). Note that these relations imply that it is impossible to measure simultaneously the three components of an angular momentum; however, $\mathbf{J}^2$ and any component of $\mathbf{J}$ are compatible.

B-3. Statement of the problem

Let us return to the example of a spinless particle in a central potential, mentioned in the introduction. We shall see in Chapter VII that, in this case, the three components of the angular momentum $\mathbf{L}$ of the particle commute with the Hamiltonian H; thus, this is also true for the operator $\mathbf{L}^2$. We then have at our disposal four constants of the motion: $\mathbf{L}^2$, L_x, L_y, L_z. But these four operators do not all commute; to form a complete set of commuting observables with H, we must pick only $\mathbf{L}^2$ and one of the three other operators, L_z for example. For a particle subjected to a central potential, we can then look for eigenstates of the Hamiltonian H which are also eigenvectors of $\mathbf{L}^2$ and L_z, without restricting the generality of the problem. However, it is impossible to obtain a basis of the state space composed of eigenvectors common to the three components of $\mathbf{L}$, as these three observables do not commute.

The situation is the same in the general case: since the components of an arbitrary angular momentum $\mathbf{J}$ do not commute, they are not simultaneously diagonalizable. We shall therefore seek the system of eigenvectors common to $\mathbf{J}^2$ and J_z, observables corresponding to the square of the absolute value of the angular momentum and to its component along the Oz axis.

C. General theory of angular momentum

In this section, we shall determine the spectrum of $\mathbf{J}^2$ and J_z for the general case and study their common eigenvectors. The reasoning will be analogous to the one used in Chapter V for the harmonic oscillator.

C-1. Definitions and notation

C-1-a. The J_+ and J_- operators

Instead of using the components J_x and J_y of the angular momentum $\mathbf{J}$, it is more convenient to introduce the following linear combinations:

$$\boxed{\begin{aligned}J_+ &= J_x + iJ_y \\ J_- &= J_x - iJ_y\end{aligned}} \qquad \text{(C-1)}$$

Like the operators a and $a^\dagger$ of the harmonic oscillator, J_+ and J_- are not Hermitian: they are adjoints of each other.

In the rest of this section, we shall use only the operators J_+, J_-, J_z and $\mathbf{J}^2$. It is straightforward, using (B-9) and (B-11), to show that these operators satisfy the commutation relations:

$$[J_z, J_+] = \hbar J_+ \tag{C-2}$$
$$[J_z, J_-] = -\hbar J_- \tag{C-3}$$
$$[J_+, J_-] = 2\hbar J_z \tag{C-4}$$
$$[\mathbf{J}^2, J_+] = [\mathbf{J}^2, J_-] = [\mathbf{J}^2, J_z] = 0 \tag{C-5}$$

Let us calculate the products J_+J_- and J_-J_+. We find:

$$\begin{aligned} J_+J_- &= (J_x + iJ_y)(J_x - iJ_y) \\ &= J_x^2 + J_y^2 - i[J_x, J_y] \\ &= J_x^2 + J_y^2 + \hbar J_z \end{aligned} \tag{C-6a}$$

$$\begin{aligned} J_-J_+ &= (J_x - iJ_y)(J_x + iJ_y) \\ &= J_x^2 + J_y^2 + i[J_x, J_y] \\ &= J_x^2 + J_y^2 - \hbar J_z \end{aligned} \tag{C-6b}$$

Using definition (B-10) of the operator $\mathbf{J}^2$, we can write these expressions in the form:

$$J_+J_- = \mathbf{J}^2 - J_z^2 + \hbar J_z \tag{C-7a}$$
$$J_-J_+ = \mathbf{J}^2 - J_z^2 - \hbar J_z \tag{C-7b}$$

Adding relations (C-7)21, we obtain:

$$\mathbf{J}^2 = \frac{1}{2}(J_+J_- + J_-J_+) + J_z^2 \tag{C-8}$$

C-1-b. Notation for the eigenvalues of $\mathbf{J}^2$ and J_z

According to (B-10), $\mathbf{J}^2$ is the sum of the squares of three Hermitian operators. Consequently, for any ket $|\psi\rangle$, the matrix element $\langle\psi|\mathbf{J}^2|\psi\rangle$ is positive or zero:

$$\begin{aligned} \langle\psi|\mathbf{J}^2|\psi\rangle &= \langle\psi|J_x^2|\psi\rangle + \langle\psi|J_y^2|\psi\rangle + \langle\psi|J_z^2|\psi\rangle \\ &= ||J_x|\psi\rangle||^2 + ||J_y|\psi\rangle||^2 + ||J_z|\psi\rangle||^2 \geq 0 \end{aligned} \tag{C-9}$$

Note that this could have been expected, since $\mathbf{J}^2$ corresponds to the square of the absolute value of the angular momentum $\mathbf{J}$. From this we see, in particular, that *all the eigenvalues of* $\mathbf{J}^2$ *are positive or zero*, since if $|\psi\rangle$ is an eigenvector of $\mathbf{J}^2$, $\langle\psi|\mathbf{J}^2|\psi\rangle$ is the product of the corresponding eigenvalue and the square of the norm of $|\psi\rangle$, which is always positive.

We shall write the eigenvalues of $\mathbf{J}^2$ in the form $j(j+1)\hbar^2$, with the *convention*:

$$j \geq 0 \tag{C-10}$$

This notation is intended to simplify the arguments which follow; it does not influence the result. Since $\mathbf{J}$ has the dimensions of $\hbar$, an eigenvalue of $\mathbf{J}^2$ is necessarily of the form $\lambda\hbar^2$, where λ is a real dimensionless number. We have just seen that λ must be positive or zero; it can then be shown that the second-degree equation in j:

$$j(j+1) = \lambda \tag{C-11}$$

always has one and only one positive or zero root. Therefore, if we use relation (C-10), the specification of λ determines j uniquely; any eigenvalue of $\mathbf{J}^2$ can thus be written $j(j+1)\hbar^2$, with j positive or zero.

As for the eigenvalues of J_z, which have the same dimensions as $\hbar$, they are traditionally written $m\hbar$, where m is a dimensionless number.

C-1-c. Eigenvalue equations for $\mathbf{J}^2$ and J_z

We shall label the eigenvectors common to $\mathbf{J}^2$ and J_z by the indices j and m which characterize the associated eigenvalues. However, $\mathbf{J}^2$ and J_z do not in general constitute a C.S.C.O. (see, for example, § A of Chapter VII), and it is necessary to introduce a third index in order to distinguish between the different eigenvectors corresponding to the same eigenvalues $j(j+1)\hbar^2$ and $m\hbar$ of $\mathbf{J}^2$ and J_z (this point will be expanded in § C-3-a below). We shall call this index k (which does not necessarily imply that it is always a discrete index).

We shall therefore try to solve the simultaneous eigenvalue equations:

$$\begin{aligned} \mathbf{J}^2|k,j,m\rangle &= j(j+1)\hbar^2|k,j,m\rangle \\ J_z|k,j,m\rangle &= m\hbar|k,j,m\rangle \end{aligned} \tag{C-12}$$

C-2. Eigenvalues of $\mathbf{J}^2$ and J_z

As in § B-2 of Chapter V, we shall begin by proving three lemmas which will then enable us to determine the spectrum of $\mathbf{J}^2$ and J_z.

C-2-a. Lemmas

α. Lemma I (Properties of the eigenvalues of $\mathbf{J}^2$ and J_z)

If $j(j+1)\hbar^2$ and $m\hbar$ are the eigenvalues of $\mathbf{J}^2$ and J_z associated with the same eigenvector $|k,j,m\rangle$, then j and m satisfy the inequality:

$$-j \leq m \leq j \tag{C-13}$$

To prove this, consider the vectors $J_+|k,j,m\rangle$ and $J_-|k,j,m\rangle$, and note that the square of their norms is positive or zero:

$$||J_+|k,j,m\rangle||^2 = \langle k,j,m|J_-J_+|k,j,m\rangle \geq 0 \tag{C-14a}$$

$$||J_-|k,j,m\rangle||^2 = \langle k,j,m|J_+J_-|k,j,m\rangle \geq 0 \tag{C-14b}$$

To calculate the left-hand sides of these inequalities, we can use formulas (C-7). We find (if we assume $|k, j, m\rangle$ to be normalized):

$$\begin{aligned}\langle k, j, m|J_- J_+|k, j, m\rangle &= \langle k, j, m|(\mathbf{J}^2 - J_z^2 - \hbar J_z)|k, j, m\rangle \\ &= j(j+1)\hbar^2 - m^2\hbar^2 - m\hbar^2 \qquad \text{(C-15a)}\\ \langle k, j, m|J_+ J_-|k, j, m\rangle &= \langle k, j, m|(\mathbf{J}^2 - J_z^2 + \hbar J_z)|k, j, m\rangle \\ &= j(j+1)\hbar^2 - m^2\hbar^2 + m\hbar^2 \qquad \text{(C-15b)}\end{aligned}$$

Substituting these expressions into inequalities (C-14), we obtain:

$$j(j+1) - m(m+1) = (j-m)(j+m+1) \geq 0 \qquad \text{(C-16a)}$$
$$j(j+1) - m(m-1) = (j-m+1)(j+m) \geq 0 \qquad \text{(C-16b)}$$

that is:

$$-(j+1) \leq m \leq j \qquad \text{(C-17a)}$$
$$-j \leq m \leq j+1 \qquad \text{(C-17b)}$$

These two conditions are satisfied simultaneously only if m satisfies inequality (C-13).

β. Lemma II (Properties of the vector $J_-|k, j, m\rangle$)

Let $|k, j, m\rangle$ be an eigenvector of $\mathbf{J}^2$ and J_z with the eigenvalues $j(j+1)\hbar^2$ and $m\hbar$.

(*i*) If $m = -j$, $J_-|k, j, -j\rangle = 0$.

(*ii*) If $m > -j$, $J_-|k, j, m\rangle$ is a non-null eigenvector of $\mathbf{J}^2$ and J_z with the eigenvalues $j(j+1)\hbar^2$ and $(m-1)\hbar$.

(*i*) According to (C-15b), the square of the norm of the ket $J_-|k, j, m\rangle$ is equal to $\hbar^2[j(j+1) - m(m-1)]$ and therefore goes to zero for $m = -j$. Since the norm of a vector goes to zero if and only if this vector is a null vector, we conclude that all vectors $J_-|k, j, -j\rangle$ are null:

$$m = -j \Longrightarrow J_-|k, j, -j\rangle = 0 \qquad \text{(C-18)}$$

It is easy to establish the converse of (C-18):

$$J_-|k, j, m\rangle = 0 \Longrightarrow m = -j \qquad \text{(C-19)}$$

Letting J_+ act on both sides of the equation appearing in (C-19), and using (C-7a), we obtain:

$$\hbar^2[j(j+1) - m^2 + m]|k, j, m\rangle = \hbar^2(j+m)(j-m+1)|k, j, m\rangle = 0 \qquad \text{(C-20)}$$

Using (C-13), (C-20) has only one solution, $m = -j$.

(*ii*) Now assume m to be greater than $-j$. According to (C-15b), $J_-|k, j, m\rangle$ is then a non-null vector since the square of its norm is different from zero.

Let us show that it is an eigenvector of $\mathbf{J}^2$ and J_z. The operators J_- and $\mathbf{J}^2$ commute; consequently:

$$[\mathbf{J}^2, J_-]|k,j,m\rangle = 0 \tag{C-21}$$

which can be written:

$$\begin{aligned}\mathbf{J}^2 J_-|k,j,m\rangle &= J_-\mathbf{J}^2|k,j,m\rangle \\ &= j(j+1)\hbar^2 J_-|k,j,m\rangle\end{aligned} \tag{C-22}$$

This relation expresses the fact that $J_-|k,j,m\rangle$ is an eigenvector of $\mathbf{J}^2$ with the eigenvalue $j(j+1)\hbar^2$.

Moreover, if we apply operator equation (C-3) to $|k,j,m\rangle$:

$$[J_z, J_-]|k,j,m\rangle = -\hbar J_-|k,j,m\rangle \tag{C-23}$$

that is:

$$\begin{aligned}J_z J_-|k,j,m\rangle &= J_- J_z|k,j,m\rangle - \hbar J_-|k,j,m\rangle \\ &= m\hbar\ J_-|k,j,m\rangle - \hbar J_-|k,j,m\rangle \\ &= (m-1)\hbar J_-|k,j,m\rangle\end{aligned} \tag{C-24}$$

$J_-|k,j,m\rangle$ is therefore an eigenvector of J_z with the eigenvalue $(m-1)\hbar$.

γ. *Lemma III (Properties of the vector $J_+|k,j,m\rangle$)*

Let $|k,j,m\rangle$ be an eigenvector of $\mathbf{J}^2$ and J_z with the eigenvalues $j(j+1)\hbar^2$ and $m\hbar$.

(i) If $m = j$, $J_+|k,j,j\rangle = 0$.

(ii) If $m < j$, $J_+|k,j,m\rangle$ is a non-null eigenvector of $\mathbf{J}^2$ and J_z with the eigenvalues $j(j+1)\hbar^2$ and $(m+1)\hbar$.

(i) The argument is similar to that of (§ C-2-a-β). According to (C-14a), the square of the norm of $J_+|k,j,m\rangle$ is zero if $m = j$. Therefore:

$$m = j \Longrightarrow J_+|k,j,j\rangle = 0 \tag{C-25}$$

The converse can be proved in the same way:

$$J_+|k,j,m\rangle = 0 \Leftrightarrow m = j \tag{C-26}$$

(ii) If $m < j$, an argument analogous to that of § C-2-a-β-ii yields, using formulas (C-5) and (C-2):

$$\mathbf{J}^2 J_+|k,j,m\rangle = j(j+1)\hbar^2 J_+|k,j,m\rangle \tag{C-27}$$

$$J_z J_+|k,j,m\rangle = (m+1)\hbar J_+|k,j,m\rangle \tag{C-28}$$

C-2-b. Determination of the spectrum of $\mathbf{J}^2$ and J_z

We shall now show that the three lemmas above enable us to determine the possible values of j and m.

Let $|k, j, m\rangle$ be a non-null eigenvector of $\mathbf{J}^2$ and J_z with the eigenvalues $j(j+1)\hbar^2$ and $m\hbar$. According to lemma *I*, $-j \leq m \leq j$. It is therefore certain that a positive or zero integer p exists such that:

$$-j \leq m - p < -j + 1 \tag{C-29}$$

Now consider the series of vectors:

$$|k, j, m\rangle, \quad J_-|k, j, m\rangle, \quad ..., \ (J_-)^p|k, j, m\rangle \tag{C-30}$$

According to lemma *II*, each of the vectors $(J_-)^n|k, j, m\rangle$ of this series $(n = 0, 1, \ldots, p)$ is a non-null eigenvector of $\mathbf{J}^2$ and J_z with the eigenvalues $j(j+1)\hbar^2$ and $(m-n)\hbar$.

The proof is by iteration. By hypothesis, $|k, j, m\rangle$ is non-null and corresponds to the eigenvalues $j(j+1)\hbar^2$ and $m\hbar$. $(J_-)^n|k, j, m\rangle$ is obtained by the action of J_- on $(J_-)^{n-1}|k, j, m\rangle$, which is an eigenvector of $\mathbf{J}^2$ and J_z with the eigenvalues $j(j+1)\hbar^2$ and $(m-n+1)\hbar$. The latter eigenvalue is necessarily greater than $-j$ since, according to (C-29):

$$m - n + 1 \geq m - p + 1 \geq -j + 1 \tag{C-31}$$

It follows, according to point (ii) of lemma *II*, that $(J_-)^n|k, j, m\rangle$ is a non-null eigenvector of $\mathbf{J}^2$ and J_z, the corresponding eigenvalues being $j(j+1)\hbar^2$ and $(m-n)\hbar$.

Now let J_- act on $(J_-)^p\ |k, j, m\rangle$. Let us first assume that the eigenvalue $(m-p)\hbar$ of J_z associated with $(J_-)^p|k, j, m\rangle$ is greater than $-j\hbar$, that is, that:

$$m - p > -j \tag{C-32}$$

By point (ii) of lemma *II*, $J_-(J_-)^p|k, j, m\rangle$ is then non-null and corresponds to the eigenvalues $j(j+1)\hbar^2$ and $(m-p-1)\hbar$. This is in contradiction with lemma *I* since, according to (C-29):

$$m - p - 1 < -j \tag{C-33}$$

We must therefore have $m-p$ equal to $-j$. In this case, $(J_-)^p|k, j, m\rangle$ corresponds to the eigenvalue $-j$ of J_z, and, according to point (i) of lemma *II*, $J_-(J_-)^p|k, j, m\rangle$ is zero. The vector series (C-30) obtained by the repeated action of J_- on $|k, j, m\rangle$ is therefore limited and the contradiction with lemma *I* is removed.

We have now shown that there exists a positive or zero integer p such that:

$$m - p = -j \tag{C-34}$$

An analogous argument, based on lemma *III*, would show that there exists a positive or zero integer q such that:

$$m + q = j \tag{C-35}$$

since the vector series:

$$|k, j, m\rangle, \ J_+|k, j, m\rangle, ..., (J_+)^q|k, j, m\rangle \tag{C-36}$$

must be limited if there is to be no contradiction with lemma *I*.

Combining (C-34) and (C-35), we obtain:

$$p + q = 2j \tag{C-37}$$

j is therefore equal to a positive or zero integer divided by 2. It follows that j is necessarily integral or half-integral[2]. Furthermore, if there exists a non-null vector $|k, j, m\rangle$, all the vectors of series (C-30) and (C-36) are also non-null and eigenvectors of $\mathbf{J}^2$ with the eigenvalue $j(j+1)\hbar^2$, as well as of J_z with the eigenvalues:

$$-j\hbar,\ (-j+1)\hbar,\ (-j+2)\hbar,\ ...,\ (j-2)\hbar,\ (j-1)\hbar,\ j\hbar \tag{C-38}$$

We summarize the results obtained above as follows:

Let $\mathbf{J}$ be an arbitrary angular momentum, obeying the commutation relations (B-9). If $j(j+1)\hbar^2$ and $m\hbar$ denote the eigenvalues of $\mathbf{J}^2$ and J_z, then:

- the only values possible for j are positive integers or half-integers or zero, that is: 0, 1/2, 1, 3/2, 2, ... (these values are the only ones possible, but they are not all necessarily realized for all angular momenta).
- for a fixed value of j, the only values possible for m are the $(2j+1)$ numbers: $-j, -j+1, ..., j-1, j$; m is therefore integral if j is integral and half-integral if j is half-integral. All these values of m are realized once one of them is.

C-3. "Standard" $\{|k, j, m\rangle\}$ representations

We shall now study the eigenvectors common to $\mathbf{J}^2$ and J_z, which form a basis of the state space since $\mathbf{J}^2$ and J_z are, by hypothesis, observables.

C-3-a. The basis states

Consider an angular momentum $\mathbf{J}$ acting in a state space $\mathcal{E}$. We shall show how to construct an orthonormal basis in $\mathcal{E}$ composed of eigenvectors common to $\mathbf{J}^2$ and J_z.

Take a pair of eigenvalues, $j(j+1)\hbar^2$ and $m\hbar$, that are actually found in the case we are considering. The set of eigenvectors associated with this pair of eigenvalues forms a vector subspace of $\mathcal{E}$ which we shall denote by $\mathcal{E}(j, m)$; the dimension $g(j, m)$ of this subspace may well be greater than 1, since $\mathbf{J}^2$ and J_z do not generally constitute a C.S.C.O. We choose in $\mathcal{E}(j, m)$ an arbitrary orthonormal basis, $\{|k, j, m\rangle; k = 1, 2, ..., g(j, m)\}$.

If m is not equal to j, there must exist another subspace $\mathcal{E}(j, m+1)$ in $\mathcal{E}$ composed of eigenvectors of $\mathbf{J}^2$ and J_z associated with the eigenvalues $j(j+1)\hbar^2$ and $(m+1)\hbar$. Similarly, if m is not equal to $-j$, there exists a subspace $\mathcal{E}(j, m-1)$. In the case where m is not equal to j or $-j$, we shall construct orthonormal bases in $\mathcal{E}(j, m+1)$ and in $\mathcal{E}(j, m-1)$, starting with the one chosen in $\mathcal{E}(j, m)$.

First, let us show that, if k_1 is not equal to k_2, $J_+|k_1, j, m\rangle$ and $J_+|k_2, j, m\rangle$ are orthogonal, as are $J_-|k_1, j, m\rangle$ and $J_-|k_2, j, m\rangle$. We can find the scalar product

[2] A number is said to be "half-integral" if it is equal to an odd number divided by 2.

of $J_\pm|k_1, j, m\rangle$ and $J_\pm|k_2, j, m\rangle$ by using formulas (C-7):

$$\begin{aligned}\langle k_2, j, m|J_\mp J_\pm|k_1, j, m\rangle &= \langle k_2, j, m|(\mathbf{J}^2 - J_z^2 \mp \hbar J_z)|k_1, j, m\rangle \\ &= [j(j+1) - m(m \pm 1)]\hbar^2\langle k_2, j, m|k_1, j, m\rangle \end{aligned} \tag{C-39}$$

These scalar products are therefore zero if $k_1 \neq k_2$ since the basis of $\mathcal{E}(j, m)$ is orthonormal; if $k_1 = k_2$, the square of the norm of $J_\pm|k_1, j, m\rangle$ is equal to:

$$[j(j+1) - m(m \pm 1)]\hbar^2$$

Now let us consider the set of the $g(j, m)$ vectors defined by:

$$|k, j, m+1\rangle = \frac{1}{\hbar\sqrt{j(j+1) - m(m+1)}} J_+|k, j, m\rangle \tag{C-40}$$

Because of what we have just shown, these vectors are orthonormal. We shall show that they constitute a basis in $\mathcal{E}(j, m+1)$. Assume that there exists, in $\mathcal{E}(j, m+1)$, a vector $|\alpha, j, m+1\rangle$ orthogonal to all the $|k, j, m+1\rangle$ obtained from (C-40). The vector $J_-|\alpha, j, m+1\rangle$ would not be null since $(m+1)$ cannot be equal to $-j$; it would belong to $\mathcal{E}(j, m)$ and would be orthogonal to all vectors $J_-|k, j, m+1\rangle$. Now, according to relation (C-40), the ket $J_-|k, j, m+1\rangle$ is proportional to $J_-J_+|k, j, m\rangle$, that is, to $|k, j, m\rangle$ [formula (C-7b)]. Therefore, $J_-|\alpha, j, m+1\rangle$ would be a non-null vector of $\mathcal{E}(j, m)$ which would be orthogonal to all vectors of the $|k, j, m\rangle$ basis. But this is impossible. Consequently, the set of vectors (C-40) constitutes a basis in $\mathcal{E}(j, m+1)$.

It can be shown, using an analogous argument, that the vectors $|k, j, m-1\rangle$ defined by:

$$|k, j, m-1\rangle = \frac{1}{\hbar\sqrt{j(j+1) - m(m-1)}} J_-|k, j, m\rangle \tag{C-41}$$

form an orthonormal basis in $\mathcal{E}(j, m-1)$.

We see, in particular, that the dimension of subspaces $\mathcal{E}(j, m+1)$ and $\mathcal{E}(j, m-1)$ is equal to that of $\mathcal{E}(j, m)$. In other words, this dimension is independent[3] of m:

$$g(j, m+1) = g(j, m-1) = g(j, m) = g(j) \tag{C-42}$$

We then proceed as follows. For each value of j actually found in the problem under consideration, we choose one of the subspaces associated with this value of j, for example $\mathcal{E}(j, j)$ corresponding to $m = j$. In this subspace, we choose an arbitrary orthonormal basis, $\{|k, j, j, \rangle\,;\ k = 1, 2, ..., g(j)\}$. Then, using formula (C-41), we construct, by iteration, the basis to which each of the other $2j$ subspaces $\mathcal{E}(j, m)$ will be related: the arrows of table (VI-1) indicate the method used. By treating all the values of j found in the problem in this way, we arrive at what is called a *standard basis* of the state space $\mathcal{E}$. The orthonormalization and closure relations for such a basis are:

$$\langle k, j, m|k', j', m'\rangle = \delta_{kk'}\ \delta_{jj'}\ \delta_{mm'} \tag{C-43a}$$

$$\sum_j \sum_{m=-j}^{+j} \sum_{k=1}^{g(j)} |k, j, m\rangle\langle k, j, m| = 1 \tag{C-43b}$$

[3] If this dimension is infinite, the result must be interpreted as follows: there is a one-to-one correspondence between the basis vectors of two subspaces corresponding to the same value of j.

Comments:

(*i*) The use of formula (C-41) implies a choice of phases: the basis vectors in $\mathcal{E}(j, m-1)$ are chosen to be proportional, with a real and positive coefficient, to the vectors obtained by application of J_- to the basis of $\mathcal{E}(j, m)$.

(*ii*) Formulas (C-40) and (C-41) are compatible, since, if we apply J_+ to both sides of (C-41) and take (C-7a) into account, we find (C-40) [with m replaced by $(m-1)$]. This means that one is not obliged to start, as we did, with the maximum value $m = j$ and (C-41) in order to construct bases of the subspaces $\mathcal{E}(j, m)$ corresponding to a given value of j.

$$
\begin{array}{cc|cccc}
 & & \multicolumn{4}{c}{\overbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{g(j)\text{ different values of }k}} \\
 & & k=1 & , \quad k=2 & , \ \ldots\ , & g(j) \\
\hline
 & \mathcal{E}(j, m=j) & |1,j,j\rangle & |2,j,j\rangle \ldots & & |g(j),j,j\rangle \\
 & \Downarrow J_- & \Downarrow J_- & \Downarrow J_- & & \Downarrow J_- \\
 & \mathcal{E}(j, m=j-1) & |1,j,j-1\rangle & |2,j,j-1\rangle \ldots & & |g(j),j,j-1\rangle \\
 & \Downarrow J_- & \Downarrow J_- & \Downarrow J_- & & \Downarrow J_- \\
(2j+1)\ \text{spaces}\ \mathcal{E}(j,m) & \vdots & \vdots & \vdots & & \vdots \\
 & \mathcal{E}(j, m) & |1,j,m\rangle & |2,j,m\rangle \ldots & & |g(j),j,m\rangle \\
 & \Downarrow J_- & \Downarrow J_- & \Downarrow J_- & & \Downarrow J_- \\
 & \vdots & \vdots & \vdots & & \vdots \\
 & \mathcal{E}(j, -j) & |1,j,-j\rangle & |2,j,-j\rangle \ldots & & |g(j),j,-j\rangle \\
\hline
 & & \underbrace{\mathcal{E}(k=1,j) \quad \mathcal{E}(k=2,j) \ , \ldots\ , \ \mathcal{E}(k=g(j),j)}_{g(j)\text{ spaces }\mathcal{E}(k,j)} & & &
\end{array}
$$

Table VI-1: Schematic representation of the construction of the $(2j+1)$ $g(j)$ *vectors of a "standard basis" associated with a fixed value of* j*. Starting with each of the* $g(j)$ *vectors* $|k,j,j\rangle$ *of the first line, one uses the action of* J_- *to construct the* $(2j+1)$ *vectors of the corresponding column.*

Each subspace $\mathcal{E}(j,m)$ *is spanned by the* $g(j)$ *vectors situated in the same row. Each subspace* $\mathcal{E}(k,j)$ *is spanned by the* $(2j+1)$ *vectors of the corresponding column.*

In most cases, in order to define a standard basis, one uses observables $A, B, ...$ that commute[4] with the three components of J, and form a C.S.C.O. with $\mathbf{J}^2$ and J_z (we shall see a concrete example of this in § A of Chapter VII):

$$[A, \mathbf{J}] = [B, \mathbf{J}] = \ldots = \mathbf{0} \tag{C-44}$$

For the sake of simplicity, we shall assume that only one of these observables A is required to make a C.S.C.O. with $\mathbf{J}^2$ and J_z. Under these conditions, each of the subspaces $\mathcal{E}(j,m)$ defined above is globally invariant under the action of A: if $|\psi_{j,m}\rangle$ is an arbitrary vector of $\mathcal{E}(j,m)$, $A|\psi_{j,m}\rangle$ is still, according to (C-44), an eigenvector of $\mathbf{J}^2$ and J_z:

$$\begin{aligned} &\mathbf{J}^2 A|\psi_{j,m}\rangle = A\,\mathbf{J}^2|\psi_{j,m}\rangle = j(j+1)\hbar^2\ A|\psi_{j,m}\rangle \\ &J_z A|\psi_{j,m}\rangle = A\,J_z|\psi_{j,m}\rangle = m\hbar\ A|\psi_{j,m}\rangle \end{aligned} \tag{C-45}$$

with the same eigenvalues as $|\psi_{j,m}\rangle$. Thus $A|\psi_{j,m}\rangle$ also belongs to $\mathcal{E}(j,m)$. If we then choose a value of j, we can diagonalize A inside the corresponding subspace $\mathcal{E}(j,j)$. We denote by $a_{k,j}$ the various eigenvalues found in this way: the index j indicates in which space $\mathcal{E}(j,j)$ they were found, and the index k (assumed to be discrete, for simplicity) distinguishes between them. A single vector (written $|k,j,j\rangle$) of $\mathcal{E}(j,j)$ is associated with each eigenvalue $a_{k,j}$, since A, $\mathbf{J}^2$ and J_z form, by hypothesis, a C.S.C.O.:

$$A|k,j,j\rangle = a_{k,j}\,|k,j,j\rangle \tag{C-46}$$

The set $\{|k,j,j\rangle;\ j \text{ fixed};\ k = 1,2,...,g(j)\}$ constitutes an orthonormal basis in $\mathcal{E}(j,j)$, from which we construct, using the method described above, a basis in the other subspaces $\mathcal{E}(j,m)$ related to the value of j chosen. By applying this procedure successively for all values of j, we arrive at a "standard" basis, $\{|k,j,m\rangle\}$ of the state space, all of whose vectors are eigenstates, not only of $\mathbf{J}^2$ and J_z, but also of A:

$$A|k,j,m\rangle = a_{k,j}\,|k,j,m\rangle \tag{C-47}$$

This can be shown as follows. If hypothesis (C-44) is satisfied, A commutes with J_-, which means that $J_-|k,j,j\rangle$, that is $|k,j,j-1\rangle$, is an eigenvector of A with the same eigenvalue as $|k,j,j\rangle$:

$$AJ_-|k,j,j\rangle = J_- A|k,j,j\rangle = a_{k,j}\,J_-|k,j,j\rangle \tag{C-48}$$

By repeating this process, it is easy to prove relation (C-47).

Comments:

(*i*) An observable that commutes with $\mathbf{J}^2$ and J_z does not necessarily commute with J_x and J_y (J_z is itself an example). Consequently, it should not be necessary, in order to form a C.S.C.O. with $\mathbf{J}^2$ and J_z, to choose observables that commute with the three components of $\mathbf{J}$ as in (C-44). However, if A did not commute with J_+ and J_- (that is, with J_x and J_y), $J_\pm|k,j,m\rangle$ would not necessarily be an eigenvector of A with the same eigenvalue as $|k,j,m\rangle$.

(*ii*) The spectrum of A is the same in all the subspaces $\mathcal{E}(j,m)$ associated with the same value of j. However, the eigenvalues $a_{k,j}$ generally depend on j (this point will be illustrated by concrete examples in §§ A and C of Chapter VII).

[4] An operator that commutes with the three components of the total angular momentum of a physical system is said to be "scalar" (*cf.* Complement $\mathrm{B_{VI}}$).

C-3-b. The spaces $\mathcal{E}(k, j)$

In the preceding section, we introduced a "standard basis" of the state space by starting with a basis chosen in the subspace $\mathcal{E}(j, m = j)$ and constructing a basis of $\mathcal{E}(j, m = j - 1)$, then one of $\mathcal{E}(j, m = j - 2)$,..., $\mathcal{E}(j, m)$, etc... The state space can be considered to be the direct sum of all the orthogonal subspaces $\mathcal{E}(j, m)$, where m varies by integral jumps from $-j$ to $+j$ and j takes on all the values actually found in the problem. This means that any vector of $\mathcal{E}$ can be written in one and only one way as a sum of vectors, each belonging to a particular subspace $\mathcal{E}(j, m)$.

Nevertheless, the use of the subspaces $\mathcal{E}(j, m)$ presents certain disadvantages. First of all, their dimension $g(j)$ depends on the physical system being considered and is not necessarily known. In addition, the subspaces $\mathcal{E}(j, m)$ are not invariant under the action of $\mathbf{J}$ since, by the very means of construction of the vectors $|k, j, m\rangle$, J_+ and J_- have non-zero matrix elements between vectors of $\mathcal{E}(j, m)$ and those of $\mathcal{E}(j, m \pm 1)$.

We shall therefore introduce other subspaces of $\mathcal{E}$, the spaces $\mathcal{E}(k, j)$. Instead of grouping together the kets $|k, j, m\rangle$ with fixed indices j and m [which span $\mathcal{E}(j, m)$], we shall now group together those for which k and j have given values, and we shall call $\mathcal{E}(k, j)$ the subspace which they span. This amounts to associating, in table (VI-1), the $(2j + 1)$ vectors of one column [instead of the $g(j)$ vectors of one row].

$\mathcal{E}$ can then be seen to be the direct sum of the orthogonal subspaces $\mathcal{E}(k, j)$, which have the simpler properties:

- the dimension of $\mathcal{E}(k, j)$ is $(2j + 1)$, whatever the value of k and whatever the physical system under consideration.

- $\mathcal{E}(k, j)$ is globally invariant under the action of $\mathbf{J}$: any component J_u of $\mathbf{J}$ [or a function $F(\mathbf{J})$ of $\mathbf{J}$], acting on a ket of $\mathcal{E}(k, j)$, yields another ket also belonging[5] to $\mathcal{E}(k, j)$. This result is not difficult to establish, since J_u [or $F(\mathbf{J})$] can always be expressed in terms of J_z, J_+ and J_-. Now, J_z, acting on $|k, j, m\rangle$, yields a ket proportional to $|k, j, m\rangle$; J_+, a ket proportional to $|k, j, m + 1\rangle$; and J_-, a ket proportional to $|k, j, m - 1\rangle$. The existence of the property in question therefore follows from the very means of construction of the "standard basis" $\{|k, j, m\rangle\}$.

C-3-c. Matrices representing the angular momentum operators

Using the subspaces $\mathcal{E}(k, j)$ considerably simplifies the search for the matrix which represents, in a "standard" basis, a component J_u of $\mathbf{J}$ [or an arbitrary function $F(\mathbf{J})$].

The matrix elements of such operators between two basis kets belonging to two different subspaces $\mathcal{E}(k, j)$ are zero. The matrix therefore has the following form:

[5] It can also be shown that $\mathcal{E}(k, j)$ is "irreducible" with respect to $\mathbf{J}$: there exists no subspace of $\mathcal{E}(k, j)$ other than $\mathcal{E}(k, j)$ itself which is globally invariant under the action of the various components of $\mathbf{J}$.

	$\mathcal{E}(k,j)$	$\mathcal{E}(k',j)$	$\mathcal{E}(k',j')$	...
$\mathcal{E}(k,j)$	matrix $(2j+1)\times(2j+1)$	0	0	0
$\mathcal{E}(k',j)$	0	matrix $(2j+1)\times(2j+1)$	0	0
$\mathcal{E}(k',j')$	0	0	matrix $(2j'+1)\times(2j'+1)$	0
$\vdots$	0	0	0	...

(C-49)

All we must then do is calculate the finite-dimensional matrices that represent the operator under consideration inside each of the subspaces $\mathcal{E}(k,j)$.

Another very important simplification arises from the fact that each of these finite submatrices is independent of k and of the physical system under study; it depends only on j and, of course, on the operator which we want to represent. To see this, note that the definition of the $|k,j,m\rangle$ [*cf.* (C-12), (C-40) and (C-41)] implies that:

$$\boxed{\begin{aligned} J_z|k,j,m\rangle &= m\hbar\ |k,j,m\rangle \\ J_+|k,j,m\rangle &= \hbar\sqrt{j(j+1)-m(m+1)}\ |k,j,m+1\rangle \\ J_-|k,j,m\rangle &= \hbar\sqrt{j(j+1)-m(m-1)}\ |k,j,m-1\rangle \end{aligned}} \tag{C-50}$$

that is:

$$\begin{aligned} \langle k,j,m|J_z|k',j',m'\rangle &= m\hbar\ \delta_{kk'}\ \delta_{jj'}\ \delta_{mm'} \\ \langle k,j,m|J_\pm|k',j',m'\rangle &= \hbar\sqrt{j(j+1)-m'(m'\pm1)}\ \delta_{kk'}\ \delta_{jj'}\ \delta_{m,m'\pm1} \end{aligned} \tag{C-51}$$

These relations show that the matrix elements representing the components of **J** depend only on j and m and not on k.

In order to know, in all cases, the matrix associated with an arbitrary component J_u in a standard basis, all we need to do is calculate, once and for all, the "universal" matrices $(J_u)^{(j)}$ which represent J_u inside the subspaces $\mathcal{E}(k,j)$ for all possible values of j ($j = 0,\ 1/2,\ 1,\ 3/2,\ ...$). When we study a particular physical system and its angular momentum **J**, we shall determine the values of j actually found in the problem, as well as the number of subspaces $\mathcal{E}(k,j)$ associated with each of them [that is, its degree of degeneracy $(2j+1)g(j)$]. We know that the matrix representing J_u in this particular case has the "block-diagonal" form (C-49), and we can therefore construct it from the "universal" matrices which we have just defined: for each value of j, we shall have $g(j)$ "blocks" identical to $(J_u)^{(j)}$.

Let us give some examples of $(J_u)^{(j)}$ matrices:

(i) $j = 0$

The subspaces $\mathcal{E}(k, j=0)$ are one-dimensional, since zero is the only possible value for m. The $(J_u)^{(0)}$ matrices therefore reduce to numbers, which, according to (C-51), are zero.

(ii) $j = 1/2$

The subspaces $\mathcal{E}(k, j = 1/2)$ are two-dimensional ($m = 1/2$ or $-1/2$). If we choose the basis vectors in this order ($m = 1/2, m = -1/2$), we find, using (C-51):

$$(J_z)^{(1/2)} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{C-52}$$

and:

$$(J_+)^{(1/2)} = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad (J_-)^{(1/2)} = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{C-53}$$

that is, using (C-1):

$$(J_x)^{(1/2)} = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad (J_y)^{(1/2)} = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{C-54}$$

The matrix representing $\mathbf{J}^2$ is therefore:

$$(\mathbf{J}^2)^{(1/2)} = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{C-55}$$

We thus find the matrices that were introduced without justification in Chapter IV, § A-2.

(iii) $j = 1$

We now have (order of the basis vectors: $m = 1, m = 0, m = -1$):

$$(J_z)^{(1)} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{C-56}$$

$$(J_+)^{(1)} = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \quad (J_-)^{(1)} = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \tag{C-57}$$

and therefore:

$$(J_x)^{(1)} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (J_y)^{(1)} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \tag{C-58}$$

and:

$$(\mathbf{J}^2)^{(1)} = 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{C-59}$$

It can be verified that matrices (C-56) and (C-58) satisfy the commutation relations (B-9).

(*iv*) *j arbitrary*

We use relations (C-51), which, according to (C-1), can also be written:

$$\langle k,j,m|J_x|k',j',m'\rangle = \frac{\hbar}{2}\,\delta_{kk'}\,\delta_{jj'} \times \left[\sqrt{j(j+1)-m'(m'+1)}\;\delta_{m,m'+1}+\sqrt{j(j+1)-m'(m'-1)}\;\delta_{m,m'-1}\right] \quad \text{(C-60)}$$

and:

$$\langle k,j,m|J_y|k',j',m'\rangle = \frac{\hbar}{2i}\,\delta_{kk'}\,\delta_{jj'} \times \left[\sqrt{j(j+1)-m'(m'+1)}\;\delta_{m,m'+1}-\sqrt{j(j+1)-m'(m'-1)}\;\delta_{m,m'-1}\right] \quad \text{(C-61)}$$

As for the matrix $(J_z^{(j)})$, it is diagonal and its elements are the $(2j+1)$ values of $m\hbar$. The only non-zero matrix elements of $(J_x)^{(j)}$ and $(J_y)^{(j)}$ are those directly above and directly below the diagonal: $(J_x)^{(j)}$ is symmetrical and real, and $(J_y)^{(j)}$ is antisymmetrical and pure imaginary.

Since the kets $|k,j,m\rangle$ are, by construction, eigenvectors of $\mathbf{J}^2$, we have:

$$\langle k,j,m|\mathbf{J}^2|k',j',m'\rangle = j(j+1)\hbar^2\;\delta_{kk'}\,\delta_{jj'}\,\delta_{mm'} \quad \text{(C-62)}$$

The matrix $(\mathbf{J}^2)^{(j)}$ is therefore proportional to the $(2j+1)\times(2j+1)$ unit matrix: its diagonal elements are all equal to $j(j+1)\hbar^2$.

Comment:

The Oz axis which we have chosen as the "quantization axis" is entirely arbitrary. All directions in space are physically equivalent, and we should expect the eigenvalues of J_x or J_y to be the same as those of J_z (their eigenvectors, however, are different, since J_x and J_y do not commute with J_z). It can indeed be verified that the eigenvalues of the $(J_x)^{(1/2)}$ and $(J_y)^{(1/2)}$ matrices [formulas (C-54)] are $\pm\frac{\hbar}{2}$, and that those of the $(J_x)^{(1)}$ and $(J_y)^{(1)}$ matrices [formulas (C-58)] are $+\hbar$, 0, $-\hbar$. This result is general: inside a given subspace $\mathcal{E}(k,j)$, the eigenvalues of J_x and J_y (like those of the component $J_u=\mathbf{J}.\mathbf{u}$ of $\mathbf{J}$ along an arbitrary unit vector $\mathbf{u}$) are $j\hbar,(j-1)\hbar,...,(-j+1)\hbar,-j\hbar$. The corresponding eigenvectors (eigenvectors common to $\mathbf{J}^2$ and J_x, $\mathbf{J}^2$ and J_y, or $\mathbf{J}^2$ and J_u) are linear combinations of the $|k,j,m\rangle$ with k and j fixed.

To conclude this section devoted to "standard" representations, we summarize:

An orthonormal basis $\{|k,j,m\rangle\}$ of the state space, composed of eigenvectors common to $\mathbf{J}^2$ and J_z:

$$\mathbf{J}^2|k,j,m\rangle = j(j+1)\hbar^2\ |k,j,m\rangle$$
$$J_z|k,j,m\rangle = m\hbar\ |k,j,m\rangle$$

is called a "standard basis " if the action of the operators J_+ and J_- on the basis vectors is given by:

$$J_+|k,j,m\rangle = \hbar\sqrt{j(j+1)-m(m+1)}\ |k,j,m+1\rangle$$
$$J_-|k,j,m\rangle = \hbar\sqrt{j(j+1)-m(m-1)}\ |k,j,m-1\rangle$$

D. Application to orbital angular momentum

In § C, we studied the general properties of angular momenta, derived uniquely from the commutation relations (B-9). We shall now return to the orbital angular momentum $\mathbf{L}$ of a spinless particle [formula (B-3)] and see how the general theory just developed applies to this particular case. Using the $\{|\mathbf{r}\rangle\}$ representation, we shall show that the eigenvalues of the operator $\mathbf{L}^2$ are the numbers $l(l+1)\hbar^2$ corresponding to *all positive integral or zero* l: of the possible values for j found in § C-2-b, the only ones allowed in this case are the integral values, all of which are present. Then we shall indicate the eigenfunctions common to $\mathbf{L}^2$ and L_z and their principal properties. Finally, we shall study these eigenstates from a physical point of view.

D-1. Eigenvalues and eigenfunctions of $\mathbf{L}^2$ and L_z

D-1-a. Eigenvalue equation in the $\{|\mathbf{r}\rangle\}$ representation

In the $\{|\mathbf{r}\rangle\}$ representation, the observables $\mathbf{R}$ and $\mathbf{P}$ correspond respectively to multiplication by $\mathbf{r}$ and to the differential operator $(\hbar/i)\boldsymbol{\nabla}$. The three components of the angular momentum $\mathbf{L}$ can then be written:

$$L_x = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \tag{D-1a}$$

$$L_y = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \tag{D-1b}$$

$$L_z = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \tag{D-1c}$$

It is more convenient to work in spherical (or polar) coordinates, since, as we shall see, the various angular momentum operators act only on the angular variables θ and φ, and not on the variable r. Instead of characterizing the vector $\mathbf{r}$ by its Cartesian components x, y, z, we label the corresponding point M in space ($\mathbf{OM} = \mathbf{r}$) by its

spherical coordinates r, θ, φ (Fig. 1):

$$\begin{cases} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \end{cases}$$

with : (D-2)

$$\begin{cases} r \geq 0 \\ 0 \leq \theta \leq \pi \\ 0 \leq \varphi < 2\pi \end{cases}$$

The volume element $\mathrm{d}^3r = \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z$ is written in spherical coordinates:

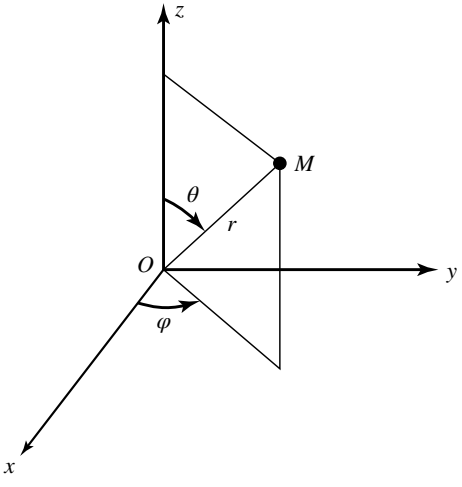


Figure 1: Definition of the spherical coordinates r, θ, φ of an arbitrary point in space.

$$\begin{aligned} \mathrm{d}^3r &= r^2\sin\theta\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}\varphi \\ &= r^2\,\mathrm{d}r\ \mathrm{d}\Omega \end{aligned} \tag{D-3}$$

where:

$$\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\ \mathrm{d}\varphi \tag{D-4}$$

is the solid angle element about the direction of polar angles θ and φ.

Applying the classical technique of changing variables, we obtain, from formulas (D-1) and (D-2), the following expressions (the calculations are rather time-consuming

but pose no great problem):

$$L_x = i\hbar \left(\sin\varphi \frac{\partial}{\partial\theta} + \frac{\cos\varphi}{\tan\theta} \frac{\partial}{\partial\varphi} \right) \tag{D-5a}$$

$$L_y = i\hbar \left(-\cos\varphi \frac{\partial}{\partial\theta} + \frac{\sin\varphi}{\tan\theta} \frac{\partial}{\partial\varphi} \right) \tag{D-5b}$$

$$L_z = \frac{\hbar}{i} \frac{\partial}{\partial\varphi} \tag{D-5c}$$

which yields:

$$\mathbf{L}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\tan\theta} \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right) \tag{D-6a}$$

$$L_+ = \hbar e^{i\varphi} \left(\frac{\partial}{\partial\theta} + i\ \cot\theta \frac{\partial}{\partial\varphi} \right) \tag{D-6b}$$

$$L_- = \hbar e^{-i\varphi} \left(-\frac{\partial}{\partial\theta} + i\ \cot\theta \frac{\partial}{\partial\varphi} \right) \tag{D-6c}$$

In the $\{|\mathbf{r}\rangle\}$ representation, the eigenfunctions associated with the eigenvalues $l(l+1)\hbar^2$ of $\mathbf{L}^2$ and $m\hbar$ of L_z are the solutions of the partial differential equations:

$$\begin{cases} -\left\{ \dfrac{\partial^2}{\partial\theta^2} + \dfrac{1}{\tan\theta} \dfrac{\partial}{\partial\theta} + \dfrac{1}{\sin^2\theta} \dfrac{\partial^2}{\partial\varphi^2} \right\} \psi(r,\theta,\varphi) = l(l+1)\ \psi(r,\theta,\varphi) & \text{(D-7a)} \\ -i \dfrac{\partial}{\partial\varphi} \psi(r,\theta,\varphi) = m\ \psi(r,\theta,\varphi) & \text{(D-7b)} \end{cases}$$

Since the general results of § C are applicable to the orbital angular momentum, we already know that l is integral or half-integral and that, for fixed l, m can take on only the values $-l$, $-l+1$, ..., $l-1$, l.

In equations (D-7), r does not appear in any differential operator, so we can consider it to be a parameter and *take into account only the θ- and φ-dependence of ψ*. Thus, we denote by $Y_l^m(\theta,\varphi)$ a common eigenfunction of $\mathbf{L}^2$ and L_z which corresponds to the eigenvalues $l(l+1)\hbar^2$ and $m\hbar$:

$$\mathbf{L}^2 Y_l^m(\theta,\varphi) = l(l+1)\hbar^2\ Y_l^m(\theta,\varphi) \tag{D-8a}$$

$$L_z Y_l^m(\theta,\varphi) = m\hbar\ Y_l^m(\theta,\varphi) \tag{D-8b}$$

To be completely rigorous, we would have to introduce an additional index in order to distinguish between the various solutions of (D-8) corresponding to the same pair of values of l and m. In fact, as we shall see further on, these equations have only one solution (to within a constant factor) for each pair of allowed values of l and m; this is why the indices l and m are sufficient.

Comments:

(i) Equations (D-8) give the θ- and φ-dependence of the eigenfunctions of $\mathbf{L}^2$ and L_z. Once the solution $Y_l^m(\theta,\varphi)$ of these equations has been found, these eigenfunctions will be obtained in the form:

$$\psi_{l,m}(r,\theta,\varphi) = f(r)\ Y_l^m(\theta,\varphi) \tag{D-9}$$

where $f(r)$ is a function[6] of r that appears as an integration constant for the partial differential equations (D-7). The fact that $f(r)$ is arbitrary shows that $\mathbf{L}^2$ and L_z do not form a C.S.C.O. in the space $\mathcal{E}_r$ of functions of $\mathbf{r}$ (or of r, θ, φ).

(*ii*) In order to normalize $\psi_{l,m}(r, \theta, \varphi)$, it is convenient to normalize $Y_l^m(\theta, \varphi)$ and $f(r)$ separately (as we shall do here). We then have, taking (D-4) into account:

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \; |Y_l^m(\theta, \varphi)|^2 \, \mathrm{d}\theta = 1 \tag{D-10}$$

and:

$$\int_0^{\infty} r^2 \, |f(r)|^2 \; \mathrm{d}r = 1 \tag{D-11}$$

D-1-b. Values of l and m

α. *l and m must be integral*

Using expression (D-5c) for L_z, we can write (D-8b) in the form:

$$\frac{\hbar}{i} \frac{\partial}{\partial \varphi} Y_l^m(\theta, \varphi) = m\hbar \; Y_l^m(\theta, \varphi) \tag{D-12}$$

which shows that $Y_l^m(\theta, \varphi)$ is equal to:

$$Y_l^m(\theta, \varphi) = F_l^m(\theta) \; \mathrm{e}^{im\varphi} \tag{D-13}$$

We can cover all space by letting φ vary between 0 and 2π. Since a wave function must be continuous at all points in space[7], we must have, in particular:

$$Y_l^m(\theta, \varphi = 0) = Y_l^m(\theta, \varphi = 2\pi) \tag{D-14}$$

which implies that:

$$\mathrm{e}^{2im\pi} = 1 \tag{D-15}$$

According to the results of § C, m is integral or half-integral. Relation (D-15) shows that, *in the case of an orbital angular momentum, m must be an integer* ($\mathrm{e}^{2im\pi}$ would be equal to -1 if m were half-integral). But we know that m and l are either both integral or both half-integral: it follows that *l must also be an integer.*

β. *All integral values (positive or zero) of l can be found*

Choose an integral value of l (positive or zero). We know from the general theory of § C that $Y_l^l(\theta, \varphi)$ must satisfy:

$$L_+ Y_l^l(\theta, \varphi) = 0 \tag{D-16}$$

[6] $f(r)$ must be such that $\psi_{l,m}(r, \theta, \varphi)$ is square-integrable.

[7] If $Y_l^m(\theta, \varphi)$ were not continuous for $\varphi = 0$, it would not be differentiable and could not be an eigenfunction of the differential operators (D-5c) and (D-6a). For example, $\frac{\partial}{\partial \theta} Y_l^m(\theta, \varphi)$ would produce a function $\delta(\varphi)$, which is incompatible with (D-12).

which yields, taking (D-6b) and (D-13) into account:

$$\left\{ \frac{\mathrm{d}}{\mathrm{d}\theta} - l \ \cot\theta \right\} F_l^l(\theta) = 0 \tag{D-17}$$

This first order equation can be integrated immediately if we notice that:

$$\cot\theta \ \mathrm{d}\theta = \frac{\mathrm{d}(\sin\theta)}{\sin\theta} \tag{D-18}$$

Its general solution is:

$$F_l^l(\theta) = c_l(\sin\theta)^l \tag{D-19}$$

where c_l is a normalization constant[8].

Consequently, for each positive or zero integral value of l, there exists a function $Y_l^l(\theta,\varphi)$ which is unique (to within a constant factor):

$$Y_l^l(\theta,\varphi) = c_l(\sin\theta)^l \mathrm{e}^{il\varphi} \tag{D-20}$$

Through the repeated action of L_-, we construct $Y_l^{l-1}, ..., Y_l^m, ..., Y_l^{-l}$. Thus we see that there corresponds to the pair of eigenvalues $l(l+1)\hbar^2$ and $m\hbar$ (where l is an arbitrary positive integer or zero and m is another integer such that $-l \leq m \leq l$), *one and only one eigenfunction* $Y_l^m(\theta,\varphi)$, which can be unambiguously calculated from (D-20). The eigenfunctions $Y_l^m(\theta,\varphi)$ are called *spherical harmonics*.

D-1-c. Fundamental properties of the spherical harmonics

The spherical harmonics $Y_l^m(\theta,\varphi)$ will be studied in greater detail in Complement $\mathrm{A_{VI}}$. Here we shall confine ourselves to summarizing this study by stating without proof its principal results.

α. Recurrence relations

According to the general results of § C, we have:

$$L_\pm Y_l^m(\theta,\varphi) = \hbar\sqrt{l(l+1) - m(m\pm 1)} \ Y_l^{m\pm 1}(\theta,\varphi) \tag{D-21}$$

Using expressions (D-6b) and (D-6c) for the operators L_+ and L_- and the fact that $Y_l^m(\theta,\varphi)$ is the product of a function of θ alone and $\mathrm{e}^{im\varphi}$, we obtain:

$$\mathrm{e}^{i\varphi}\left(\frac{\partial}{\partial\theta} - m \ \cot\theta\right) Y_l^m(\theta,\varphi) = \sqrt{l(l+1) - m(m+1)} \ Y_l^{m+1}(\theta,\varphi) \tag{D-22a}$$

$$\mathrm{e}^{-i\varphi}\left(-\frac{\partial}{\partial\theta} - m \ \cot\theta\right) Y_l^m(\theta,\varphi) = \sqrt{l(l+1) - m(m-1)} \ Y_l^{m-1}(\theta,\varphi) \tag{D-22b}$$

[8] Inversely, one can easily show that the function obtained in this way is actually an eigenfunction of $\mathbf{L}^2$ and L_z with eigenvalues $l(l+1)\hbar^2$ and $l\hbar$. According to (D-5c) and (D-13), it is immediately seen that $L_z Y_l^l(\theta,\varphi) = l\hbar Y_l^l(\theta,\varphi)$. Then, using this equation and (D-16), as well as (C-7b), one can show that $Y_l^l(\theta,\varphi)$ is also an eigenfunction of $\mathbf{L}^2$ with the expected eigenvalue.

β. Orthonormalization and closure relations

Equations (D-7) determine the spherical harmonics only to within a constant factor. We now choose this factor so as to orthonormalize the $Y_l^m(\theta,\varphi)$ (as functions of the angular variables θ and φ):

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \ \mathrm{d}\theta \ Y_{l'}^{m'*}(\theta,\varphi) \ Y_l^m(\theta,\varphi) = \delta_{l'l} \ \delta_{m'm} \tag{D-23}$$

Furthermore, any function of θ and φ, $f(\theta,\varphi)$, can be expanded in terms of the spherical harmonics:

$$f(\theta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} c_{l,m} \ Y_l^m(\theta,\varphi) \tag{D-24}$$

with:

$$c_{l,m} = \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \ \mathrm{d}\theta \ Y_l^{m*}(\theta,\varphi) \ f(\theta,\varphi) \tag{D-25}$$

The spherical harmonics therefore constitute an orthonormal basis in the space $\mathcal{E}_\Omega$ of functions of θ and φ. This fact is expressed by the closure relation:

$$\begin{aligned} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(\theta,\varphi) \ Y_l^{m*}(\theta',\varphi') &= \delta(\cos\theta - \cos\theta') \ \delta(\varphi - \varphi') \\ &= \frac{1}{\sin\theta} \ \delta(\theta - \theta') \ \delta(\varphi - \varphi') \end{aligned} \tag{D-26}$$

[it is $\delta(\cos\theta - \cos\theta')$, and not $\delta(\theta - \theta')$, that enters into the right-hand side of the closure relation, because the integrations over the variable θ are performed using the differential element $\sin\theta \mathrm{d}\theta = -\mathrm{d}(\cos\theta)$].

γ. Parity and complex conjugation

First of all, recall that the change from $\mathbf{r}$ to $-\mathbf{r}$ (reflection through the coordinate origin) is expressed in spherical coordinates by (Fig. 2):

$$\begin{aligned} r &\Longrightarrow r \\ \theta &\Longrightarrow \pi - \theta \\ \varphi &\Longrightarrow \pi + \varphi \end{aligned} \tag{D-27}$$

It is simple (see Complement $\mathrm{A_{VI}}$) to show that:

$$Y_l^m(\pi - \theta, \pi + \varphi) = (-1)^l \ Y_l^m(\theta,\varphi) \tag{D-28}$$

The spherical harmonics are therefore functions with a definite parity, which is independent of m; they are even if l is even and odd if l is odd.

Also, it can easily be seen that:

$$[Y_l^m(\theta,\varphi)]^* = (-1)^m \ Y_l^{-m}(\theta,\varphi) \tag{D-29}$$

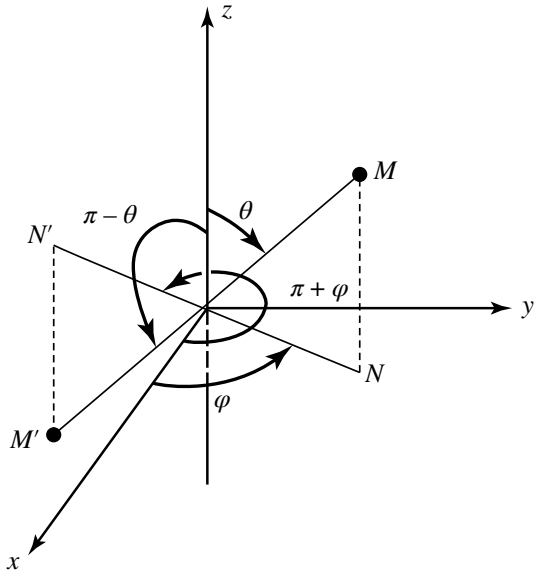


Figure 2: Transformation in spherical coordinates of an arbitrary point by reflection through the origin; r is not changed, θ becomes $\pi - \theta$, and φ becomes $\pi + \varphi$.

D-1-d. "Standard" bases of the wave function space of a spinless particle

As we have already noted [comment (i) of § D-1-a], $\mathbf{L}^2$ and L_z do not constitute a C.S.C.O. in the wave function space of a spinless particle. We shall now indicate, relying on the reasoning and results of § C-3, the form of the "standard" bases of this space.

Let $\mathcal{E}(l, m = l)$ be the subspace of eigenfunctions common to $\mathbf{L}^2$ and L_z, of eigenvalues $l(l+l)\hbar^2$ and $l\hbar$, where l is a fixed positive integer or zero. The first step in the construction of a "standard" basis (*cf.* § C-3) consists of choosing an arbitrary orthonormal basis in each of the $\mathcal{E}(l, m = l)$. We shall denote by $\psi_{k,l,l}(\mathbf{r})$ the functions that constitute the basis chosen in $\mathcal{E}(l, m = l)$, the index k (assumed to be discrete for simplicity) serving to distinguish between the various functions of this basis. By repeated application of the operator L_- on the $\psi_{k,l,l}(\mathbf{r})$, we then construct the functions $\psi_{k,l,m}(\mathbf{r})$ which complete the "standard" basis for $m \neq l$; they satisfy equations (C-12) and (C-50), which become here:

$$\begin{aligned} &\mathbf{L}^2 \psi_{k,l,m}(\mathbf{r}) = l(l+1)\hbar^2 \ \psi_{k,l,m}(\mathbf{r}) \\ &L_z \psi_{k,l,m}(\mathbf{r}) = m\hbar \ \psi_{k,l,m}(\mathbf{r}) \end{aligned} \tag{D-30}$$

and:

$$L_\pm \psi_{k,l,m}(\mathbf{r}) = \hbar\sqrt{l(l+1) - m(m \pm 1)} \ \psi_{k,l,m\pm1}(\mathbf{r}) \tag{D-31}$$

But we saw in § D-1-a that all eigenfunctions common to $\mathbf{L}^2$ and L_z that correspond to given eigenvalues $l(l+1)\hbar^2$ and $m\hbar$ have the same angular dependence, that of $Y_l^m(\theta, \varphi)$; only their radial dependence differs. From equations (D-30), we therefore deduce that $\psi_{k,l,m}(\mathbf{r})$ has the form:

$$\psi_{k,l,m}(\mathbf{r}) = R_{k,l,m}(r) \ Y_l^m(\theta, \varphi) \tag{D-32}$$

Let us now show that, if the $\psi_{k,l,m}(\mathbf{r})$ constitute a "standard" basis, the radial functions $R_{k,l,m}(\mathbf{r})$ are independent of m. Since the differential operators $L_\pm$ do not act on the r-dependence, we have, according to (D-21):

$$\begin{aligned} L_\pm\ \psi_{k,l,m}(\mathbf{r}) &= R_{k,l,m}(\mathbf{r})\ L_\pm\ Y_l^m(\theta,\varphi) \\ &= \hbar\sqrt{l(l+1)-m(m\pm1)}\ R_{k,l,m}(r)\ Y_l^{m\pm1}(\theta,\varphi) \end{aligned} \tag{D-33}$$

Comparison with (D-31) shows that the radial functions must satisfy, for all r:

$$R_{k,l,m\pm1}(r) = R_{k,l,m}(r) \tag{D-34}$$

and are consequently independent of m. The functions $\psi_{k,l,m}(\mathbf{r})$ of a "standard" basis of the wave function space of a (spinless) particle are therefore necessarily of the form:

$$\psi_{k,l,m}(\mathbf{r}) = R_{k,l}(r)\ Y_l^m(\theta,\varphi) \tag{D-35}$$

The orthonormalization relation for such a basis is:

$$\begin{aligned} \int \mathrm{d}^3r\ \psi^*_{k,l,m}(\mathbf{r})\,\psi_{k',l',m'}(\mathbf{r}) &= \int_0^\infty r^2\mathrm{d}r\ R^*_{k,l}(r)\,R_{k',l'}(r) \\ &\times \int_0^{2\pi}\mathrm{d}\varphi \int_0^\pi \sin\theta\,\mathrm{d}\theta\ Y_l^{m*}(\theta,\varphi)\,Y_{l'}^{m'}(\theta,\varphi) = \delta_{kk'}\ \delta_{ll'}\ \delta_{mm'} \end{aligned} \tag{D-36}$$

Since the spherical harmonics are orthonormal [formula (D-23)], we obtain, finally:

$$\int_0^\infty r^2\mathrm{d}r\ R^*_{k,l}(r)\,R_{k',l}(r) = \delta_{kk'} \tag{D-37}$$

The radial functions $R_{k,l}(r)$ are therefore normalized with respect to the variable r; moreover, two radial functions corresponding to the same value of l but to different indices k are orthogonal.

Comments:

(*i*) Relation (D-37) is simply a consequence of the orthonormality of the functions $\psi_{k,l,m}(\mathbf{r}) = R_{k,l}(r)\,Y_l^l(\theta,\varphi)$, which have been chosen as a basis in the subspace $\mathcal{E}(l, m=l)$. It is therefore essential that the index l be the same for the two functions $R_{k,l}$ appearing on the left-hand side. For $l \neq l'$, $\psi_{k,l,m}(\mathbf{r})$ and $\psi_{k',l',m'}(\mathbf{r})$ are orthogonal anyway because of their angular dependence (they are eigenfunctions of the Hermitian operator $\mathbf{L}^2$ with different eigenvalues). The integral:

$$\int_0^\infty r^2\mathrm{d}r\ R^*_{k,l}(r)\,R_{k',l'}(r) \tag{D-38}$$

may therefore take on any value *a priori* if l and l' are different.

(*ii*) In general, the radial functions $R_{k,l}(r)$ depend on l, for the following reason. A function of the form $f(r)g(\theta,\varphi)$ can be continuous at the coordinate origin ($r=0$, θ and φ arbitrary) only if $g(\theta,\varphi)$ reduces to a constant or if $f(r)$ goes to zero at $r=0$ [since if $g(\theta,\varphi)$ depends on θ and φ, the limit of $f(r)g(\theta,\varphi)$ when $r \longrightarrow 0$ depends on the direction along which one approaches the origin if $f(0)$ is not zero]. Consequently, if we want the basis functions $\psi_{k,l,m}(\mathbf{r})$ to be continuous, only the radial functions corresponding to $l=0$ can be non-zero at $r=0$ [$Y_0^0(\theta,\varphi)$ is indeed a constant]. Similarly, if we require the $\psi_{k,l,m}(\mathbf{r})$ to be differentiable (once or several times) at the origin, we obtain conditions for the $R_{k,l}(\mathbf{r})$ that depend on the value of l.

D-2. Physical considerations

D-2-a. Study of a $|k,l,m\rangle$ state

Consider a (spinless) particle in an eigenstate $|k,l,m\rangle$ of $\mathbf{L}^2$ and L_z [whose associated wave function is $\psi_{k,l,m}(\mathbf{r})$], that is, a state in which the square of its angular momentum and the projection of this angular momentum along the Oz axis have well-defined values [$l(l+1)\hbar^2$ and $m\hbar$ respectively].

Suppose that we want to measure the component along the Ox or Oy axis of the angular momentum of this particle. Since L_x and L_y do not commute with L_z, $|k,l,m\rangle$ is an eigenstate neither of L_x nor of L_y; we cannot, therefore, predict with certainty the result of such a measurement. Let us calculate the mean values and root mean square deviations of L_x and L_y in the state $|k,l,m\rangle$.

These calculations can be performed very simply if we express L_x and L_y in terms of L_+ and L_-. We invert formulas (C-1):

$$\begin{aligned} L_x &= \frac{1}{2}(L_+ + L_-) \\ L_y &= \frac{1}{2i}(L_+ - L_-) \end{aligned} \tag{D-39}$$

Thus we see that $L_x\,|k,l,m\rangle$ and $L_y\,|k,l,m\rangle$ are linear combinations of $|k,l,m+1\rangle$ and $|k,l,m-1\rangle$; this leads to:

$$\langle k,l,m|L_x|k,l,m\rangle = \langle k,l,m|L_y|k,l,m\rangle = 0 \tag{D-40}$$

Furthermore:

$$\begin{aligned} \langle k,l,m|L_x^2|k,l,m\rangle &= \frac{1}{4}\langle k,l,m|(L_+^2 + L_-^2 + L_+L_- + L_-L_+)|k,l,m\rangle \\ \langle k,l,m|L_y^2|k,l,m\rangle &= -\frac{1}{4}\langle k,l,m|(L_+^2 + L_-^2 - L_+L_- - L_-L_+)|k,l,m\rangle \end{aligned} \tag{D-41}$$

The terms in L_+^2 and L_-^2 do not contribute to the result, since $L_\pm^2|k,l,m\rangle$ is proportional to $|k,l,m\pm 2\rangle$. In addition, formula (C-8) yields:

$$L_+L_- + L_-L_+ = 2(\mathbf{L}^2 - L_z^2) \tag{D-42}$$

We therefore obtain:

$$\begin{aligned}\langle k,l,m|L_x^2|k,l,m\rangle &= \langle k,l,m|L_y^2|k,l,m\rangle \\ &= \frac{1}{2}\langle k,l,m|(\mathbf{L}^2 - L_z^2)|k,l,m\rangle \\ &= \frac{\hbar^2}{2}[l(l+1) - m^2]\end{aligned} \tag{D-43}$$

Thus, in the state $|k,l,m\rangle$:

$$\langle L_x\rangle = \langle L_y\rangle = 0 \tag{D-44a}$$

$$\Delta L_x = \Delta L_y = \hbar\sqrt{\frac{1}{2}[l(l+1) - m^2]} \tag{D-44b}$$

These results suggest the following picture. Consider a *classical* angular momentum, whose modulus is equal to $\hbar\sqrt{l(l+1)}$ and whose projection along Oz is $m\hbar$ (Fig. 3):

$$\begin{aligned}|\mathbf{OL}| &= \hbar\sqrt{l(l+1)} \\ \overline{\mathrm{OH}} &= m\hbar\end{aligned} \tag{D-45}$$

We denote by Θ and Φ the polar angles that characterize its direction. Since the triangle OLJ has a right angle at J, and $OH = JL$, we have:

$$OJ = \sqrt{OL^2 - OH^2} = \hbar\sqrt{l(l+1) - m^2} \tag{D-46}$$

Consequently, the components of such a classical angular momentum would be:

$$\begin{aligned}\overline{OI} &= \hbar\sqrt{l(l+1) - m^2}\cos\Phi \\ \overline{OK} &= \hbar\sqrt{l(l+1) - m^2}\sin\Phi \\ \overline{OH} &= \hbar\sqrt{l(l+1)}\cos\Theta = m\hbar\end{aligned} \tag{D-47}$$

Now let us assume that $|\mathbf{OL}|$ and Θ are known and that Φ is a random variable which can take on any value in the interval $[0, 2\pi]$, all these values being equally probable (an evenly distributed random variable). We then have, averaging over Φ:

$$\langle\overline{OI}\rangle \propto \int_0^{2\pi} \cos\Phi \, \mathrm{d}\Phi = 0 \tag{D-48a}$$

$$\langle\overline{OK}\rangle \propto \int_0^{2\pi} \sin\Phi \, \mathrm{d}\Phi = 0 \tag{D-48b}$$

which corresponds to (D-44a). In addition:

$$\langle\overline{OI}^2\rangle = \frac{1}{2\pi}\hbar^2[l(l+1) - m^2]\int_0^{2\pi}\cos^2\Phi \, \mathrm{d}\Phi = \frac{\hbar^2}{2}[l(l+1) - m^2] \tag{D-49}$$

and, similarly:

$$\langle\overline{OK}^2\rangle = \frac{\hbar^2}{2}[l(l+1) - m^2] \tag{D-50}$$

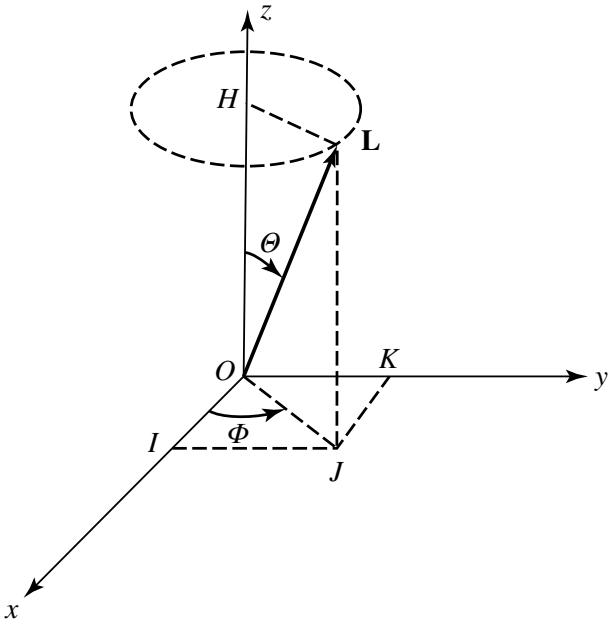


Figure 3: A classical model for the orbital angular momentum of a particle in a state $|l, m\rangle$. *We assume that the distance /OL/ and the angle* Θ *are known, but that* Φ *is a random variable whose probability density is constant inside the interval* $[0, 2\pi]$. *The classical mean values of the components of OL, as well as those of the squares of these components, are then equal to the corresponding quantum mechanical mean values.*

These mean values are identical to the ones we found in (D-44). Consequently, the angular momentum of a particle in the state $|k, l, m\rangle$, behaves, insofar as the mean values of its components and their squares are concerned, like a classical angular momentum of magnitude $\hbar\sqrt{l(l+1)}$ having a projection $m\hbar$ along Oz, but for which Φ is a random variable evenly distributed between 0 and 2π.

Of course, this picture must be used carefully: we have shown throughout this chapter how much the quantum mechanical properties of angular momenta differ from their classical properties. In particular, we must stress the fact that an individual measurement of L_x or L_y on a particle in the state $|k, l, m\rangle$ cannot yield an arbitrary value between $-\hbar\sqrt{l(l+1) - m^2}$ and $+\hbar\sqrt{l(l+1) - m^2}$, as the preceding model might lead us to believe. The only possible results are the eigenvalues of L_x or L_y (we saw at the end of § C that these are the same as those of L_z), that is, since l is fixed here, one of the $(2l + 1)$ values $l\hbar$, $(l - 1)\hbar$, ..., $(-l + 1)\hbar$, $-l\hbar$.

D-2-b. Calculation of the physical predictions concerning measurements of $\mathbf{L}^2$ and L_z

Consider a particle whose state is described by the (normalized) wave function:

$$\langle \mathbf{r}|\psi\rangle = \psi(\mathbf{r}) = \psi(r,\theta,\varphi) \tag{D-51}$$

We know that a measurement of $\mathbf{L}^2$ can yield only the results 0, $2\hbar^2$, $6\hbar^2$, ..., $l(l+1)\hbar^2$, ..., and a measurement of L_z, only the results 0, $\pm\hbar$, $\pm 2\hbar$, ..., $m\hbar$, ... How can we calculate the probabilities of these different results from the wave function $\psi(r,\theta,\varphi)$?

α. General formulas

Let us denote by $\mathscr{P}_{\mathbf{L}^2,L_z}(l,m)$ the probability of finding, in a simultaneous measurement of $\mathbf{L}^2$ and L_z, the results $l(l+1)\hbar^2$ and $m\hbar$. This probability can be obtained by expanding $\psi(\mathbf{r})$ on a basis composed of eigenfunctions of $\mathbf{L}^2$ and L_z; we shall choose a "standard" basis of the type introduced in § D-1-d:

$$\psi_{k,l,m}(\mathbf{r}) = R_{k,l}(r)\ Y_l^m(\theta,\varphi) \tag{D-52}$$

$\psi(\mathbf{r})$ can then be written:

$$\psi(\mathbf{r}) = \sum_k \sum_l \sum_m c_{k,l,m}\ R_{k,l}(r)\ Y_l^m(\theta,\varphi) \tag{D-53}$$

where the coefficients $c_{k,l,m}$ can be calculated by using the usual formula:

$$\begin{aligned} c_{k,l,m} &= \int \mathrm{d}^3r\ \psi^*_{k,l,m}(\mathbf{r})\ \psi(\mathbf{r}) \\ &= \int_0^\infty r^2\mathrm{d}r\ R^*_{k,l}(r) \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta\,\mathrm{d}\theta\ Y_l^{m*}(\theta,\varphi)\ \psi(r,\theta,\varphi) \end{aligned} \tag{D-54}$$

According to the postulates of Chapter III, the probability $\mathscr{P}_{\mathbf{L}^2,L_z}(l,m)$ is given, under these conditions, by:

$$\mathscr{P}_{\mathbf{L}^2,L_z}(l,m) = \sum_k |c_{k,l,m}|^2 \tag{D-55}$$

If we measure only $\mathbf{L}^2$, the probability of finding the result $l(l+1)\hbar^2$ is equal to:

$$\mathscr{P}_{\mathbf{L}^2}(l) = \sum_{m=-l}^{+l} \mathscr{P}_{\mathbf{L}^2,L_z}(l,m) = \sum_k \sum_{m=-l}^{+l} |c_{k,l,m}|^2 \tag{D-56}$$

Similarly, if it is only L_z that we wish to measure, the probability of obtaining $m\hbar$ is:

$$\mathscr{P}_{L_z}(m) = \sum_{l\geq|m|} \mathscr{P}_{\mathbf{L}^2,L_z}(l,m) = \sum_k \sum_{l\geq|m|} |c_{k,l,m}|^2 \tag{D-57}$$

(the restriction $l \geq |m|$ is automatically satisfied, since there are no coefficients $c_{k,l,m}$ for which $|m|$ would be greater than l).

Actually, since $\mathbf{L}^2$ and L_z act only on θ and φ, we see that it is the θ- and φ-dependence of the wave function $\psi(\mathbf{r})$ that count in the preceding probability calculations. To be more precise, consider $\psi(r,\theta,\varphi)$ as a function of θ and φ depending on the

parameter r. Like any other function of θ and φ, ψ can then be expanded in terms of the spherical harmonics:

$$\psi(r,\theta,\varphi) = \sum_l \sum_m a_{l,m}(r)\ Y_l^m(\theta,\varphi) \tag{D-58}$$

The coefficients $a_{l,m}$ of this expansion depend on the "parameter" r and are given by:

$$a_{l,m}(r) = \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta\,\mathrm{d}\theta\ Y_l^{m*}(\theta,\varphi)\ \psi(r,\theta,\varphi) \tag{D-59}$$

If we compare expressions (D-58) and (D-53), we see that the $c_{k,l,m}$ are the coefficients of the expansion of $a_{l,m}(r)$ on the functions $R_{k,l}(r)$:

$$a_{l,m}(r) = \sum_k c_{k,l,m}\ R_{k,l}(r) \tag{D-60}$$

with, taking (D-54) and (D-59) into account:

$$c_{k,l,m} = \int_0^{\infty} r^2\mathrm{d}r\ R_{k,l}^*(r) a_{l,m}(r) \tag{D-61}$$

Using (D-37) and (D-60), we also obtain:

$$\int_0^{\infty} r^2\mathrm{d}r\ |a_{l,m}(r)|^2 = \sum_k |c_{k,l,m}|^2 \tag{D-62}$$

The probability $\mathscr{P}_{\mathbf{L}^2,L_z}(l,m)$ [formula (D-55)] can therefore also be written in the form:

$$\mathscr{P}_{\mathbf{L}^2,L_z}(l,m) = \int_0^{\infty} r^2\mathrm{d}r\ |a_{l,m}(r)|^2 \tag{D-63}$$

From this we can deduce, as in (D-56) and (D-57):

$$\mathscr{P}_{\mathbf{L}^2}(l) = \sum_{m=-l}^{+l} \int_0^{\infty} r^2\mathrm{d}r\ |a_{l,m}(r)|^2 \tag{D-64}$$

and:

$$\mathscr{P}_{L_z}(m) = \sum_{l\geq|m|} \int_0^{\infty} r^2\mathrm{d}r\ |a_{l,m}(r)|^2 \tag{D-65}$$

[here again, $a_{l,m}(r)$ exists only for $l \geq |m|$]. Consequently, to obtain the physical predictions concerning measurements of $\mathbf{L}^2$ and L_z, we may consider the wave function as depending only on θ and φ. We then expand it in terms of the spherical harmonics as in (D-58) and apply formulas (D-63), (D-64) and (D-65).

Similarly, since L_z acts only on φ, it is the φ-dependence of the wave function $\psi(\mathbf{r})$ that counts in the calculation of $\mathscr{P}_{L_z}(m)$. To see this, we shall use the fact that the

spherical harmonics are products of a function of θ alone and a function of φ alone. We shall write them in the form:

$$Y_l^m(\theta,\varphi) = Z_l^m(\theta)\ \frac{\mathrm{e}^{im\varphi}}{\sqrt{2\pi}} \tag{D-66}$$

so that each of the functions of the product is normalized, since we have:

$$\int_0^{2\pi} \mathrm{d}\varphi\ \frac{\mathrm{e}^{-im\varphi}}{\sqrt{2\pi}}\ \frac{\mathrm{e}^{im'\varphi}}{\sqrt{2\pi}} = \delta_{mm'} \tag{D-67}$$

Substituting this formula into the orthonormalization relation (D-23) for the spherical harmonics, we find:

$$\int_0^{\pi} \sin\theta\,\mathrm{d}\theta\ Z_l^{m*}(\theta)\ Z_{l'}^m(\theta) = \delta_{ll'} \tag{D-68}$$

[for reasons analogous to those indicated in comment (i) of § D-1-d, the same value of m is involved in both functions Z_l^m of the left-hand side].

If we consider $\psi(r,\theta,\varphi)$ to be a function of φ defined in the interval $[0,2\pi]$ and depending on the "parameters" r and θ, we can expand it in a Fourier series:

$$\psi(r,\theta,\varphi) = \sum_m b_m(r,\theta)\ \frac{\mathrm{e}^{im\varphi}}{\sqrt{2\pi}} \tag{D-69}$$

where the coefficients $b_m(r,\theta)$ can be calculated from the formula:

$$b_m(r,\theta) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \mathrm{d}\varphi\ \mathrm{e}^{-im\varphi}\ \psi(r,\theta,\varphi) \tag{D-70}$$

If we compare formulas (D-69) and (D-70) with (D-58) and (D-59), we see that the $a_{l,m}(r)$ for fixed m are the coefficients of the expansion of $b_m(r,\theta)$ on the functions Z_l^m corresponding to the same value of m:

$$b_m(r,\theta) = \sum_l a_{l,m}(r)\ Z_l^m(\theta) \tag{D-71}$$

with:

$$a_{l,m}(r) = \int_0^{\pi} \sin\theta\,\mathrm{d}\theta\ Z_l^{m*}(\theta)\ b_m(r,\theta) \tag{D-72}$$

With (D-68) taken into account, expansion (D-71) requires that:

$$\int_0^{\pi} \sin\theta\,\mathrm{d}\theta\ |b_m(r,\theta)|^2 = \sum_l |a_{l,m}(r)|^2 \tag{D-73}$$

Substituting this formula into (D-65), we obtain $\mathscr{P}_{L_z}(m)$ in the form:

$$\mathscr{P}_{L_z}(m) = \int_0^{\infty} r^2\mathrm{d}r \int_0^{\pi} \sin\theta\,\mathrm{d}\theta\ |b_m(r,\theta)|^2 \tag{D-74}$$

Therefore, as far as measurements of L_z alone are concerned, all we need to do is consider the wave function as depending solely on φ and expand it in a Fourier series as in (D-69) in order to calculate the probabilities of the various possible results.

We might be tempted to think that an argument analogous to the preceding ones would give $\mathscr{P}_{\mathbf{L}^2}(l)$ in terms of the expansion of $\psi(r,\theta,\varphi)$ with respect to the variable θ alone. In fact, this is not the case: predictions concerning a measurement of $\mathbf{L}^2$ alone involve both the θ- and the φ-dependence of the wave function; this is related to the fact that $\mathbf{L}^2$ acts on both θ and φ. We must therefore use formula (D-64).

β. Special cases and examples

Suppose that the wave function $\psi(\mathbf{r})$ representing the state of the particle appears in the form of a product of a function of r alone and a function of θ and φ:

$$\psi(r,\theta,\varphi) = f(r)\ g(\theta,\varphi) \tag{D-75}$$

We can always assume $f(r)$ and $g(\theta,\varphi)$ to be separately normalized:

$$\int_0^\infty r^2 \mathrm{d}r\ |f(r)|^2 = 1 \tag{D-76a}$$

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta\, \mathrm{d}\theta\ |g(\theta,\varphi)|^2 = 1 \tag{D-76b}$$

To obtain the expansion (D-58) of such a wave function, all we must do is expand $g(\theta,\varphi)$ in terms of the spherical harmonics:

$$g(\theta,\varphi) = \sum_l \sum_m d_{l,m}\ Y_l^m(\theta,\varphi) \tag{D-77}$$

with:

$$d_{l,m} = \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta\, \mathrm{d}\theta\ Y_l^{m*}(\theta,\varphi)\ g(\theta,\varphi) \tag{D-78}$$

In this case, therefore, the coefficients $a_{l,m}(r)$ of formula (D-58) are all proportional to $f(r)$:

$$a_{l,m}(r) = d_{l,m}\ f(r) \tag{D-79}$$

With (D-76a) taken into account, expression (D-63) for the probability $\mathscr{P}_{\mathbf{L}^2,L_z}(l,m)$ becomes simply:

$$\mathscr{P}_{\mathbf{L}^2,L_z}(l,m) = |d_{l,m}|^2 \tag{D-80}$$

This probability is totally *independent of the radial part $f(r)$ of the wave function.*

Similarly, let us consider the case in which the wave function $\psi(r,\theta,\varphi)$ is the product of three functions of a single variable:

$$\psi(r,\theta,\varphi) = f(r)\ h(\theta)\ k(\varphi) \tag{D-81}$$

which we shall assume to be separately normalized:

$$\int_0^\infty r^2 \mathrm{d}r \; |f(r)|^2 = \int_0^\pi \sin\theta \, \mathrm{d}\theta \; |h(\theta)|^2 = \int_0^{2\pi} \mathrm{d}\varphi \; |k(\varphi)|^2 = 1 \tag{D-82}$$

Of course, (D-81) is a special case of (D-75), and the results we have just established apply here. But, in addition, if we are interested only in a measurement of L_z, all we must do is expand $k(\varphi)$ in the form:

$$k(\varphi) = \sum_m e_m \; \frac{\mathrm{e}^{im\varphi}}{\sqrt{2\pi}} \tag{D-83}$$

where:

$$e_m = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \mathrm{d}\varphi \; \mathrm{e}^{-im\varphi} \; k(\varphi) \tag{D-84}$$

in order to obtain the equivalent of formula (D-69), with:

$$b_m(r,\theta) = e_m \; f(r) \; h(\theta) \tag{D-85}$$

According to (D-82), $\mathscr{P}_{L_z}(m)$ is then given by (D-74) as:

$$\mathscr{P}_{L_z}(m) = |e_m|^2 \tag{D-86}$$

The preceding considerations can be illustrated by some very simple examples. First, let us assume that the wave function $\psi(\mathbf{r})$ is in fact independent of θ and φ, so that:

$$\begin{cases} h(\theta) = \dfrac{1}{\sqrt{2}} \\ k(\varphi) = \dfrac{1}{\sqrt{2\pi}} \end{cases} \tag{D-87}$$

We then have:

$$g(\theta,\varphi) = \frac{1}{\sqrt{4\pi}} = Y_0^0(\theta,\varphi) \tag{D-88}$$

Thus a measurement of $\mathbf{L}^2$ or of L_z must yield zero.

Now, let us modify only the θ-dependence, choosing:

$$\begin{cases} h(\theta) = \sqrt{\dfrac{3}{2}} \cos\theta \\ k(\varphi) = \dfrac{1}{\sqrt{2\pi}} \end{cases} \tag{D-89}$$

In this case:

$$g(\theta,\varphi) = \sqrt{\frac{3}{4\pi}} \cos\theta = Y_1^0(\theta,\varphi) \tag{D-90}$$

We are again sure of the results of measuring $\mathbf{L}^2$ or L_z. For $\mathbf{L}^2$, we can only obtain $2\hbar^2$; for L_z, 0. It can be verified that this modification of the θ-dependence has not changed the physical predictions concerning the measurement of L_z.

On the other hand, if we modify the φ-dependence by setting, for example:

$$\begin{cases} h(\theta) = \dfrac{1}{\sqrt{2}} \\ g(\varphi) = \dfrac{\mathrm{e}^{i\varphi}}{\sqrt{2\pi}} \end{cases} \tag{D-91}$$

$g(\theta, \varphi)$ is no longer equal to a single spherical harmonic. According to (D-86), all the probabilities $\mathscr{P}_{L_z}(m)$ are zero except for:

$$\mathscr{P}_{L_z}(m = 1) = 1 \tag{D-92}$$

But the predictions concerning a measurement of $\mathbf{L}^2$ are also changed with respect to the case (D-87). In order to calculate these predictions, we must expand the function:

$$g(\theta, \varphi) = \frac{1}{\sqrt{4\pi}} \mathrm{e}^{i\varphi} \tag{D-93}$$

on the spherical harmonics. It can be verified that all the $Y_l^m(\theta, \varphi)$, with odd l and $m = 1$, actually appear in the expansion of the function (D-93). Consequently, we are no longer sure of the result of a measurement of $\mathbf{L}^2$ (the probabilities of the various possible results can be calculated from the expression for the spherical harmonics). We therefore conclude from this example that, as pointed out at the end of § D-2-b-α, the φ-dependence of the wave function also enters into the calculation of predictions concerning measurements of $\mathbf{L}^2$.

References and suggestions for further reading:

Dirac (1.13), §§ 35 and 36; Messiah (1.17), Chap. XIII; Rose (2.19); Edmonds (2.21).

COMPLEMENTS OF CHAPTER VI, READER'S GUIDE

A_{VI} : SPHERICAL HARMONICS	Detailed study of the spherical harmonics $Y_l^m(\theta, \varphi)$; establishes certain properties used in Chapter VI, as well as in certain subsequent complements.
B_{VI} : ANGULAR MOMENTUM AND ROTATIONS	Brings out the close relation that exists between the angular momentum **J** of a quantum mechanical system and the spatial rotations that can be performed on it. Shows that the relation commutations between the components of **J** express purely geometrical properties of these rotations; introduces the concept of a scalar or vector observable, which will reappear in other complements (especially D_X). Important theoretically; however, sometimes difficult; can be reserved for later study.
C_{VI} : ROTATION OF DIATOMIC MOLECULES	A simple and direct application of quantum mechanical properties of angular momentum: pure rotational spectra of heteropolar diatomic molecules, and Raman rotational spectra. Elementary level. Because of the importance of the phenomena studied in physics and chemistry, can be recommended for a first reading.
D_{VI} : ANGULAR MOMENTUM OF STATIONNARY STATES OF A TWO-DIMENSIONAL HARMONIC OSCILLATOR	Can be considered as a worked example. Studies the stationnary states of the two-dimensional harmonic oscillator; in order to class these states by angular momentum, introduces the concept of "circular quanta". Not theoretically difficult. Some results will be used in Complement E_{VI}.
E_{VI} : A CHARGED PARTICLE IN A MAGNETIC FIELD: LANDAU LEVELS	A general study of the quantum mechanical properties of a charged particle in a magnetic field, followed by a study of the important special case in which the magnetic field is uniform (Landau levels). Not theoretically difficult. Recommmended for a first reading, which can, however, be confined to §§ 1-a and 1-b, 2-a and 2-b, 3-a.
F_{VI} : EXERCISES	.

Complement A_{VI}

Spherical harmonics

1 Calculation of spherical harmonics 705
1-a Determination of $Y_l^l(\theta, \varphi)$ 705
1-b General expression for $Y_l^m(\theta, \varphi)$ 707
1-c Explicit expressions for $l = 0$, 1 and 2 709
2 Properties of spherical harmonics 710
2-a Recurrence relations 710
2-b Orthonormalization and closure relations 711
2-c Parity 711
2-d Complex conjugation 712
2-e Relation between the spherical harmonics and the Legendre polynomials and associated Legendre functions 712

This complement is devoted to the study of the form and principal properties of spherical harmonics. It includes the proofs of certain results that were stated without proof in § D-1-c of Chapter VI.

1. Calculation of spherical harmonics

In order to calculate the various spherical harmonics $Y_l^m(\theta, \varphi)$, we shall use the method indicated in Chapter VI (§ D-1-c); starting with the expression for $Y_l^l(\theta, \varphi)$, we shall use the operator L_- to obtain by iteration the spherical harmonics corresponding to the same value of l and the $(2l+1)$ values of m associated with it. Recall that the operators L_+ and L_- act only on the angular dependence of a wave function and can be written:

$$L_\pm = \hbar\, \mathrm{e}^{\pm i\varphi} \left[\pm \frac{\partial}{\partial \theta} + i\, \cot\theta \frac{\partial}{\partial \varphi} \right] \tag{1}$$

1-a. Determination of $Y_l^l(\theta, \varphi)$

We have seen (§ D-1-c of Chapter VI) that $Y_l^l(\theta, \varphi)$ can be calculated from the equation:

$$L_+\, Y_l^l(\theta, \varphi) = 0 \tag{2}$$

and from the fact that:

$$Y_l^l(\theta, \varphi) = F_l^l(\theta)\ \mathrm{e}^{il\varphi} \tag{3}$$

Thus we obtained:

$$Y_l^l(\theta, \varphi) = c_l (\sin\theta)^l\ \mathrm{e}^{il\varphi} \tag{4}$$

where c_l is an arbitrary constant.

First, let us determine the absolute value of c_l by requiring $Y_l^l(\theta, \varphi)$ to be normalized with respect to the angular variables θ and φ:

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \,\mathrm{d}\theta \, |Y_l^l(\theta,\varphi)|^2 = |c_l|^2 \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \,\mathrm{d}\theta \, (\sin\theta)^{2l} = 1 \tag{5}$$

We obtain:

$$|c_l|^2 = 1/(2\pi I_l) \tag{6}$$

where I_l is given by:

$$I_l = \int_0^{\pi} \sin\theta \,\mathrm{d}\theta \, (\sin\theta)^{2l} = \int_{-1}^{+1} \mathrm{d}u \, (1-u^2)^l \tag{7}$$

(setting $u = \cos\theta$). I_l can easily be calculated by recurrence since:

$$I_l = \int_{-1}^{+1} \mathrm{d}u \, (1-u^2)(1-u^2)^{l-1} = I_{l-1} - \int_{-1}^{+1} \mathrm{d}u \, u^2 (l-u^2)^{l-1} \tag{8}$$

An integration by parts of the last integral yields:

$$I_l = I_{l-1} - \frac{1}{2l} I_l \tag{9}$$

We therefore have:

$$I_l = \frac{2l}{2l+1} I_{l-1} \tag{10}$$

with:

$$I_0 = \int_{-1}^{+1} \mathrm{d}u = 2 \tag{11}$$

From this, we can immediately derive the value of I_l:

$$I_l = \frac{(2l)!!}{(2l+1)!!} I_0 = \frac{2^{2l+1}(l!)^2}{(2l+1)!} \tag{12}$$

$Y_l^l(\theta, \varphi)$ is then normalized if:

$$|c_l| = \frac{1}{2^l l!} \sqrt{\frac{(2l+1)!}{4\pi}} \tag{13}$$

In order to define c_l completely, we must choose its phase. It is customary to choose:

$$c_l = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)!}{4\pi}} \tag{14}$$

We shall see later that, with this convention, $Y_l^0(\theta)$ (which is independent of φ) has a real positive value for $\theta = 0$.

1-b. General expression for $Y_l^m(\theta,\varphi)$

We shall obtain the other spherical harmonics $Y_l^m(\theta,\varphi)$ by successive application of the operator L_- to the $Y_l^l(\theta,\varphi)$ we have just determined. First, we shall prove a convenient formula that will enable us to simplify the calculations.

α. The action of $(L_\pm)^p$ on a function of the form $e^{in\varphi}F(\theta)$

The action of the operators L_+ and L_- on a function of the form $e^{in\varphi}F(\theta)$ (where n is any integer) is given by:

$$L_\pm[e^{in\varphi}F(\theta)] = \mp\hbar e^{i(n\pm1)\varphi}(\sin\theta)^{1\pm n}\frac{d}{d(\cos\theta)}[(\sin\theta)^{\mp n}\ F(\theta)] \tag{15}$$

More generally:

$$(L_\pm)^p[e^{in\varphi}F(\theta)] = (\mp\hbar)^p e^{i(n\pm p)\varphi}(\sin\theta)^{p\pm n}\frac{d^p}{d(\cos\theta)^p}[(\sin\theta)^{\mp n}\ F(\theta)] \tag{16}$$

First, let us prove formula (15). We know that:

$$\frac{d}{d(\cos\theta)} = \frac{d\theta}{d(\cos\theta)}\frac{d}{d\theta} = -\frac{1}{\sin\theta}\frac{d}{d\theta} \tag{17}$$

and therefore:

$$\begin{aligned}(\sin\theta)^{1\pm n}&\frac{d}{d(\cos\theta)}\left[(\sin\theta)^{\mp n}F(\theta)\right]\\ &= (\sin\theta)^{1\pm n}\left(-\frac{1}{\sin\theta}\right)\left[\mp n(\sin\theta)^{\mp n-1}\cos\theta\ F(\theta) + (\sin\theta)^{\mp n}\ \frac{dF(\theta)}{d\theta}\right]\\ &= -\left[\mp n\ \cot\theta\ F(\theta) + \frac{dF(\theta)}{d\theta}\right]\end{aligned} \tag{18}$$

Consequently:

$$\begin{aligned}\mp e^{i(n\pm1)\varphi}(\sin\theta)^{1\pm n}\frac{d}{d(\cos\theta)}[(\sin\theta)^{\mp n}F(\theta)] &= \left[-n\ \cot\theta \pm \frac{\partial}{\partial\theta}\right]e^{i(n\pm1)\varphi}F(\theta)\\ &= e^{\pm i\varphi}\left[\pm\frac{\partial}{\partial\theta} + i\ \cot\theta\frac{\partial}{\partial\varphi}\right]e^{in\varphi}F(\theta)\end{aligned} \tag{19}$$

We recognize expression (1) for the operators L_+ and L_-; relation (19) is therefore identical to (15).

Now, to establish formula (16), we can reason by recurrence, since, for $p=1$, (16) reduces to (15), which we have just proved. Let us therefore assume that relation (16) is true for $(p-1)$:

$$(L_\pm)^{p-1}[e^{in\varphi}F(\theta)] = (\mp\hbar)^{p-1}e^{i(n\pm p\mp1)\varphi}(\sin\theta)^{p-1\pm n}\times\frac{d^{p-1}}{d(\cos\theta)^{p-1}}[(\sin\theta)^{\mp n}F(\theta)] \tag{20}$$

and let us show that it is then also valid for p. To do so, we apply $L_\pm$ to both sides of (20); for the right-hand side, we can use formula (15), making the substitutions:

$$\begin{aligned}n &\Longrightarrow n\pm p\mp1\\ F(\theta) &\Longrightarrow (\sin\theta)^{p-1\pm n}\frac{d^{p-1}}{d(\cos\theta)^{p-1}}[(\sin\theta)^{\mp n}F(\theta)]\end{aligned} \tag{21}$$

We then obtain:

$$(L_\pm)^p[e^{in\varphi}F(\theta)] = (\mp\hbar)^p e^{i(n\pm p)\varphi}(\sin\theta)^{\pm n+p} \times \frac{d}{d(\cos\theta)}\left\{(\sin\theta)^{\mp n-p+1}(\sin\theta)^{p-1\pm n}\frac{d^{p-1}}{d(\cos\theta)^{p-1}}[(\sin\theta)^{\mp n}F(\theta)]\right\}$$
$$= (\mp\hbar)^p e^{i(n\pm p)\varphi}(\sin\theta)^{p\pm n}\frac{d^p}{d(\cos\theta)^p}[(\sin\theta)^{\mp n}F(\theta)] \quad (22)$$

Formula (16) is therefore proven by recurrence.

β. Calculation of $Y_l^m(\theta,\varphi)$ from $Y_l^l(\theta,\varphi)$

As we have already indicated (Chapter VI, § D-1-c-α), the spherical harmonics $Y_l^m(\theta,\varphi)$ must satisfy:

$$L_\pm Y_l^m(\theta,\varphi) = \hbar\sqrt{l(l+1)-m(m\pm1)}\; Y_l^{m\pm1}(\theta,\varphi)$$
$$= \hbar\sqrt{l(l\mp m)(l\pm m+1)}\; Y_l^{m\pm1}(\theta,\varphi) \quad (23)$$

These relations automatically insure that $Y_l^{m\pm1}$ is normalized if Y_l^m is. Also, they fix the relative phases of spherical harmonics corresponding to the same value of l and different values of m.

In particular, we can calculate $Y_l^m(\theta,\varphi)$ from $Y_l^l(\theta,\varphi)$ by using the operator L_- given by (1) and formula (23). Thus, we shall obtain directly a normalized function $Y_l^m(\theta,\varphi)$ whose phase will be determined by the convention used for $Y_l^l(\theta,\varphi)$ [formula (14)]. To go from $Y_l^l(\theta,\varphi)$ to $Y_l^m(\theta,\varphi)$, we must apply $(l-m)$ times the operator L_-; according to (23), we thus obtain:

$$(L_-)^{l-m}Y_l^l(\theta,\varphi) = (h)^{l-m}\sqrt{(2l)(1)\times(2l-1)(2)\times\ldots\times(l+m+1)(l-m)}\; Y_l^m(\theta,\varphi) \quad (24)$$

that is:

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{(l+m)!}{(2l)!(l-m)!}}\left(\frac{L_-}{\hbar}\right)^{l-m} Y_l^l(\theta,\varphi) \quad (25)$$

Finally, using expression (4) for $Y_l^l(\theta,\varphi)$ [where the coefficient c_l is given by (14)] and formula (16) (with $n=l$ and $p=l-m$), we can write (25) explicitly in the form:

$$Y_l^m(\theta,\varphi) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{2l+l}{4\pi}\frac{(l+m)!}{(l-m)!}}\; e^{im\varphi}(\sin\theta)^{-m}\frac{d^{l-m}}{d(\cos\theta)^{l-m}}(\sin\theta)^{2l} \quad (26)$$

γ. Calculation of $Y_l^m(\theta,\varphi)$ from $Y_l^{-l}(\theta,\varphi)$

In order to obtain expression (26), we started with the result of § 1-a. It is, of course, just as easy to calculate $Y_l^{-l}(\theta,\varphi)$ first and then use the operator L_+. The expression thus obtained for Y_l^m is different from (26), although the two are completely equivalent.

Let us therefore calculate $Y_l^{-l}(\theta,\varphi)$ from (26)[1]. Since:

$$(\sin\theta)^{2l} = (1-\cos^2\theta)^l \tag{27}$$

is a polynomial of degree $2l$ in $\cos\theta$, only its highest-order term contributes to $Y_l^{-l}(\theta,\varphi)$:

$$\frac{\mathrm{d}^{2l}}{\mathrm{d}(\cos\theta)^{2l}}\,(\sin\theta)^{2l} = (-1)^l\,(2l)! \tag{28}$$

We therefore immediately find that:

$$Y_l^{-l}(\theta,\varphi) = \frac{1}{2^l l!}\sqrt{\frac{(2l+l)!}{4\pi}}\;\mathrm{e}^{-il\varphi}(\sin\theta)^l \tag{29}$$

$Y_l^m(\theta,\varphi)$ can then be obtained by applying $(l+m)$ times the operator L_+ to $Y_l^{-l}(\theta,\varphi)$. Using (23) and (16), we finally arrive at:

$$Y_l^m(\theta,\varphi) = \frac{(-1)^{l+m}}{2^l l!}\;\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\;\mathrm{e}^{im\varphi}(\sin\theta)^m\frac{\mathrm{d}^{l+m}}{\mathrm{d}(\cos\theta)^{l+m}}\,(\sin\theta)^{2l} \tag{30}$$

1-c. Explicit expressions for $l = 0$, 1 and 2

General formulas (26) and (30) yield the expressions of the spherical harmonics for the first values of l:

$$Y_0^0 = \frac{1}{\sqrt{4\pi}} \tag{31}$$

$$\begin{cases} Y_1^{\pm1}(\theta,\varphi) = \mp\sqrt{\dfrac{3}{8\pi}}\;\sin\theta\;\mathrm{e}^{\pm i\varphi} \\ Y_1^0(\theta,\varphi) = \sqrt{\dfrac{3}{4\pi}}\;\cos\theta \end{cases} \tag{32}$$

$$\begin{cases} Y_2^{\pm2}(\theta,\varphi) = \sqrt{\dfrac{15}{32\pi}}\;\sin^2\theta\;\mathrm{e}^{\pm2i\varphi} \\ Y_2^{\pm1}(\theta,\varphi) = \mp\sqrt{\dfrac{15}{8\pi}}\;\sin\theta\cos\theta\;\mathrm{e}^{\pm i\varphi} \\ Y_2^0(\theta,\varphi) = \sqrt{\dfrac{5}{16\pi}}\;(3\cos^2\theta-1) \end{cases} \tag{33}$$

[1] We could obviously calculate Y_l^{-l} from the equation:

$L_- Y_l^{-l}(\theta,\varphi) = 0$

However, its phase would then remain arbitrary. By using (26), we shall determine $Y_l^{-l}(\theta,\varphi)$ completely, and its phase will be a consequence of the convention chosen in § 1-a.

2. Properties of spherical harmonics

2-a. Recurrence relations

By their very construction, the spherical harmonics satisfy relations (23); that is, using (1):

$$e^{\pm i\varphi}\left[\pm\frac{\partial}{\partial\theta}-m\cot\theta\right]Y_l^m(\theta,\varphi)=\sqrt{l(l+1)-m(m\pm1)}\;Y_l^{m\pm1}(\theta,\varphi) \tag{34}$$

Also note the following formula, which is often useful:

$$\cos\theta\;Y_l^m(\theta,\varphi)=\sqrt{\frac{(l+m+1)(l-m+1)}{(2l+1)(2l+3)}}\;Y_{l+1}^m(\theta,\varphi) + \sqrt{\frac{(l+m)(l-m)}{(2l+1)(2l-1)}}\;Y_{l-1}^m(\theta,\varphi) \tag{35}$$

Here is an outline of its proof. According to (25):

$$\cos\theta\;Y_l^m=\sqrt{\frac{(l+m)!}{(2l)!(l-m)!}}\cos\theta\left(\frac{L_-}{\hbar}\right)^{l-m}Y_l^l(\theta,\varphi) \tag{36}$$

Now, using expression (1) for L_-, it is easy to verify that:

$$[L_-\ ,\ \cos\theta]=\hbar e^{-i\varphi}\sin\theta \tag{37}$$

and:

$$[L_-\ ,\ e^{-i\varphi}\sin\theta]=0 \tag{38}$$

Using a recurrence argument, we can then calculate the commutator of $(L_-/\hbar)^k$ and $\cos\theta$, since, if we assume that:

$$\left[\left(\frac{L_-}{\hbar}\right)^{k-1},\ \cos\theta\right]=(k-1)e^{i\varphi}\sin\theta\left(\frac{L_-}{\hbar}\right)^{k-2} \tag{39}$$

we obtain:

$$\begin{aligned}\left[\left(\frac{L_-}{\hbar}\right)^{k},\ \cos\theta\right]&=\left(\frac{L_-}{\hbar}\right)^{k-1}\left[\frac{L_-}{\hbar}\ ,\ \cos\theta\right]+\left[\left(\frac{L_-}{\hbar}\right)^{k-1},\ \cos\theta\right]\frac{L_-}{\hbar}\\&=\left(\frac{L_-}{\hbar}\right)^{k-1}e^{-i\varphi}\sin\theta+(k-1)e^{-i\varphi}\sin\theta\left(\frac{L_-}{\hbar}\right)^{k-1}\end{aligned} \tag{40}$$

that is:

$$\left[\left(\frac{L_-}{\hbar}\right)^{k},\ \cos\theta\right]=k\;e^{-i\varphi}\sin\theta\left(\frac{L_-}{\hbar}\right)^{k-1}=k\;\left(\frac{L_-}{\hbar}\right)^{k-1}e^{-i\varphi}\sin\theta \tag{41}$$

This relation has therefore been established by recurrence. We can use it to write (36) in the form:

$$\cos\theta\;Y_l^m=\sqrt{\frac{(l+m)!}{(2l)!(l-m)!}}\left[\left(\frac{L_-}{\hbar}\right)^{l-m}\cos\theta\;Y_l^l-(l-m)\left(\frac{L_-}{\hbar}\right)^{l-m-1}e^{-i\varphi}\sin\theta\;Y_l^l\right] \tag{42}$$

Using (4) and (14), we can easily show that:

$$\mathrm{e}^{-i\varphi} \sin\theta \; Y_l^l = -\sqrt{\frac{2l+1}{2l}} \; (1-\cos^2\theta) \; Y_{l-1}^{l-1} \tag{43}$$

If we then calculate the explicit expressions for Y_{l+1}^l and Y_{l+1}^{l-1} from the general expression (26), we find that:

$$\cos\theta \; Y_l^l = \frac{1}{\sqrt{2l+3}} \; Y_{l+1}^l \tag{44a}$$

$$\cos^2\theta \; Y_{l-1}^{l-1} = \frac{2}{2l+1}\sqrt{\frac{l}{2l+3}} \; Y_{l+1}^{l-1} + \frac{1}{2l+1} \; Y_{l-1}^{l-1} \tag{44b}$$

Substituting relations (43) and (44) into (42) and using (23), we obtain (35).

2-b. Orthonormalization and closure relations

Because of the way we constructed them, the spherical harmonics constitute a set of normalized functions; they are also orthogonal, since they are eigenfunctions of the Hermitian operators $\mathbf{L}^2$ and L_z with different eigenvalues. The corresponding orthonormalization relation is:

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \, \mathrm{d}\theta \; Y_l^{m*}(\theta,\varphi) \; Y_{l'}^{m'}(\theta,\varphi) = \delta_{ll'} \; \delta_{mm'} \tag{45}$$

It can be shown (here, we shall assume) that any square-integrable function of θ and φ can be expanded in one and only one way on the spherical harmonics:

$$f(\theta,\varphi) = \sum_l \sum_m c_{l,m} \; Y_l^m(\theta,\varphi) \tag{46}$$

with:

$$c_{l,m} = \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \, d\theta \; Y_l^{m*}(\theta,\varphi) f(\theta,\varphi) \tag{47}$$

The set of spherical harmonics therefore constitutes an orthonormal basis of the space of square-integrable functions of θ and φ. This can be expressed by the closure relation:

$$\sum_l \sum_m Y_l^m(\theta,\varphi) \; Y_l^{m*}(\theta',\varphi') = \delta(\cos\theta - \cos\theta') \; \delta(\varphi - \varphi') \tag{48}$$

2-c. Parity

The parity operation on a function defined in ordinary space (*cf.* Complement $\mathrm{F_{II}}$) consists of replacing in this function the coordinates of any point in space by those of the point symmetrical to it with respect to the origin of the reference frame:

$$\mathbf{r} \Longrightarrow -\mathbf{r} \tag{49}$$

In spherical coordinates, this operation is expressed by the substitutions (Fig. 2 of Chapter VI):

$$\begin{aligned} r &\Longrightarrow r \\ \theta &\Longrightarrow \pi - \theta \\ \varphi &\Longrightarrow \pi + \varphi \end{aligned} \tag{50}$$

Consequently, if we are using a standard basis for the wave function space of a spinless particle (§ D-1-d of Chapter VI), the radial part of the basis functions $\psi_{k,l,m}(r)$ is unchanged by the parity operation. The only transformation is that of the spherical harmonics, which we shall now describe.

First, note that in the substitution of (50):

$$\begin{aligned} \sin\theta &\Longrightarrow \sin\theta \\ \cos\theta &\Longrightarrow -\cos\theta \\ \mathrm{e}^{im\varphi} &\Longrightarrow (-1)^m \mathrm{e}^{im\varphi} \end{aligned} \tag{51}$$

Under these conditions, the function $Y_l^l(\theta, \varphi)$ which we calculated in § 1-a is transformed into:

$$Y_l^l(\pi - \theta, \pi + \varphi) = (-1)^l \; Y_l^l(\theta, \varphi) \tag{52}$$

Moreover:

$$\begin{aligned} \frac{\partial}{\partial\theta} &\Longrightarrow -\frac{\partial}{\partial\theta} \\ \frac{\partial}{\partial\varphi} &\Longrightarrow \frac{\partial}{\partial\varphi} \end{aligned} \tag{53}$$

Relations (51) and (53) show that the operators L_+ and L_- [formulas (1)] remain unchanged [which means that L_+ and L_- are even operators, in the sense defined in Complement F_{II} (§ 2-a)]. Consequently, according to result (52) and formula (25), which enables us to calculate $Y_l^m(\theta, \varphi)$:

$$Y_l^m(\pi - \theta, \pi + \varphi) = (-1)^l \; Y_l^m(\theta, \varphi) \tag{54}$$

The spherical harmonics are therefore functions whose parity is well-defined and independent of m: they are even for l even and odd for l odd.

2-d. Complex conjugation

Because of their φ-dependence, the spherical harmonics are complex-valued functions. It can be seen directly, by comparing (26) and (30), that:

$$[Y_l^m(\theta, \varphi)]^* = (-1)^m \; Y_l^{-m}(\theta, \varphi) \tag{55}$$

2-e. Relation between the spherical harmonics and the Legendre polynomials and associated Legendre functions

The θ-dependence of the spherical harmonics resides in functions known as Legendre polynomials and associated Legendre functions. We shall neither prove nor even enumerate all the properties of these functions, but shall simply indicate their relation with the spherical harmonics.

α. *$Y_l^0(\theta)$ is proportional to a Legendre polynomial*

For $m = 0$, formulas (26) and (30) yield:

$$Y_l^0(\theta) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{2l+1}{4\pi}}\frac{\mathrm{d}^l}{\mathrm{d}(\cos\theta)^l}(\sin\theta)^{2l} \tag{56}$$

which can be written in the form:

$$Y_l^0(\theta) = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta) \tag{57}$$

setting:

$$P_l(u) = \frac{(-1)^l}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}u^l}(1-u^2)^l \tag{58}$$

According to its definition (58), $P_l(u)$ is an lth order polynomial in u of parity[2] $(-1)^l$:

$$P_l(-u) = (-1)^l\, P_l(u) \tag{59}$$

$P_l(u)$ is the lth order Legendre polynomial. It is easy to show that it has l zeros in the interval $[-1, +1]$, and that the numerical coefficient in (58) insures that:

$$P_l(1) = 1 \tag{60}$$

It can also be proven that the Legendre polynomials form a set of orthogonal functions:

$$\int_{-1}^{+1}\mathrm{d}u\ P_l(u)\ P_{l'}(u) = \int_0^\pi \sin\theta\,\mathrm{d}\theta\ P_l(\cos\theta)\ P_{l'}(\cos\theta) = \frac{2}{2l+1}\ \delta_{ll'} \tag{61}$$

on which can be expanded functions of θ alone:

$$f(\theta) = \sum_l c_l\ P_l(\cos\theta) \tag{62}$$

with:

$$c_l = \frac{2l+1}{2}\int_0^\pi \sin\theta\,\mathrm{d}\theta\ P_l(\cos\theta)\ f(\theta) \tag{63}$$

Comment:

According to (57) and (60):

$$Y_l^0(0) = \sqrt{\frac{2l+1}{4\pi}} \tag{64}$$

As we pointed out in § 1-a, the phase convention chosen for $Y_l^l(\theta,\varphi)$ gives a real positive value to $Y_l^0(0)$.

[2] Parity with respect to the variable u. Note, however, that the parity operation in space [formulas (50)] amounts to changing $\cos\theta$ to $-\cos\theta$; property (59) can be expressed by:

$Y_l^0(\pi-\theta) = (-1)^l\ Y_l^0(\theta)$

which is a special case of (54).

β. $Y_l^m(\theta,\varphi)$ is proportional to an associated Legendre function

For m positive, $Y_l^m(\theta,\varphi)$ can be obtained by applying L_+ to $Y_l^0(\theta)$; using (23):

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{(1-m)!}{(1+m)!}} \left(\frac{L_+}{\hbar}\right)^m Y_l^0(\theta) \qquad (m \geq 0) \tag{65}$$

Using formulas (1) and (16), we then find:

$$Y_l^m(\theta,\varphi) = (-1)^m \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(1+m)!}}\, P_l^m(\cos\theta)\, \mathrm{e}^{im\varphi} \qquad (m \geq 0) \tag{66}$$

where P_l^m is an associated Legendre function, defined by:

$$P_l^m(u) = \sqrt{(1-u^2)^m}\, \frac{\mathrm{d}^m}{\mathrm{d}u^m} P_l(u) \qquad (-1 \leq u \leq +1) \tag{67}$$

$P_l^m(u)$ is the product of $\sqrt{(1-u^2)^m}$ and a polynomial of degree $(l-m)$ and parity $(-1)^{l-m}$; $P_l^0(u)$ is the Legendre polynomial $P_l(u)$. The set of $P_l^m(u)$ for fixed m constitutes an orthogonal system of functions:

$$\int_{-1}^{+1} \mathrm{d}u\; P_l^m(u)\; P_{l'}^m(u) = \int_0^\pi \sin\theta\, \mathrm{d}\theta\; P_l^m(\cos\theta)\; P_{l'}^m(\cos\theta) = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\,\delta_{ll'} \tag{68}$$

on which can be expanded functions of θ alone.

Formula (66) is valid for m positive (or zero); for negative m, it suffices to use relation (55) to obtain:

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l+m)!}{(1-m)!}}\, P_l^{-m}(\cos\theta)\, \mathrm{e}^{im\varphi} \qquad (m < 0) \tag{69}$$

γ. Spherical harmonic addition theorem

Consider two arbitrary directions in space, Ou' and Ou'', defined respectively by the polar angles (θ',φ') and (θ'',φ''), and call the angle between them α. The following relation can be proven:

$$\frac{2l+1}{4\pi}\, P_l(\cos\alpha) = \sum_{m=-l}^{+l} (-1)^m\; Y_l^m(\theta',\varphi')\; Y_l^{-m}(\theta'',\varphi'') \tag{70}$$

(where P_l is the lth-order Legendre polynomial). It is known as the "spherical harmonics addition theorem".

We shall indicate the main steps of an elementary proof of relation (70). First of all, note that, if $\cos\alpha$ is expressed in terms of the polar angles (θ',φ') and (θ'',φ''), the left-hand side of (70) can be considered to be a function of θ' and φ'; it can therefore be expanded on the spherical harmonics $Y_{l'}^{m'}(\theta',\varphi')$. The coefficients of this expansion, which are, of course, functions of the other two variables, θ'' and φ'', can also be expanded on the spherical harmonics $Y_{l''}^{m''}(\theta'',\varphi'')$. We must therefore have:

$$\frac{2l+1}{4\pi}\, P_l(\cos\alpha) = \sum_{l',m'} \sum_{l'',m''} c_{l',m';l'',m''}\; Y_{l'}^{m'}(\theta',\varphi')\; Y_{l''}^{m''}(\theta'',\varphi'') \tag{71}$$

where the problem is to calculate the coefficients $c_{l',m';l'',m''}$. They can be obtained by the following process:

(i) in the first place, these coefficients are different from zero only for:

$$l' = l'' = l \tag{72}$$

To show this, first fix the direction Ou'' ; $P_l(\cos\alpha)$ then depends only on θ' and φ'. If the Oz axis is chosen along Ou'', $\cos\alpha = \cos\theta'$ and $P_l(\cos\alpha)$ is proportional to $Y_l^0(\theta')$ [*cf.* relation (57)]. To generalize to the case in which the direction of Ou'' is arbitrary, we perform a rotation which takes Oz onto this direction: $\cos\alpha$ remains unchanged, as does $P_l(\cos\alpha)$. Since the rotation operators (Complement B_{VI}, § 3-c-γ) commute with $\mathbf{L}^2$, the transform of $Y_l^0(\theta')$ remains an eigenfunction of $\mathbf{L}^2$ with the eigenvalue $l(l+1)\hbar^2$, that is, a linear combination of the spherical harmonics $Y_{l'}^{m'}(\theta', \varphi')$; we therefore have $l' = l$. Similarly, it can be established that $l'' = l$.

(ii) under a rotation of both directions Ou' and Ou'' through an angle β about Oz, the angle α is not changed, and neither are θ' and θ'', while φ' and φ'' become $\varphi' + \beta$ and $\varphi'' + \beta$. The left-hand side of (71) therefore does not change in value, and each term of the right-hand side is multiplied by $e^{i(m'+m'')\beta}$. Consequently, the only non-zero coefficients in the sum of the right-hand side are those which satisfy:

$$m' + m'' = 0 \tag{73}$$

(iii) combining results (72) and (73), we see that formula (71) can be written in the form:

$$\frac{2l+1}{4\pi}\ P_l(\cos\alpha) = \sum_{m=-l}^{+l} (-1)^m c_m\ Y_l^m(\theta', \varphi')\ Y_l^{-m}(\theta'', \varphi'') \tag{74}$$

If we set $\theta' = \theta''$ and $\varphi' = \varphi''$, we obtain, according to (60):

$$\frac{2l+1}{4\pi} = \sum_{m=-l}^{+l} (-1)^m c_m\ Y_l^m(\theta', \varphi')\ Y_l^{-m}(\theta', \varphi') \tag{75}$$

Since $(-1)^m Y_l^{-m}$ is simply Y_l^{m*}, the integration of (75) with respect to $d\Omega' = \sin\theta' d\theta' d\varphi'$ yields, with the orthonormalization equation (45):

$$2l + 1 = \sum_{m=-l}^{+l} c_m \tag{76}$$

We now take the square of the modulus of both sides of (74) and integrate over $d\Omega'$ and $d\Omega''$. Using relation (45), it is easy to see that the right-hand side yields $\sum_{m=-l}^{+l} |c_m|^2$. As far as the left-hand side is concerned, we can again take advantage of the invariance of the angle α with respect to a rotation in order to show that $\int d\Omega' |P_l(\cos\alpha)|^2$ is actually independent of (θ'', φ''). If we then choose Ou'' along Oz to evaluate this integral, we find, according to relation (61):

$$\int d\Omega'\ |P_l(\cos\alpha)|^2 = \int d\Omega\ |P_l(\cos\alpha)|^2 = 2\pi \times \frac{2}{2l+1} \tag{77}$$

Integrating over $d\Omega''$, we find a second relation between the coefficients c_m:

$$2l + 1 = \sum_{m=-l}^{+l} |c_m|^2 \tag{78}$$

(*iv*) equations (76) and (78) suffice for the determination of the $(2l+1)$ coefficients c_m: they are all equal to 1. To show this, consider, in a normed $(2l+1)$-dimensional vector space, the vector $\mathbf{X}$ of components $x_m = c_m/\sqrt{2l+1}$ and the vector $\mathbf{Y}$ of components $y_{\mathbf{m}} = 1/\sqrt{2l+1}$. The Schwartz inequality indicates that:

$$(\mathbf{X}^* \cdot \mathbf{X})(\mathbf{Y}^* \cdot \mathbf{Y}) \geq |\mathbf{Y}^* \cdot \mathbf{X}|^2 \tag{79}$$

where there is equality if and only if $\mathbf{X}$ and $\mathbf{Y}$ are proportional. Now, (76) and (78) show that this is the case: x_m and c_m are therefore independent of m, as is y_m, and we have, necessarily, $c_m = 1$. This concludes the proof of formula (70).

References

Messiah (1.17), App. B, § IV; Arfken (10.4), Chap. 12; Edmonds (2.21), Table 1; Butkov (10.8), Chap. 9, §§ 5 and 8; Whittaker and Watson (10.12), Chap. XV; Bateman (10.39), Chap. III; Bass (10.1), vol. I, § 17-7.

Complement B_{VI}

Angular momentum and rotations

1 Introduction 717

2 Brief study of geometrical rotations $\mathscr{R}$ 718

2-a Definition. Parametrization 718

2-b Infinitesimal rotations 719

3 Rotation operators in state space. Example: a spinless particle 720

3-a Existence and definition of rotation operators 720

3-b Properties of rotation operators R 721

3-c Expressing rotation operators in terms of angular momentum observables 723

4 Rotation operators in the state space of an arbitrary system 727

4-a System of several spinless particles 727

4-b An arbitrary system 729

5 Rotation of observables 730

5-a General transformation law 730

5-b Scalar observables 732

5-c Vector observables 732

6 Rotation invariance 734

6-a Invariance of physical laws 734

6-b Consequence: conservation of angular momentum 736

6-c Applications 737

1. Introduction

We indicated in Chapter VI (§ B-2) that the commutation relations between the components of an angular momentum are actually the expression of the geometrical properties of rotations in ordinary three-dimensional space. This is what we intend to show in this complement, where we investigate the relation between rotations and angular momentum operators.

Consider a physical system (S) whose quantum mechanical state, at a given time, is characterized by the ket $|\psi\rangle$ of the state space $\mathcal{E}$. We perform a rotation $\mathscr{R}$ on this system; in this new position, the state of the system is described by a ket $|\psi'\rangle$ which is different from $|\psi\rangle$. Given the geometrical transformation $\mathscr{R}$, the problem is to determine $|\psi'\rangle$ from $|\psi\rangle$. We shall see that it has the following solution: with every geometrical rotation $\mathscr{R}$ can be associated a linear operator R acting in the state space $\mathcal{E}$ such that:

$$|\psi'\rangle = R|\psi\rangle \tag{1}$$

Let us immediately stress the necessity of distinguishing between the geometrical rotation $\mathscr{R}$, which operates in ordinary space, and its "image" R, which acts in the state space:

$$\mathscr{R} \Longrightarrow R \tag{2}$$

We shall begin (§ 2) by reviewing the principal properties of geometrical rotations $\mathscr{R}$. We shall not embark upon a detailed study of them; rather, we shall simply note some results which will be useful to us later. Then, in § 3, we shall use the example of a spinless particle to define the rotation operators R precisely, to study their most important properties, and to determine their relation to the angular momentum operators $\mathbf{L}$. We shall then be able to interpret the commutation relations amongst the components of the angular momentum $\mathbf{L}$ as the image, in the space $\mathcal{E}_{\mathbf{r}}$, of purely geometrical characteristics of rotations $\mathscr{R}$. We shall then generalize (§ 4) these concepts to arbitrary quantum mechanical systems. In § 5, we shall examine the behavior of the observables describing the physical quantities measurable in this system, upon rotation of the system. This will lead us to classify observables according to how they transform under a rotation (scalar, vector, tensor observables). Finally, in § 6, we shall briefly consider the problem of rotation invariance and indicate some important consequences of this invariance.

2. Brief study of geometrical rotations $\mathscr{R}$

2-a. Definition. Parametrization

A rotation $\mathscr{R}$ is a one-to-one transformation of three-dimensional space that conserves a point of this space, the angles and the distances, as well as the handedness of the reference frames[1]. We shall be concerned here with the set of rotations that conserve a given point O, which we shall choose as the origin of the reference frame. A rotation can then be characterized by the *axis of rotation* (given by its unit vector $\mathbf{u}$ or its polar angles θ and φ) and the angle of rotation $\alpha (0 \leq \alpha < 2\pi)$. To determine a rotation, three parameters are required; they can be chosen to be the components of the vector:

$$\boldsymbol{\alpha} = \alpha\, \mathbf{u} \tag{3}$$

whose absolute value is equal to the angle of rotation and whose direction defines the axis of rotation. Note that a rotation can also be characterized by three angles, called Euler angles. We shall denote by $\mathscr{R}_{\mathbf{u}}(\alpha)$ the geometrical rotation through an angle α about the axis defined by the unit vector $\mathbf{u}$.

The set of rotations $\mathscr{R}$ constitutes a group: the product of two rotations (that is, the transformation resulting from the successive application of these two rotations) is also a rotation; there exists an identity rotation (rotation through a zero angle about an arbitrary axis); for every rotation $\mathscr{R}_{\mathbf{u}}(\alpha)$ there is an inverse rotation, $\mathscr{R}_{-\mathbf{u}}(\alpha)$. The group of rotations is not commutative: in general, the product of two rotations depends on the order in which they are performed[2]:

$$\mathscr{R}_{\mathbf{u}}(\alpha)\, \mathscr{R}_{\mathbf{u}'}(\alpha') \neq \mathscr{R}_{\mathbf{u}'}(\alpha')\, \mathscr{R}_{\mathbf{u}}(\alpha) \tag{4}$$

[1] This last property is imposed in order to exclude reflections with respect to a point or a plane.

[2] When one writes $\mathscr{R}_2 \mathscr{R}_1$, this means that rotation $\mathscr{R}_1$ must be performed first, $\mathscr{R}_2$ being applied subsequently to the result obtained.

Recall, however, that two rotations performed about the same axis always commute:

$$\mathscr{R}_{\mathbf{u}}(\alpha)\,\mathscr{R}_{\mathbf{u}}(\alpha') = \mathscr{R}_{\mathbf{u}}(\alpha')\,\mathscr{R}_{\mathbf{u}}(\alpha) = \mathscr{R}_{\mathbf{u}}(\alpha+\alpha') \tag{5}$$

(if necessary, 2π is subtracted from $\alpha+\alpha'$, to keep it within the interval $[0, 2\pi]$).

2-b. Infinitesimal rotations

An infinitesimal rotation is defined as a rotation that is infinitesimally close to the identity rotation, that is, a rotation $\mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha)$ through an infinitesimal angle $\mathrm{d}\alpha$ about an arbitrary axis $\mathbf{u}$. It is easy to see that the transform of a vector $\mathbf{OM}$ under the infinitesimal rotation $\mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha)$ can be written, to first order in $\mathrm{d}\alpha$:

$$\mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha)\,\mathbf{OM} = \mathbf{OM} + \mathrm{d}\alpha\ \mathbf{u}\times\mathbf{OM} \tag{6}$$

Every finite rotation can be decomposed into an infinite number of infinitesimal rotations, since the angle of rotation can vary continuously and since, according to (5):

$$\mathscr{R}_{\mathbf{u}}(\alpha+\mathrm{d}\alpha) = \mathscr{R}_{\mathbf{u}}(\alpha)\,\mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha) = \mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha)\,\mathscr{R}_{\mathbf{u}}(\alpha) \tag{7}$$

where $\mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha)$ is an infinitesimal rotation. Thus, the study of the rotation group can be reduced to an examination of infinitesimal rotations[3].

Before ending this rapid survey of the properties of geometrical rotations, we note the following relation which will be useful to us later:

$$\mathscr{R}_{\mathbf{e}_y}(-\mathrm{d}\alpha')\,\mathscr{R}_{\mathbf{e}_x}(\mathrm{d}\alpha)\,\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha')\,\mathscr{R}_{\mathbf{e}_x}(-\mathrm{d}\alpha) = \mathscr{R}_{\mathbf{e}_z}(\mathrm{d}\alpha\,\mathrm{d}\alpha') \tag{8}$$

where $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$ denote the unit vectors of the three coordinate axes Ox, Oy and Oz respectively. If $\mathrm{d}\alpha$ and $\mathrm{d}\alpha'$ are first-order infinitesimal angles, this relation is correct to the second order. It describes, in a special case, the non-commutative structure of the rotation group.

To prove relation (8), let us apply its left-hand side to an arbitrary vector $\mathbf{OM}$. We use formula (6) to find the vector $\mathbf{OM}'$, the transform of $\mathbf{OM}$ under the succesive action of the four infinitesimal rotations. It can be seen immediately that if $\mathrm{d}\alpha$ is zero, the left-hand side of (8) reduces to the product $\mathscr{R}_{\mathbf{e}_y}(-\mathrm{d}\alpha')\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha')$, which is equal to the identity rotation [see (5)]; the vector $\mathbf{OM}'-\mathbf{OM}$ must therefore be proportional to $\mathrm{d}\alpha$. For an analogous reason, it must also be proportional to $\mathrm{d}\alpha'$. Consequently, the difference $\mathbf{OM}'-\mathbf{OM}$ is proportional to $\mathrm{d}\alpha\,\mathrm{d}\alpha'$.

Therefore, to calculate $\mathbf{OM}'$ to second order, we may restrict ourselves to first order in each of the two infinitesimal angles $\mathrm{d}\alpha$ and $\mathrm{d}\alpha'$. First of all, according to (6):

$$\mathscr{R}_{\mathbf{e}_x}(-\mathrm{d}\alpha)\ \mathbf{OM} = \mathbf{OM} - \mathrm{d}\alpha\ \mathbf{e}_x\times\mathbf{OM} \tag{9}$$

We must then apply $\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha')$ to this vector; this can be done by again using (6):

$$\begin{aligned}\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha')\ \mathscr{R}_{\mathbf{e}_x}(-\mathrm{d}\alpha)\ \mathbf{OM}&\\ =&(\mathbf{OM} - \mathrm{d}\alpha\ \mathbf{e}_x\times\mathbf{OM}) + \mathrm{d}\alpha'\ \mathbf{e}_y\times(\mathbf{OM} - \mathrm{d}\alpha\ \mathbf{e}_x\times\mathbf{OM})\\ =&\mathbf{OM} - \mathrm{d}\alpha\ \mathbf{e}_x\times\mathbf{OM} + \mathrm{d}\alpha'\ \mathbf{e}_y\times\mathbf{OM} - \mathrm{d}\alpha\mathrm{d}\alpha'\ \mathbf{e}_y\times(\mathbf{e}_x\times\mathbf{OM})\end{aligned} \tag{10}$$

[3]However, limiting ourselves to infinitesimal rotations, we lose sight of a "global" property of the finite rotation group: the fact that a rotation through an angle of 2π is the identity transformation. The rotation operators (see § 3) constructed from infinitesimal operators do not always have this global property. In certain cases (see Complement $\mathrm{A_{IX}}$), the operator associated with a 2π rotation is not the unit operator but its opposite.

The action of $\mathscr{R}_{\mathbf{e}_x}(\mathrm{d}\alpha)$ on the vector appearing on the right-hand side of (10) results in the addition of the following infinitesimal terms to this vector:

$$\mathrm{d}\alpha\ \mathbf{e}_x \times \mathbf{OM} + \mathrm{d}\alpha\,\mathrm{d}\alpha'\ \mathbf{e}_x \times (\mathbf{e}_y \times \mathbf{OM}) \tag{11}$$

obtained by the vector multiplication of the right-hand side of (10) by $\mathrm{d}\alpha\ \mathbf{e}_x$, with only the first order terms in $\mathrm{d}\alpha$ being retained. Therefore:

$$\begin{aligned}\mathscr{R}_{\mathbf{e}_x}(\mathrm{d}\alpha)\,\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha')\mathscr{R}_{\mathbf{e}_x}(-\mathrm{d}\alpha)\ \mathbf{OM}&\\ = \mathbf{OM} + \mathrm{d}\alpha'\ \mathbf{e}_y \times \mathbf{OM} + \mathrm{d}\alpha\,\mathrm{d}\alpha'[\mathbf{e}_x \times (\mathbf{e}_y \times \mathbf{OM}) - \mathbf{e}_y \times (\mathbf{e}_x \times \mathbf{OM})]&\end{aligned} \tag{12}$$

Finally, $\mathbf{OM}'$ is equal to the sum of the vector just obtained and its vector product by $-\mathrm{d}\alpha'\mathbf{e}_y$. To first order in $\mathrm{d}\alpha'$, this vector product can be written simply:

$$-\mathrm{d}\alpha'\ \mathbf{e}_y \times \mathbf{OM} \tag{13}$$

which means that:

$$\begin{aligned}\mathscr{R}_{\mathbf{e}_y}(-\mathrm{d}\alpha')\,\mathscr{R}_{\mathbf{e}_x}(\mathrm{d}\alpha)\,\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha')\,\mathscr{R}_{\mathbf{e}_x}(-\mathrm{d}\alpha)\ \mathbf{OM}&\\ = \mathbf{OM} + \mathrm{d}\alpha\,\mathrm{d}\alpha'[\mathbf{e}_x \times (\mathbf{e}_y \times \mathbf{OM}) - \mathbf{e}_y \times (\mathbf{e}_x \times \mathbf{OM})]&\end{aligned} \tag{14}$$

It is then easy to transform the double vector products; we find:

$$\begin{aligned}\mathscr{R}_{\mathbf{e}_y}(-\mathrm{d}\alpha')\,\mathscr{R}_{\mathbf{e}_x}(\mathrm{d}\alpha)\,\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha')\,\mathscr{R}_{\mathbf{e}_x}(-\mathrm{d}\alpha)\ \mathbf{OM} &= \mathbf{OM} + \mathrm{d}\alpha\,\mathrm{d}\alpha'\ \mathbf{e}_z \times \mathbf{OM}\\ &= \mathscr{R}_{\mathbf{e}_z}(\mathrm{d}\alpha\,\mathrm{d}\alpha')\ \mathbf{OM}\end{aligned} \tag{15}$$

Since this relation is true for any vector $\mathbf{OM}$, expression (8) is verified.

3. Rotation operators in state space. Example: a spinless particle

In this section, we consider a physical system composed of a single (spinless) particle in three-dimensional space.

3-a. Existence and definition of rotation operators

At a given time, the quantum mechanical state of the particle is characterized, in the state space $\mathcal{E}_{\mathbf{r}}$, by the ket $|\psi\rangle$ with that is associated the wave function $\psi(\mathbf{r}) = \langle\mathbf{r}|\psi\rangle$. Let us perform a rotation $\mathscr{R}$ on this system which associates with the point $\mathbf{r}_0(x_0, y_0, z_0)$ of space the point $\mathbf{r}'_0(x'_0, y'_0, z'_0)$ such that:

$$\mathbf{r}'_0 = \mathscr{R}\mathbf{r}_0 \tag{16}$$

Let $|\psi'\rangle$ be the state vector of the system after rotation, and $\psi'(\mathbf{r}) = \langle\mathbf{r}|\psi'\rangle$, the corresponding wave function. It is natural to assume that the value of the initial wave function $\psi(\mathbf{r})$ at the point $\mathbf{r}_0$ will be found, after rotation, to be the value of the final wave function $\psi'(\mathbf{r})$ at the point $\mathbf{r}'_0$ given by (16):

$$\psi'(\mathbf{r}'_0) = \psi(\mathbf{r}_0) \tag{17}$$

that is:

$$\psi'(\mathbf{r}'_0) = \psi(\mathscr{R}^{-1}\mathbf{r}'_0) \tag{18}$$

Since this equation is valid for any point $(\mathbf{r}'_0)$ in space, it can be written in the form:

$$\psi'(\mathbf{r}) = \psi(\mathscr{R}^{-1}\mathbf{r}) \tag{19}$$

By definition, the operator R in the state space $\mathcal{E}_\mathbf{r}$ associated with the geometrical rotation $\mathscr{R}$ being treated is the one that acts on the state $|\psi\rangle$ before rotation to yield the state $|\psi'\rangle$ after the rotation $\mathscr{R}$:

$$|\psi'\rangle = R|\psi\rangle \tag{20}$$

R is called a "rotation operator". Relation (19) characterizes its action in the $\{|\mathbf{r}\rangle\}$ representation:

$$\langle\mathbf{r}|R|\psi\rangle = \langle\mathscr{R}^{-1}\mathbf{r}|\psi\rangle \tag{21}$$

where $|\mathscr{R}^{-1}\mathbf{r}\rangle$ is the basis ket of this representation determined by the components of the vector $\mathscr{R}^{-1}\mathbf{r}$.

Comment:

If the state of the particle after rotation were $\mathrm{e}^{i\theta}|\psi'\rangle$ (where θ is an arbitrary real number) instead of $|\psi'\rangle$, its physical properties would not be modified. In other words, relation (17) could be replaced by:

$$\psi'(r'_0) = \mathrm{e}^{i\theta}\psi(r_0) \tag{22}$$

θ would obviously be independent of r_0, but could depend on the rotation $\mathscr{R}$. We shall not treat this difficulty here.

3-b. Properties of rotation operators R

α. R is a linear operator

This essential property of rotation operators follows from their very definition. If the state $|\psi\rangle$ before the rotation is a linear superposition of states, for example:

$$|\psi\rangle = \lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle \tag{23}$$

formula (21) indicates that:

$$\begin{aligned}\langle\mathbf{r}|R|\psi\rangle &= \lambda_1\langle\mathscr{R}^{-1}\mathbf{r}|\psi_1\rangle + \lambda_2\langle\mathscr{R}^{-1}\mathbf{r}|\psi_2\rangle \\ &= \lambda_1\langle\mathbf{r}|R|\psi_1\rangle + \lambda_2\langle\mathbf{r}|R|\psi_2\rangle\end{aligned} \tag{24}$$

Since this relation is true for any ket of the $\{|\mathbf{r}\rangle\}$ basis, we deduce that R is a linear operator:

$$R|\psi\rangle = R[\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle] = \lambda_1 R|\psi_1\rangle + \lambda_2 R|\psi_2\rangle \tag{25}$$

β. R is unitary

In formula (21), the ket $|\psi\rangle$ can be arbitrary. The action of the operator R on the bra $\langle \mathbf{r}|$ is therefore given by:

$$\langle \mathbf{r}|R = \langle \mathscr{R}^{-1}\mathbf{r}| \tag{26}$$

Taking the Hermitian conjugate of both sides of equation (26), we obtain:

$$R^{\dagger}|\mathbf{r}\rangle = |\mathscr{R}^{-1}\mathbf{r}\rangle \tag{27}$$

Moreover, if we recall that the ket $|\mathbf{r}\rangle$ represents a state in which the particle is perfectly localized at the point $\mathbf{r}$, we see that:

$$R|\mathbf{r}\rangle = |\mathscr{R}\mathbf{r}\rangle \tag{28}$$

This equation simply expresses the fact that if the particle was localized at the point $\mathbf{r}$ before the rotation, it will be localized at the point $\mathbf{r}' = \mathscr{R}\mathbf{r}$ after the rotation. To get (28) from (21), we choose a basis state $|\mathbf{r}_0\rangle$ for $|\psi\rangle$:

$$\langle \mathbf{r}|R|\mathbf{r}_0\rangle = \langle \mathscr{R}^{-1}\mathbf{r}|\mathbf{r}_0\rangle = \delta[(\mathscr{R}^{-1}\mathbf{r}) - \mathbf{r}_0] \tag{29}$$

where we have used the orthonormalization relation of the $|\mathbf{r}\rangle$ basis. Furthermore[4]:

$$\delta[(\mathscr{R}^{-1}\mathbf{r}) - \mathbf{r}_0] = \delta[\mathbf{r} - (\mathscr{R}\mathbf{r}_0)] \tag{30}$$

Substituting (30) into (29), we indeed find that:

$$\langle \mathbf{r}|R|r_0\rangle = \delta[\mathbf{r} - (\mathscr{R}\mathbf{r}_0)] = \langle \mathbf{r}|\mathscr{R}\mathbf{r}_0\rangle \tag{31}$$

that is, since $\{|\mathbf{r}\rangle\}$ is a basis of $\mathcal{E}_{\mathbf{r}}$:

$$R|\mathbf{r}_0\rangle = |\mathscr{R}\mathbf{r}_0\rangle \tag{32}$$

Starting with formulas (27) and (28), it is easy to show that:

$$RR^{\dagger} = R^{\dagger}R = 1 \tag{33}$$

since the action of $RR^{\dagger}$ or $R^{\dagger}R$ on any vector of the $\{|\mathbf{r}\rangle\}$ basis yields the same vector; for example:

$$RR^{\dagger}|\mathbf{r}\rangle = R|\mathscr{R}^{-1}\mathbf{r}\rangle = |\mathscr{R}\mathscr{R}^{-1}\mathbf{r}\rangle = |\mathbf{r}\rangle \tag{34}$$

The operator R is therefore unitary.

Comment:

The operator R therefore conserves the scalar product and the norm of vectors that it transforms:

$$\left.\begin{array}{l} |\psi'\rangle = R|\psi\rangle \\ |\varphi'\rangle = R|\varphi\rangle \end{array}\right\} \qquad \Longrightarrow \qquad \langle \varphi'|\psi'\rangle = \langle \varphi|\psi\rangle \tag{35}$$

This property is very important from the physical point of view since the probability amplitudes which yield physical predictions appear in the form of scalar products of two kets.

[4]Relation (30) can easily be established by using the definition of delta "functions" and the fact that a rotation conserves the infinitesimal volume element.

γ. The set of operators R constitutes a representation of the rotation group

We have pointed out (§ 2) that the geometrical rotations form a group; in particular, the product of two rotations $\mathscr{R}_1$ and $\mathscr{R}_2$ is always a rotation:

$$\mathscr{R}_2\mathscr{R}_1 = \mathscr{R}_3 \tag{36}$$

With the three geometrical rotations $\mathscr{R}_1$, $\mathscr{R}_2$ and $\mathscr{R}_3$ are associated, in the state space $\mathcal{E}_\mathbf{r}$, three rotation operators R_1, R_2 and R_3, respectively. If the three geometrical rotations satisfy relation (36), we shall show that the corresponding rotation operators are such that:

$$R_2R_1 = R_3 \tag{37}$$

(R_2R_1 is a product of operators of $\mathcal{E}_\mathbf{r}$ as defined in Chapter II, § B-3-a).

Consider a particle whose state is described by an arbitrary ket $|\mathbf{r}\rangle$ of the basis characterizing the $\{|\mathbf{r}\rangle\}$ representation. If we perform the rotation $\mathscr{R}_1$ on this particle, its state becomes:

$$R_1|\mathbf{r}\rangle = |\mathscr{R}_1\mathbf{r}\rangle \tag{38}$$

by definition of R_1. Now, we perform the rotation $\mathscr{R}_2$ on the new state we have just obtained; the state of the particle after this second rotation is, according to (38) and the definition of R_2:

$$R_2R_1|\mathbf{r}\rangle = R_2|\mathscr{R}_1\mathbf{r}\rangle = |\mathscr{R}_2\mathscr{R}_1\mathbf{r}\rangle \tag{39}$$

If we take (36) into account, we see that relation (39) is equivalent to:

$$R_2R_1|\mathbf{r}\rangle = |\mathscr{R}_3\mathbf{r}\rangle \tag{40}$$

Now, the operator R_3, associated with the rotation $\mathscr{R}_3$, is such that:

$$R_3|\mathbf{r}\rangle = |\mathscr{R}_3\mathbf{r}\rangle \tag{41}$$

Relation (37) is therefore proven, since the ket $|\mathbf{r}\rangle$ under consideration can be chosen in an arbitrary way from the kets of the $\{|\mathbf{r}\rangle\}$ basis.

To express the important result we have just established, one says that the correspondence $\mathscr{R} \Longrightarrow R$ between geometrical rotations and rotation operators conserves the group law, or that the set of operators R constitutes a "representation" of the rotation group. Of course, with the identity rotation, we associate the identity operator in $\mathcal{E}_\mathbf{r}$, and with the rotation $\mathscr{R}^{-1}$ (the inverse of a rotation $\mathscr{R}$), the operator R^{-1}, which is the inverse of the one corresponding to $\mathscr{R}$ (we showed in § 3-b-β that $R^{-1} = R^\dagger$).

3-c. Expression for rotation operators in terms of angular momentum observables

α. Infinitesimal rotation operators

First of all, let us consider an infinitesimal rotation about the Oz axis, $\mathscr{R}_{\mathbf{e}_z}(\mathrm{d}\alpha)$. If we apply it to a particle whose state is described by the wave function $\psi(\mathbf{r})$, we know

from (19) that the wave function $\psi'(\mathbf{r})$ associated with the state of the particle after rotation satisfies:

$$\psi'(\mathbf{r}) = \psi[\mathscr{R}_{\mathbf{e}_z}^{-1}(\mathrm{d}\alpha)\mathbf{r}] \tag{42}$$

But if (x, y, z) are the components of $\mathbf{r}$, those of $\mathscr{R}_{e_z}^{-1}(\mathrm{d}\alpha)\mathbf{r}$ can easily be calculated from (6):

$$\mathscr{R}_{\mathbf{e}_z}^{-1}(\mathrm{d}\alpha)\mathbf{r} = \mathscr{R}_{-\mathbf{e}_z}(\mathrm{d}\alpha)\mathbf{r} = (\mathbf{r} - \mathrm{d}\alpha\ \mathbf{e}_z \times \mathbf{r}) \qquad \begin{cases} x + y\ \mathrm{d}\alpha \\ y - x\ \mathrm{d}\alpha \\ z \end{cases} \tag{43}$$

Equation (42) can then be written in the form:

$$\psi'(x, y, z) = \psi(x + y\ \mathrm{d}\alpha, y - x\ \mathrm{d}\alpha, z) \tag{44}$$

which yields, to first order in $\mathrm{d}\alpha$:

$$\begin{aligned} \psi'(x, y, z) &= \psi(x, y, z) + \mathrm{d}\alpha\ \left[y\frac{\partial \psi}{\partial x} - x\frac{\partial \psi}{\partial y} \right] \\ &= \psi(x, y, z) - \mathrm{d}\alpha\ \left[x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right]\ \psi(x, y, z) \end{aligned} \tag{45}$$

Within the brackets, we recognize, to within a factor of $\hbar/i$, the expression in the $\{|\mathbf{r}\rangle\}$ representation for the operator $L_z = XP_y - YP_x$. We therefore obtain the result:

$$\psi'(\mathbf{r}) = \langle \mathbf{r}|\psi'\rangle = \langle \mathbf{r}| \left(1 - \frac{i}{\hbar}\,\mathrm{d}\alpha\, L_z \right) |\psi\rangle \tag{46}$$

Now, by definition of the operator $R_{e_z}(\mathrm{d}\alpha)$ associated with the rotation $\mathscr{R}_{e_z}(\mathrm{d}\alpha)$:

$$|\psi'\rangle = R_{\mathbf{e}_z}(\mathrm{d}\alpha)|\psi\rangle \tag{47}$$

Therefore, since the original state $|\psi\rangle$ is arbitrary, we finally find that:

$$\boxed{R_{\mathbf{e}_z}(\mathrm{d}\alpha) = 1 - \frac{i}{\hbar}\ \mathrm{d}\alpha\, L_z} \tag{48}$$

The preceding argument can easily be generalized to an infinitesimal rotation about an arbitrary axis. We therefore have, in general:

$$\boxed{R_{\mathbf{u}}(\mathrm{d}\alpha) = 1 - \frac{i}{\hbar}\ \mathrm{d}\alpha\ \mathbf{L} \cdot \mathbf{u}} \tag{49}$$

Comment:

(46) can also be quickly established by using the spherical coordinates (r, θ, φ), since L_z then corresponds to the differential operator $\frac{\hbar}{i}\frac{\partial}{\partial \varphi}$.

β. Interpretation of the commutation relations for the components of the angular momentum **L**

What then is the "image" in the state space $\mathcal{E}_{\mathbf{r}}$ of relation (8)? According to the results of § 3-b-γ and the expressions we have just obtained, this relation implies that, to first order with respect to each of the angles $\mathrm{d}\alpha$ and $\mathrm{d}\alpha'$:

$$\left[1+\frac{i}{\hbar}\mathrm{d}\alpha' L_y\right]\left[1-\frac{i}{\hbar}\mathrm{d}\alpha\, L_x\right]\left[1-\frac{i}{\hbar}\mathrm{d}\alpha' L_y\right]\left[1+\frac{i}{\hbar}\mathrm{d}\alpha\, L_x\right]=1-\frac{i}{\hbar}\,\mathrm{d}\alpha\,\mathrm{d}\alpha'\; L_z \tag{50}$$

Expanding the left-hand side and setting the coefficients of $\mathrm{d}\alpha\mathrm{d}\alpha'$ equal, we easily find that relation (50) reduces to:

$$[L_x\ ,\ L_y] \quad = i\hbar\, L_z \tag{51}$$

Of course, the two other commutation relations for the components of **L** can be found, by an analogous argument, from the formulas obtained from (8) by cyclic permutation of the vectors $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$.

Thus, *the commutation relations of the orbital angular momentum of a particle can be seen to be consequences of the non-commutative structure of the geometrical rotation group.*

γ. Finite rotation operators

Now, consider a rotation $\mathscr{R}_{\mathbf{e}_z}(\alpha)$ through an arbitrary angle α about the Oz axis. According to formula (7), the operator $R_{\mathbf{e}_z}(\alpha)$ associated with such a rotation must satisfy (again using the results of § 3-b-γ):

$$R_{\mathbf{e}_z}(\alpha+\mathrm{d}\alpha)=R_{\mathbf{e}_z}(\alpha)\,R_{\mathbf{e}_z}(\mathrm{d}\alpha) \tag{52}$$

where the two operators of the right-hand side commute. But we know the expression for $R_{\mathbf{e}_z}(\mathrm{d}\alpha)$, so we have:

$$R_{\mathbf{e}_z}(\alpha+\mathrm{d}\alpha)=R_{\mathbf{e}_z}(\alpha)\left[1-\frac{i}{\hbar}\,\mathrm{d}\alpha\, L_z\right] \tag{53}$$

that is:

$$R_{\mathbf{e}_z}(\alpha+\mathrm{d}\alpha)-R_{\mathbf{e}_z}(\alpha)=-\frac{i}{\hbar}\,\mathrm{d}\alpha\, R_{\mathbf{e}_z}(\alpha)\,L_z \tag{54}$$

Here again, $R_{\mathbf{e}_z}(\alpha)$ and L_z must commute. Although we are dealing with operators, the solution of equation (54) is formally the same as it would be if we were considering an ordinary function of the variable α:

$$\boxed{R_{\mathbf{e}_z}(\alpha)=\mathrm{e}^{-\frac{i}{\hbar}\,\alpha\, L_z}} \tag{55}$$

Indeed, if we recall (*cf.* Complement $\mathrm{B_{II}}$, § 4) that the exponential of an operator is defined by the corresponding power series expansion, it is easy to verify that expression (55) is the solution of equation (54). Moreover, the "integration constant" is equal to 1, since we know that:

$$R_{\mathbf{e}_z}(0)=1 \tag{56}$$

As in § 3-c-α above, it is easy to generalize this result to a finite rotation about an arbitrary axis:

$$\boxed{R_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\,\alpha\,\mathbf{L}\cdot\mathbf{u}}} \tag{57}$$

Comments:

(*i*) Formula (57) can be written explicitly in the form:

$$R_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\alpha(L_x u_x + L_y u_y + L_z u_z)} \tag{58}$$

where u_x, u_y and u_z are the components of the unit vector $\mathbf{u}$. Recall, however, that, since L_x, L_y and L_z do not commute:

$$R_{\mathbf{u}}(\alpha) \neq \mathrm{e}^{-\frac{i}{\hbar}\alpha L_x u_x}\,\mathrm{e}^{-\frac{i}{\hbar}\alpha L_y u_y}\,\mathrm{e}^{-\frac{i}{\hbar}\alpha L_z u_z} \tag{59}$$

(*ii*) It can be seen from expression (57) that the operator $R_{\mathbf{u}}(\alpha)$ is unitary. Since the components of $\mathbf{L}$ are Hermitian:

$$[R_{\mathbf{u}}(\alpha)]^\dagger = \mathrm{e}^{\frac{i}{\hbar}\,\alpha\,\mathbf{L}\cdot\mathbf{u}} \tag{60}$$

we have (as $\mathbf{L}\cdot\mathbf{u}$ obviously commutes with itself):

$$[R_{\mathbf{u}}(\alpha)]^\dagger\, R_{\mathbf{u}}(\alpha) = R_{\mathbf{u}}(\alpha)\,[R_{\mathbf{u}}(\alpha)]^\dagger = 1 \tag{61}$$

(*iii*) In the special case envisaged in this section, we find that:

$$R_{\mathbf{u}}(2\pi) = 1 \tag{62}$$

We shall confine ourselves to proving this result for the rotation through 2π about the Oz axis (the generalization of this proof involves no difficulties). To this end, consider an arbitrary ket $|\psi\rangle$, and expand it on a basis composed of eigenvectors of the observable L_z:

$$|\psi\rangle = \sum_{m,\tau} c_{m,\tau}\,|m,\tau\rangle \tag{63}$$

with:

$$L_z|m,\tau\rangle = m\hbar\,|m,\tau\rangle \tag{64}$$

(τ symbolizes the indices other than m that are necessary to specify the vectors of the basis used, such as a "standard" basis $|k,l,m\rangle$ – *cf.* § C-3 of Chapter VI). The action of $R_{\mathbf{e}_z}(\alpha)$ on $|\psi\rangle$ is then easy to obtain:

$$\begin{aligned} R_{\mathbf{e}_z}(\alpha)|\psi\rangle &= \sum_{m,\tau} c_{m,\tau}\;\mathrm{e}^{-\frac{i}{\hbar}\alpha L_z}|m,\tau\rangle \\ &= \sum_{m,\tau} c_{m,\tau}\;\mathrm{e}^{-i\alpha m}|m,\tau\rangle \end{aligned} \tag{65}$$

But we know that, for the orbital angular momentum of a particle, *m is always integral*. Consequently, when α attains the value 2π, all the factors $e^{-i\alpha m}$ become equal to 1, and:

$$R_{\mathbf{e}_z}(2\pi)|\psi\rangle = \sum_{m,\tau} c_{m,\tau}|m,\tau\rangle = |\psi\rangle \tag{66}$$

Since this relation is satisfied for all $|\psi\rangle$, we deduce that $R_{\mathbf{e}_z}(2\pi)$ is the identity operator.

The preceding argument clearly indicates that *formula* (62) *would not be valid if half-integral values of m were not excluded*. Indeed, we shall see in Complement A_{IX} that, for a spin 1/2, the operator associated with a rotation of 2π is equal to -1 and not 1; this result is related to the fact that we constructed the finite rotations from infinitesimal rotations (*cf.* footnote 3).

4. Rotation operators in the state space of an arbitrary system

We shall now generalize the concepts we introduced and the results we obtained for a special case (in § 3).

4-a. System of several spinless particles

First of all, the arguments of § 3 can be extended without difficulty to systems composed of several spinless particles. We shall quickly demonstrate this, choosing as an example a system of two spinless particles, (1) and (2).

The state space $\mathcal{E}$ of such a system is the tensor product of the state spaces $\mathcal{E}_{\mathbf{r}_1}$ and $\mathcal{E}_{\mathbf{r}_2}$ of the two particles:

$$\mathcal{E} = \mathcal{E}_{\mathbf{r}_1} \otimes \mathcal{E}_{\mathbf{r}_2} \tag{67}$$

We shall use the same notation as in § F-4-b of Chapter II. Starting from the position and momentum observables ($\mathbf{R}_1$ and $\mathbf{P}_1$ on the one hand, $\mathbf{R}_2$ and $\mathbf{P}_2$ on the other), we can define an orbital angular momentum for each of the particles:

$$\begin{aligned} \mathbf{L}_1 &= \mathbf{R}_1 \times \mathbf{P}_1 \\ \mathbf{L}_2 &= \mathbf{R}_2 \times \mathbf{P}_2 \end{aligned} \tag{68}$$

The components of $\mathbf{L}_1$, as well as those of $\mathbf{L}_2$, satisfy the commutation relations characteristic of angular momenta.

Consider a vector which is a tensor product of a vector of $\mathcal{E}_{\mathbf{r}_1}$ and a vector of $\mathcal{E}_{\mathbf{r}_2}$:

$$|\psi\rangle = |\varphi(1)\rangle \otimes |\chi(2)\rangle \tag{69}$$

$|\psi\rangle$ represents the state of the system formed by particle (1) in the state $|\varphi(1)\rangle$ and particle (2) in the state $|\chi(2)\rangle$. If we perform a rotation through an angle α about $\mathbf{u}$ on the two-particle system, the state of the system after the rotation corresponds to the two particles in the "rotated states" $|\varphi'(1)\rangle$ and $|\chi'(2)\rangle$ respectively:

$$|\psi'\rangle = \quad |\varphi'(1)\rangle \otimes |\chi'(2)\rangle = [R^1_{\mathbf{u}}(\alpha)|\varphi(1)\rangle] \otimes [R^2_{\mathbf{u}}(\alpha)|\chi(2)\rangle] \tag{70}$$

where $R^1_{\mathbf{u}}(\alpha)$ and $R^2_{\mathbf{u}}(\alpha)$ are the rotation operators in $\mathcal{E}_{\mathbf{r_1}}$ and $\mathcal{E}_{\mathbf{r_2}}$:

$$R^1_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\,\alpha\,\mathbf{L}_1\cdot\mathbf{u}} \tag{71a}$$

$$R^2_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\,\alpha\,\mathbf{L}_2\cdot\mathbf{u}} \tag{71b}$$

Relation (70) can also be written, by definition of the tensor product of two operators (Chap. II, § F-2-b):

$$|\psi'\rangle = [R^1_{\mathbf{u}}(\alpha)\otimes R^2_{\mathbf{u}}(\alpha)]|\varphi(1)\rangle\otimes|\chi(2)\rangle \tag{72}$$

Since every vector of $\mathcal{E}$ is a linear combination of vectors analogous to (69), the rotation transform $|\psi'\rangle$ of an arbitrary vector $|\psi\rangle$ of $\mathcal{E}$ is:

$$|\psi'\rangle = [R^1_{\mathbf{u}}(\alpha)\otimes R^2_{\mathbf{u}}(\alpha)]|\psi\rangle \tag{73}$$

Using formula (F-14) of Chapter II and the fact that $\mathbf{L}_1$ and $\mathbf{L}_2$ commute (they are operators relating to different particles), we obtain for the rotation operators in $\mathcal{E}$:

$$R^1_{\mathbf{u}}(\alpha)\otimes R^2_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\alpha\mathbf{L}_1\cdot\mathbf{u}}\mathrm{e}^{-\frac{i}{\hbar}\alpha\mathbf{L}_2\cdot\mathbf{u}} = \mathrm{e}^{-\frac{i}{\hbar}\alpha\mathbf{L}\cdot\mathbf{u}} \tag{74}$$

where:

$$\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2 \tag{75}$$

is the total angular momentum of the two-particle system. All the formulas of the preceding section therefore remain valid as long as $\mathbf{L}$ represents the *total angular momentum*.

Comments:

(*i*) $\mathbf{L}$ is an operator that acts in $\mathcal{E}$. In (75), $\mathbf{L}_1$ is, rigorously, the extension of the operator $\mathbf{L}_1$ acting in $\mathcal{E}_{\mathbf{r}_1}$ into $\mathcal{E}$ (an analogous comment could be made for $\mathbf{L}_2$). To simplify the notation, we shall not use different symbols for $\mathbf{L}_1$ and its extension into $\mathcal{E}$ (*cf.* Chap. II, § F-2-c).

(*ii*) We might consider performing a rotation on only one of the two particles, for example, particle (1). In the course of such a "partial rotation", a vector such as (69) transforms into:

$$[R^1_{\mathbf{u}}(\alpha)|\varphi(1)\rangle]\otimes|\chi(2)\rangle \tag{76}$$

where only the state of particle (1) is modified. As above, it can be shown that the effect of a rotation performed only on particle (1) on an arbitrary state $|\psi\rangle$ of $\mathcal{E}$ is described by the operator:

$$R^1_{\mathbf{u}}(\alpha)\otimes\mathbb{1}(2) = \mathrm{e}^{-\frac{i}{\hbar}\alpha\mathbf{L}_1\cdot\mathbf{u}} \tag{77}$$

where $\mathbb{1}$ is the unit operator in $\mathcal{E}_{\mathbf{r}_2}$ [in (77), $\mathbf{L}_1$ acts in $\mathcal{E}$].

4-b. An arbitrary system

The starting point of the arguments elaborated thus far is equation (19), which gives the transformation law of the state vector of the system in terms of that of its wave function. In the case of an arbitrary quantum mechanical system (which does not necessarily have a classical analogue), one cannot use the same method. For example, for a particle with spin, the operators X, Y and Z no longer form a C.S.C.O., and the state of the particle can no longer be defined by a wave function $\psi(x, y, z)$ (*cf.* Chap. IX). One must reason directly in the state space $\mathcal{E}$ of the system. Without going into detail, we shall assume here that an operator R acting in $\mathcal{E}$ can be associated with any geometrical rotation $\mathscr{R}$; if the system is initially in the state $|\psi\rangle$, the rotation $\mathscr{R}$ takes it into the state:

$$|\psi'\rangle = R|\psi\rangle \tag{78}$$

where the operator R is linear and unitary (*cf.* comment of § 3-b-β).

As far as the group law of the rotations $\mathscr{R}$ is concerned, it is conserved by the operators R, but only locally: the product of two geometrical rotations, at least one of which is infinitesimal, is represented in the state space $\mathcal{E}$ by the product of the corresponding operators R (which implies, in particular, that the "image" of a rotation through an angle equal to zero is the identity operator). However, the operator associated with a geometrical rotation through an angle 2π is not necessarily the identity operator [*cf.* comment (*iii*) of § 3-c-γ and Complement $\mathrm{A_{IX}}$].

Now, let us consider an infinitesimal rotation $\mathscr{R}_{\mathbf{e}_z}(\mathrm{d}\alpha)$ about the Oz axis. Since the group law is conserved for infinitesimal rotations, the operator $R_{\mathbf{e}_z}(\mathrm{d}\alpha)$ is necessarily of the form:

$$R_{\mathbf{e}_z}(\mathrm{d}\alpha) = 1 - \frac{i}{\hbar}\,\mathrm{d}\alpha\, J_z \tag{79}$$

where J_z is a Hermitian operator since $R_{\mathbf{e}_z}(\mathrm{d}\alpha)$ is unitary (*cf.* Complement $\mathrm{C_{II}}$, § 3). This relation is the *definition* of J_z. Similarly, the Hermitian operators J_x and J_y can be introduced by starting with infinitesimal rotations about the Ox and Oy axes. The total angular momentum $\mathbf{J}$ of the system is then defined in terms of its three components J_x, J_y and J_z.

Now we can use the reasoning of § 3-c-β: the geometrical relation (8) implies that the components of $\mathbf{J}$ satisfy commutation relations which are identical to those of orbital angular momenta. Thus, *the total angular momentum of any quantum mechanical system is related to the corresponding rotation operators; the commutation relations amongst its components follow directly*, which enables us to use them, as in Chapter VI (§ B-2), to characterize any angular momentum.

Finally, let us show that, with J_x, J_y and J_z defined as we have just indicated, the operator $R_{\mathbf{u}}(\mathrm{d}\alpha)$ associated with an arbitrary infinitesimal rotation is written (u_x, u_y and u_z being the components of the unit vector $\mathbf{u}$):

$$R_{\mathbf{u}}(\mathrm{d}\alpha) = 1 - \frac{i}{\hbar}\,\mathrm{d}\alpha\,(J_x u_x + J_y u_y + J_z u_z) \tag{80}$$

which can then be condensed into the form:

$$R_{\mathbf{u}}(\mathrm{d}\alpha) = 1 - \frac{i}{\hbar}\,\mathrm{d}\alpha\,\mathbf{J}\cdot\mathbf{u} \tag{81}$$

Formula (80) is simply a consequence of the geometrical relation:

$$\mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha) = \mathscr{R}_{\mathbf{e}_x}(u_x \mathrm{d}\alpha)\, \mathscr{R}_{\mathbf{e}_y}(u_y \mathrm{d}\alpha)\, \mathscr{R}_{\mathbf{e}_z}(u_z \mathrm{d}\alpha) \tag{82}$$

valid to first order in $\mathrm{d}\alpha$, and which can be obtained directly from formula (6).

We have thus generalized expressions (48) and (49) for infinitesimal rotation operators. Since the group law is conserved locally (see above), relation (52) and the argument following it remain valid. Consequently, the finite rotation operators have expressions analogous to (55) and (57):

$$R_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\alpha\, \mathbf{J}\cdot\mathbf{u}} \tag{83}$$

5. Rotation of observables

We just showed how the vector representing the state of a quantum mechanical system transforms under rotation. But in quantum mechanics, the state of a system and the physical quantities are described independently. Therefore, we shall now indicate what happens to observables upon rotation.

5-a. General transformation law

Consider an observable A, relating to a given physical system; we shall assume, to simplify the notation, the spectrum of A to be discrete and non-degenerate:

$$A|u_n\rangle = a_n|u_n\rangle \tag{84}$$

In order to understand how this observable is affected by a rotation, we shall imagine that we have a device which can measure A in the physical system under consideration. Now, the observable A', the transform of A with respect to the geometrical rotation $\mathscr{R}$, is by definition what is measured by the device when it has been subjected to the rotation $\mathscr{R}$.

Let us assume the system to be in the eigenstate $|u_n\rangle$ of A: the device for measuring A in this system will give the result a_n without fail. But just before performing the measurement, we apply a rotation $\mathscr{R}$ to the physical system and, simultaneously, to the measurement device; their relative positions are unchanged. Consequently, if the observable A which we are considering describes a physical quantity attached only to the system which we have rotated (that is, independent of other systems or devices which we have not rotated), then, in its new position, the measurement device will still give the *same* result a_n without fail. Now, after rotation, the device, by definition, measures A', and the system is in the state:

$$|u'_n\rangle = R|u_n\rangle \tag{85}$$

We must therefore have:

$$A|u_n\rangle = a_n|u_n\rangle \Longrightarrow A'|u'_n\rangle = a_n|u'_n\rangle \tag{86}$$

Combining (85) and (86), we find:

$$A'R|u_n\rangle = a_n R|u_n\rangle \tag{87}$$

that is:

$$R^\dagger A' R|u_n\rangle = a_n|u_n\rangle \tag{88}$$

since the inverse of R is $R^\dagger$. The set of vectors $|u_n\rangle$ constitutes a basis in the state space (A is an observable), so we have:

$$R^\dagger A' R = A \tag{89}$$

that is:

$$\boxed{A' = RAR^\dagger} \tag{90}$$

In the special case of an infinitesimal rotation $\mathscr{R}_{\mathbf{u}}(\mathrm{d}\alpha)$, the general expression (81), substituted into (90), gives, to first order in $\mathrm{d}\alpha$:

$$\begin{aligned} A' &= \left(1 - \frac{i}{\hbar}\,\mathrm{d}\alpha\,\mathbf{J}\cdot\mathbf{u}\right) A \left(1 + \frac{i}{\hbar}\,\mathrm{d}\alpha\,\mathbf{J}\cdot\mathbf{u}\right) \\ &= A - \frac{i}{\hbar}\,\mathrm{d}\alpha\,[\mathbf{J}\cdot\mathbf{u}\ ,\ A] \end{aligned} \tag{91}$$

Comments:

(i) In the case of a spinless particle, relation (90) implies that:

$$\langle\mathbf{r}|A'|\mathbf{r}'\rangle = \langle\mathbf{r}|RAR^\dagger|\mathbf{r}'\rangle \tag{92}$$

Using (26) and (27), we therefore obtain:

$$\langle\mathbf{r}|A'|\mathbf{r}'\rangle = \langle\mathscr{R}^{-1}\mathbf{r}|A|\mathscr{R}^{-1}\mathbf{r}'\rangle \tag{93}$$

The transformation which enables us to obtain A' from A is therefore completely analogous to the one which gives $|\psi'\rangle$ in terms of $|\psi\rangle$ [*cf.* (19)].

(ii) Consider the case in which the observable A is associated with a classical quantity $\mathscr{A}$. $\mathscr{A}$ is then a function of the positions $\mathbf{r}_i$ and momenta $\mathbf{p}_i$ of the particles which constitute the system; the operator A is obtained from this function by applying the quantization rules given in Chapter III. We know how to find the quantity $\mathscr{A}'$ associated with $\mathscr{A}$ by a rotation $\mathscr{R}$ in classical mechanics: for example, if $\mathscr{A}$ is a scalar, $\mathscr{A}'$ is the same as $\mathscr{A}$; if $\mathscr{A}$ is the component along an axis Ou of a vectorial quantity, $\mathscr{A}'$ is the component of this same vectorial quantity along the axis which is the result of the transformation of Ou by the rotation $\mathscr{R}$. We can also construct the quantum mechanical operator corresponding to $\mathscr{A}'$, by applying the same quantization rules as above. It can be shown that this operator is the same as the operator A' given in (90); this is what is shown in Figure 1.

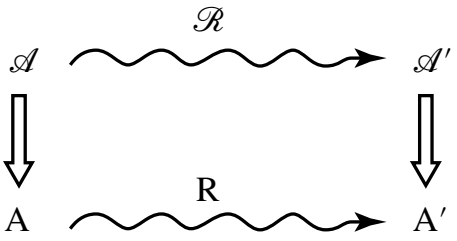


Figure 1: Behavior, under a rotation $\mathscr{R}$, of a classical physical quantity $\mathscr{A}$ and of the associated observable A.

5-b. Scalar observables

An observable A is said to be scalar if:

$$A' = A \tag{94}$$

for all R. According to (91), this implies that:

$$[A, \mathbf{J}] = 0 \tag{95}$$

A scalar observable commutes with the three components of the total angular momentum.

There are numerous examples of scalar observables. $\mathbf{J}^2$ is always a scalar (this results, as we saw in § B-2 of Chapter VI, from the commutation relations that characterize an angular momentum). For a spinless particle, $\mathbf{R}^2$, $\mathbf{P}^2$ and $\mathbf{R} \cdot \mathbf{P}$, which correspond to classical scalar quantities, are scalars. It is easy to show, moreover (*cf.* § 5-c below) that $\mathbf{R}^2$, $\mathbf{P}^2$ and $\mathbf{R} \cdot \mathbf{P}$ satisfy (95). We shall also see later (§ 6) that the Hamiltonian of an isolated physical system is a scalar.

5-c. Vector observables

A vector observable $\mathbf{V}$ is a set of three observables V_x, V_y, V_z (its Cartesian components) that is transformed by rotation according to the characteristic law of vectors. The transform, under a rotation $\mathscr{R}$, of the component $V_u = \mathbf{V} \cdot \mathbf{u}$ of $\mathbf{V}$ along a given axis Ou (of unit vector $\mathbf{u}$) must be the component $V_{\mathbf{u}'} = \mathbf{V} \cdot \mathbf{u}'$ of $\mathbf{V}$ along the axis Ou' derived from Ou by the rotation $\mathscr{R}$.

Consider, for example, the component V_x of such an observable. We shall examine its behavior under infinitesimal rotations about each of the coordinate axes. V_x is obviously unchanged by a rotation about Ox; according to (91), this can be expressed in the form:

$$[J_x \ , \ V_x] = 0 \tag{96}$$

If we perform a rotation $\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha)$ about the Oy axis, the transform of V_x is the observable $(V_x)'$ given by (91) as:

$$(V_x)' = V_x - \frac{i}{\hbar}\, \mathrm{d}\alpha\, [J_y \ , \ V_x] \tag{97}$$

But V_x is the component of $\mathbf{V}$ along the Ox axis, of unit vector e_x. The rotation $\mathscr{R}_{\mathbf{e}_y}(\mathrm{d}\alpha)$ takes $\mathbf{e}_x$ onto $\mathbf{e}'_x$ such that [formula (6)]:

$$\begin{aligned}\mathbf{e}'_x &= \mathbf{e}_x + \mathrm{d}\alpha\ \mathbf{e}_y \times \mathbf{e}_x \\ &= \mathbf{e}_x - \mathrm{d}\alpha\ \mathbf{e}_z\end{aligned} \tag{98}$$

Consequently, if $\mathbf{V}$ is a vector observable, $(V_x)'$ must be the same as $\mathbf{V} \cdot \mathbf{e}'_x$:

$$\begin{aligned}(V_x)' &= \mathbf{V} \cdot \mathbf{e}_x - \mathrm{d}\alpha\ \mathbf{V} \cdot \mathbf{e}_z \\ &= V_x - \mathrm{d}\alpha V_z\end{aligned} \tag{99}$$

Comparing (97) and (99), we see that:

$$[V_x\ ,\ J_y] = i\hbar V_z \tag{100}$$

For an infinitesimal rotation $\mathscr{R}_{\mathbf{e}_z}(\mathrm{d}\alpha)$ about the Oz axis, an argument analogous to the one above leads to the relation:

$$[J_z\ ,\ V_x] = i\hbar V_y \tag{101}$$

By studying the behavior of V_y and V_z under infinitesimal rotations, one can prove the formulas which can be derived from (96), (100) and (101) by cyclic permutation of the indices x, y, z. The set of relations obtained in this way is characteristic of a vector observable: they imply that an arbitrary infinitesimal rotation transforms $\mathbf{V} \cdot \mathbf{u}$ into $\mathbf{V} \cdot \mathbf{u}'$, where $\mathbf{u}'$ is the transform of $\mathbf{u}$ with respect to the rotation under consideration.

It is clear that the angular momentum $\mathbf{J}$ itself is a vector observable; (96), (100) and (101) then follow from the commutation relations characterizing angular momenta. For a system composed of a single spinless particle, $\mathbf{R}$ and $\mathbf{P}$ are vector observables, as can easily be verified from the canonical commutation relations. Thus the vector notation we use for $\mathbf{R}, \mathbf{P}, \mathbf{L}$ and $\mathbf{J}$ is justified.

Comments:

(i) The scalar product $\mathbf{V} \cdot \mathbf{W}$ of two vector observables, defined by the customary formula:

$$\mathbf{V} \cdot \mathbf{W} = V_x W_x + V_y W_y + V_z W_z \tag{102}$$

is a scalar operator. To verify this, we can calculate, for example, the commutator of $\mathbf{V} \cdot \mathbf{W}$ with J_x:

$$\begin{aligned}[\mathbf{V} \cdot \mathbf{W}, J_x] &= [V_y W_y, J_x] + [V_z W_z, J_x] \\ &= V_y[W_y, J_x] + [V_y, J_x]W_y + V_z[W_z, J_x] + [V_z, J_x]W_z \\ &= -i\hbar V_y W_z - i\hbar V_z W_y + i\hbar V_z W_y + i\hbar V_y W_z \\ &= 0\end{aligned} \tag{103}$$

We have already pointed out that $\mathbf{J}^2$, $\mathbf{R}^2$, $\mathbf{P}^2$ and $\mathbf{R} \cdot \mathbf{P}$ are scalar observables.

(ii) It is the *total angular momentum of the system* under study that appears in relations (96), (100) and (101). The following example illustrates the importance of this fact: if, for a two-particle system, we were to use $\mathbf{L}_1$ instead of $\mathbf{L} = \mathbf{L_1} + \mathbf{L_2}$, $\mathbf{R}_2$ would appear to be a set of three scalar observables and not a vector observable.

6. Rotation invariance

The discussion presented in the preceding sections does not have as its sole purpose the justification of the definition of angular momenta in terms of the commutation relations. The importance of rotations in physics is essentially related to the fact that physical laws are rotation invariant. We are going to explain in this section exactly what this means, and we shall indicate some of the consequences of this fundamental property.

6-a. Invariance of physical laws

Consider a physical system (S), classical or quantum mechanical, which we subject to a rotation $\mathscr{R}$ at some given time. If we take the precaution of rotating, at the same time as the system (S) under consideration, all other systems or devices that can influence it, *the physical properties and behavior of (S) are not modified*. This means that the physical laws governing the system have remained the same: the physical laws are said to be rotation invariant. Note that this property is not at all obvious: there exist transformations – those of similarity[5], for example – with respect to which the physical laws are not invariant[6]. It is therefore advisable to consider rotation invariance to be a postulate which is justified by the experimental verification of its consequences.

When we say that the physical properties and behavior of a system are unchanged by a rotation performed at the time t_0, this statement covers two observations:

(i) the properties of the system at this time are not modified (although the description of the state of the system and the physical quantities are; see preceding sections). In quantum mechanics, this implies that the transform A' of an arbitrary observable A has the same spectrum, and that the probability of finding one of the eigenvalues of this spectrum in a measurement of A' on the system after rotation is the same as it was for the measurement of A on the system before rotation. From this, it can be deduced that the operators R describing rotations in state space are linear and unitary, or antilinear and unitary (that is, anti-unitary[7]).

(ii) the time evolution of the system is not affected. To state this point more precisely, let us denote the state of the system by $|\psi(t_0)\rangle$; under a rotation performed at t_0, this state becomes:

$$|\psi'(t_0)\rangle = R|\psi(t_0)\rangle \tag{104}$$

We now let the system evolve freely and compare its state $|\psi'(t)\rangle$ at a subsequent time t to the state $|\psi(t)\rangle$ which it would have attained if it had been allowed to evolve freely from $|\psi(t_0)\rangle$. If the behavior of the system is not modified, we must have:

$$|\psi'(t)\rangle = R|\psi(t)\rangle \tag{105}$$

that is, for all t, the state $|\psi'(t)\rangle$ must be obtained from $|\psi(t)\rangle$ by the same rotation as in (104). Therefore, if $|\psi(t)\rangle$ is a solution of the Schrödinger equation, $R|\psi(t)\rangle$ is also

[5]Consider, for example, a hydrogen atom. If we multiply the distance between the proton and the electron by a constant $\lambda \neq 1$ (without modifying the charges and masses of the particles), we obtain a system whose evolution no longer obeys either classical or quantum mechanical physical laws.

[6]Let us also point out that experiments have shown that the laws governing the β-decay of nuclei are not invariant under reflection with respect to a plane (non-conservation of parity).

[7]All transformations which leave the physical laws invariant are described by unitary operators, except for time reversal, with which is associated an anti-unitary operator.

a solution of this equation: *the transform of a possible motion of the system is also a possible motion.* We shall see in § 6-b that this implies that the Hamiltonian H of the system is a scalar observable.

The invariance of physical laws under rotation is related to the *symmetry properties* of the equations which state these laws mathematically. To understand the origin of these symmetries, consider, for example, a system composed of a single (spinless) particle. The expression of the physical laws governing such a system explicitly involves the parameters $\mathbf{r}(x, y, z)$ and $\mathbf{p}(p_x, p_y, p_z)$ which characterize the position of the particle and its momentum: in classical mechanics, $\mathbf{r}$ and $\mathbf{p}$ define at each instant the state of the particle; in quantum mechanics, although the meaning of these parameters is a little less simple, they appear in the wave function $\psi(\mathbf{r})$ and its Fourier transform $\overline{\psi}(\mathbf{p})$. When the particle is subjected to an instantaneous rotation $\mathscr{R}$, $\mathbf{r}$ and $\mathbf{p}$ are transformed into $\mathbf{r}'$ and $\mathbf{p}'$ such that:

$$\begin{aligned} \mathbf{r}' &= \mathscr{R}\,\mathbf{r} \\ \mathbf{p}' &= \mathscr{R}\,\mathbf{p} \end{aligned} \tag{106}$$

If we replace $\mathbf{r}$ by $\mathscr{R}^{-1}\mathbf{r}'$ and $\mathbf{p}$ by $\mathscr{R}^{-1}\mathbf{p}'$ in the equations which express the physical laws, we obtain relations that now involve $\mathbf{r}'$ and $\mathbf{p}'$. The invariance of physical laws under the rotation $\mathscr{R}$ thus implies that *the form of the equations for* $\mathbf{r}'$ *and* $\mathbf{p}'$ *is the same as that of the equations for* $\mathbf{r}$ *and* $\mathbf{p}$: simply omitting the primes labeling the new parameters must give us back the original equations. It is clear that this considerably restrains the number of possible forms of these equations.

Comments:

(i) What happens when we perform a rotation on a system which is not isolated? Consider, for example, a particle in an external potential. If we rotate this system, *without simultaneously rotating the sources of the external potential*, the subsequent evolution of the system is, in general, modified[8]. In classical mechanics, the forces exerted on the particle are not the same in its new position. In quantum mechanics, $\psi'(\mathbf{r}, t) = \psi(\mathscr{R}^{-1}\mathbf{r}, t)$ is a solution of a Schrödinger equation in which the potential $V(\mathbf{r})$ is replaced by $V(\mathscr{R}^{-1}\mathbf{r})$, which is, in general, different from $V(\mathbf{r})$. Therefore, the transform of a possible motion is no longer a possible motion. The presence of the external potential destroys, so to speak, the homogeneity of the space in which the system under study evolves.

However, the external potential may present certain symmetries that allow certain rotations to be performed on the physical system without its behavior being modified. If there exist rotations $\mathscr{R}_0$ such that $V(\mathscr{R}_0^{-1}\mathbf{r})$ is the same as $V(\mathbf{r})$, the properties of the system are unchanged by one of these rotations $\mathscr{R}_0$. This is the case, for example, for

[8] If the particle is placed in a vector potential, its properties immediately after the rotation may be profoundly modified. Consider, for example, a spinless particle in an external magnetic field. According to the transformation law (19), the probability current given by formula (D-20) of Chapter III cannot in general be derived by rotation of the initial current, as it depends on the vector potential describing the magnetic field.The physical interpretation of this phenomenon is as follows. We can imagine that, instead of rotating the particle, we rotate the magnetic field rapidly and in the opposite direction. The wave function has not had the time to change: this corresponds to formula (19). If the physical properties are modified, it is because an induced electromotive field has appeared and acted on the particle. This action does not depend on the exact way in which we rotate the magnetic field, provided that we do so quickly enough.

central potentials, that is, those depending only on the distance to a fixed point O: the rotations $\mathscr{R}_0$ are then all rotations that conserve the point O (*cf.* Chap. VII).

(*ii*) Let us return to the case of isolated physical systems. Thus far, we have adopted an "active" viewpoint: the observer remains fixed, and the physical system is rotated.

We can also define a "passive" viewpoint: the observer rotates, and, without touching the system being studied, uses a new coordinate frame, derived from the initial frame by the given rotation. Rotation invariance is then expressed in the following way: in his new position (that is, using his new coordinate axes), the observer describes physical phenomena by laws that have the same form as in the old frame. Nothing allows him to assert that one of his positions is more fundamental than the other: it is impossible to define an absolute orientation in space by the study of any physical phenomenon. It is clear that, for an isolated system, a "passive" rotation is equivalent to the "active" rotation through an equal angle about the opposite axis.

6-b. Consequence: conservation of angular momentum

We indicated in § 6-a that rotation invariance is related to symmetry properties of the equations expressing the physical laws. Here, we shall study the case of the Schrödinger equation, and we shall show that *the Hamiltonian of an isolated physical system is a scalar observable.*

Consider an isolated system in the state $|\psi(t_0)\rangle$. We perform an arbitrary rotation $\mathscr{R}$ at time t_0; the state of the system becomes:

$$|\psi'(t_0)\rangle = R|\psi(t_0)\rangle \tag{107}$$

where R is the "image" of the rotation $\mathscr{R}$. If we now let the system evolve freely from $|\psi'(t_0)\rangle$, its state at the instant $t_0 + \mathrm{d}t$, according to the Schrödinger equation, will be:

$$|\psi'(t_0 + \mathrm{d}t)\rangle = |\psi'(t_0)\rangle + \frac{\mathrm{d}t}{i\hbar}\, H|\psi'(t_0)\rangle \tag{108}$$

Now, if we had not performed the rotation, the state of the system at time $t_0 + \mathrm{d}t$ would have been:

$$|\psi(t_0 + \mathrm{d}t)\rangle = |\psi(t_0)\rangle + \frac{\mathrm{d}t}{i\hbar}\, H|\psi(t_0)\rangle \tag{109}$$

Rotation invariance implies (*cf.* § 6-a) that:

$$|\psi'(t_0 + \mathrm{d}t)\rangle = R|\psi(t_0 + \mathrm{d}t)\rangle \tag{110}$$

where R is the same as in (107). According to the two preceding equations, this implies that:

$$RH|\psi(t_0)\rangle = H|\psi'(t_0)\rangle \tag{111}$$

that is:

$$RH|\psi(t_0)\rangle = HR|\psi(t_0)\rangle \tag{112}$$

Since $|\psi(t_0)\rangle$ is arbitrary, it follows that H commutes with all rotation operators. For this to be so, it is necessary and sufficient that H commute with the infinitesimal rotation

operators, that is, with the three components of the total angular momentum **J** of the system:

$$[H\ ,\ \mathbf{J}] = 0 \tag{113}$$

H is therefore a scalar observable.

Rotation invariance is therefore related to the fact that *the total angular momentum of an isolated system is a constant of the motion*: conservation of angular momentum can be seen to be a consequence of rotation invariance.

Comments:

(i) The Hamiltonian of a non-isolated system is not, in general, a scalar. However, if certain rotations exist that leave the system invariant [comment (i) of § 6-a], the Hamiltonian commutes with the corresponding operators. Thus, the Hamiltonian of a particle in a central potential commutes with the operator **L** associated with the angular momentum of the particle with respect to the center of forces.

(ii) For an isolated system composed of several interacting particles, the Hamiltonian commutes with the *total* angular momentum. However, it does not generally commute with the individual angular momentum of each particle. For the transform of a possible motion to remain a possible motion, the rotation must be performed on the whole system, not on only some of the particles.

6-c. Applications

We have just shown that rotation invariance means that the total angular momentum **J** of an isolated system is a constant of the motion in the quantum mechanical sense. It is therefore useful to determine the stationary states of such a system (eigenstates of the Hamiltonian) which are also eigenstates of $\mathbf{J}^2$ and J_z. We can then choose for the state space a standard basis $\{|k, j, m\rangle\}$, composed of eigenstates common to H, $\mathbf{J}^2$ and J_z:

$$\begin{aligned} H|k, j, m\rangle &= E\ |k, j, m\rangle \\ \mathbf{J}^2|k, j, m\rangle &= j(j+1)\hbar^2\ |k, j, m\rangle \\ J_z|k, j, m\rangle &= m\hbar\ |k, j, m\rangle \end{aligned} \tag{114}$$

α. *Essential rotational degeneracy*

Since the Hamiltonian H is a scalar observable, it commutes with J_+ and J_-. From this fact, we deduce that $|k, j, m+1\rangle$ and $|k, j, m-1\rangle$, which are respectively proportional to $J_+|k, j, m\rangle$ and $J_-|k, j, m\rangle$, are eigenstates of H with the same eigenvalue as $|k, j, m\rangle$ [the argument is the same as for formula (C-48) of Chapter VI]. Thus it can be shown by iteration that the $(2j + 1)$ vectors of the standard basis characterized by the given values of k and j have the same energy. The corresponding degeneracy of the eigenvalues of H is called "essential" because it arises from rotation invariance and occurs for any form of the Hamiltonian H. Of course, in certain cases, the energy levels can present additional degeneracies, which are called "accidental". We shall see an example of this in Chapter VII, § C.

β. Matrix elements of observables in a standard basis

When we study a physical quantity in an isolated system, the knowledge of the behavior of the associated observable under a rotation enables us to establish some of its properties, without having to consider its precise form. We can predict that only some of its matrix elements in a standard basis such as $\{|k, j, m\rangle\}$ will be different from zero, and we can give the relations between them. Thus, a scalar observable has matrix elements only between two basis vectors whose values of j are equal, as are their values of m [this results from the fact that this observable commutes with $\mathbf{J}^2$ and J_z; *cf.* theorems on commuting observables, § D-3-a of Chapter II]. Moreover, these non-zero elements are independent of m (since the scalar observable also commutes with J_+ and J_-). For vector or tensor observables, these properties are contained in the *Wigner-Eckart theorem*, which we shall prove later in a special case (*cf.* Complement D_X), and which is frequently used in the areas of physics in which phenomena are treated by quantum mechanics (atomic, molecular, and nuclear physics, elementary particle physics, etc.).

References and suggestions for further reading:

Symmetry and conservation laws: Feynman III (1.2), Chap. 17; Schiff (1.18), Chap. 7; Messiah (1.17), Chap. XV; see also the articles by Morrisson (2.28), Feinberg and Goldhaber (2.29), Wigner (2.30).

Relation with group theory: Messiah (1.17), Appendix D; Meijer and Bauer (2.18), Chaps. 5 and 6; Bacry (10.31), Chap. 6; Wigner (2.23), Chaps. 14 and 15. See also Omnès (16.13), in particular Chap. III.

Complement C_{VI}

Rotation of diatomic molecules

1 Introduction 739
2 Rigid rotator. Classical study 740
2-a Notation 740
2-b Motion of the rotator. Angular momentum and energy 741
2-c The fictitious particle associated with the rotator 741
3 Quantization of the rigid rotator 741
3-a The quantum mechanical state and observables of the rotator 741
3-b Eigenstates and eigenvalues of the Hamiltonian 742
3-c Study of the observable Z 744
4 Experimental evidence for the rotation of molecules 746
4-a Heteropolar molecules. Pure rotational spectrum 746
4-b Homopolar molecules. Raman rotational spectra 749

1. Introduction

In § 1 of Complement A_V, we studied the vibrations of the two nuclei of a diatomic molecule about their equilibrium position, neglecting the rotation of these two nuclei about their center of mass. We obtained stationary vibrational states of energies E_v whose wave functions $\varphi_v(r)$ depended only on the distance r between the nuclei.

Here, we shall adopt a complementary point of view: we shall study the rotation of the two nuclei about their center of mass, neglecting their vibrations. That is, we shall assume that the distance r between them remains fixed and equal to r_e (where r_e represents the distance between the two nuclei in the stable equilibrium position of the molecule; see Figure 1 of Complement A_V). The wave functions of the stationary rotational states then can depend only on the polar angles θ and φ which define the direction of the molecular axis. We shall see that these wave functions are the spherical harmonics $Y_l^m(\theta, \varphi)$ [studied in Chapter VI (§ D-1) and in Complement A_{VI}], and that they correspond to a rotational energy E_l that depends only on l.

Actually, in the center of mass frame, the molecule both rotates and vibrates, and the wave functions of its stationary states must be functions of the three variables r, θ and φ. We shall show in Complement F_{VII} that, to a first approximation, these wave functions are of the form $\varphi_v(r)Y_l^m(\theta, \varphi)/r$ and correspond to the energy $E_v + E_l$. This result justifies the approach adopted here, which consists of considering only one degree of freedom – rotational or vibrational – at a time[1]. We shall begin in § 2 by presenting the classical study of a system of two masses separated by a fixed distance (rigid rotator). The quantum mechanical treatment of this problem will then be taken up in § 3, where

[1] In Complement F_{VII}, we shall also study the corrections introduced by the coupling between the vibrational and rotational degrees of freedom.

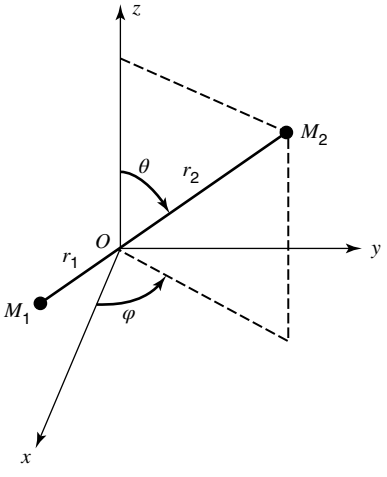


Figure 1: Parameters defining the position of the rigid rotator M_1M_2 whose center of mass is at the origin O of the reference frame; the distances r_1 and r_2 are fixed; only the polar angles θ and φ can vary.

we shall use the results of Chapter VI concerning orbital angular momentum. Finally in § 4, we shall describe some experimental manifestations of the rotation of diatomic molecules (pure and Raman rotational spectra).

2. Rigid rotator. Classical study

2-a. Notation

Two particles, of mass m_1 and m_2, are separated by a fixed distance r_e. Their center of mass O is chosen as the origin of a coordinate frame $Oxyz$ with respect to which the direction of the axis connecting them is defined by means of the polar angles θ and φ (Fig. 1). The distances OM_1 and OM_2 are denoted respectively by r_1 and r_2; by definition of the center of mass:

$$m_1 r_1 = m_2 r_2 \tag{1}$$

which allows us to write:

$$\frac{r_1}{m_2} = \frac{r_2}{m_1} = \frac{r_e}{m_1 + m_2} \tag{2}$$

The moment of inertia I of the system with respect to O is equal to:

$$I = m_1 r_1^2 + m_2 r_2^2 \tag{3}$$

Introducing the reduced mass:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{4}$$

and using (2), we can put I in the form:

$$I = \mu r_e^2 \tag{5}$$

2-b. Motion of the rotator. Angular momentum and energy

If no external force acts on the rotator, the total angular momentum $\boldsymbol{\mathcal{L}}$ of the system with respect to the point O is a constant of the motion. The rotator therefore rotates about O in a plane perpendicular to the fixed vector $\boldsymbol{\mathcal{L}}$, with a constant angular velocity ω_R. The modulus of $\boldsymbol{\mathcal{L}}$ is related to ω_R by:

$$|\boldsymbol{\mathcal{L}}| = m_1 r_1 r_1 \omega_R + m_2 r_2 r_2 \omega_R = I\omega_R \tag{6}$$

that is, using (5):

$$|\boldsymbol{\mathcal{L}}| = \mu r_e^2 \omega_R \tag{7}$$

The rotational frequency of the system $\nu_R = \omega_R/2\pi$ is proportional to the angular momentum $|\boldsymbol{\mathcal{L}}|$ and inversely proportional to the moment of inertia I.

In the center of mass frame, the total energy $\mathcal{H}$ of the system reduces to the rotational kinetic energy:

$$\mathcal{H} = \frac{1}{2} I\, \omega_R^2 \tag{8}$$

which can also be written, using (6) and (5):

$$\mathcal{H} = \frac{\boldsymbol{\mathcal{L}}^2}{2I} = \frac{\boldsymbol{\mathcal{L}}^2}{2\mu r_e^2} \tag{9}$$

2-c. The fictitious particle associated with the rotator

Formulas (5), (7) and (9) show that the problem we are studying here is formally equivalent to that of a fictitious particle of mass μ forced to remain at a fixed distance r_e from the point O, about which it rotates with the angular velocity ω_R. $\boldsymbol{\mathcal{L}}$ is the angular momentum of this fictitious particle with respect to O.

3. Quantization of the rigid rotator

3-a. The quantum mechanical state and observables of the rotator

Since r_e is fixed, the parameters defining the position of the rotator (or that of the associated fictitious particle) are the polar angles θ and φ of Figure 1. The quantum mechanical state of the rotator will then be described by a wave function $\psi(\theta, \varphi)$ which depends only on these two parameters. $\psi(\theta, \varphi)$ is square-integrable; we shall assume it to be normalized:

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta\, \mathrm{d}\theta\ |\psi(\theta, \varphi)|^2 = 1 \tag{10}$$

The physical interpretation of $\psi(\theta, \varphi)$ is the following: $|\psi(\theta, \varphi)|^2 \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi$ represents the probability of finding the axis of the rotator pointing in the solid angle element $\mathrm{d}\Omega = \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi$ about the direction of polar angles θ and φ.

Using Dirac notation, we associate with every square-integrable function $\psi(\theta, \varphi)$, a ket $|\psi\rangle$ of the state space $\mathcal{E}_\Omega$:

$$\psi(\theta, \varphi) \leftrightarrow |\psi\rangle \in \mathcal{E}_\Omega \tag{11}$$

The scalar product of $|\psi\rangle$ and $|\chi\rangle$ is, by definition:

$$\langle\chi|\psi\rangle = \int \mathrm{d}\Omega\ \chi^*(\theta,\varphi)\ \psi(\theta,\varphi) \tag{12}$$

where $\chi(\theta,\varphi)$ and $\psi(\theta,\varphi)$ are the wave functions associated with $|\chi\rangle$ and $|\psi\rangle$.

The quantum mechanical Hamiltonian H of the rotator (or of the associated fictitious particle) can be obtained by replacing $\boldsymbol{\mathcal{L}}^2$ in expression (9) for the classical energy by the operator $\mathbf{L}^2$ studied in § D of Chapter VI:

$$H = \frac{\mathbf{L}^2}{2\mu r_e^2} \tag{13}$$

H is an operator acting in $\mathcal{E}_\Omega$. According to formula (D-6a) of Chapter VI, if $|\psi\rangle$ is represented by the wave function $\psi(\theta,\varphi)$, $H|\psi\rangle$ is represented by:

$$H|\psi\rangle \leftrightarrow -\frac{\hbar^2}{2\mu r_e^2}\left[\frac{\partial^2}{\partial\theta^2} + \frac{1}{\mathrm{tg}\theta}\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\psi(\theta,\varphi) \tag{14}$$

Other observables of interest, which we shall study later, are those which correspond to the three algebraic projections x, y, z of the segment M_1M_2 (x, y, z are also the coordinates of the fictitious particle):

$$\begin{aligned} x &= r_e \sin\theta\cos\varphi \\ y &= r_e \sin\theta\sin\varphi \\ z &= r_e \cos\theta \end{aligned} \tag{15}$$

The importance of these variables will be seen in § 4-a. The observables X, Y, Z corresponding to x, y, z act in $\mathcal{E}_\Omega$. With the kets $X|\psi\rangle, Y|\psi\rangle, Z|\psi\rangle$, are associated the functions:

$$\begin{aligned} X|\psi\rangle &\leftrightarrow r_e \sin\theta\ \cos\varphi\ \psi(\theta,\varphi) \\ Y|\psi\rangle &\leftrightarrow r_e \sin\theta\ \sin\varphi\ \psi(\theta,\varphi) \\ Z|\psi\rangle &\leftrightarrow r_e \cos\theta\ \psi(\theta,\varphi) \end{aligned} \tag{16}$$

Comment:

As we have already pointed out in the introduction, the true wave functions of the molecule depend on r, θ, φ. Similarly, the observables of this molecule, obtained from the corresponding classical quantities by the quantization rules of Chapter III, act on these functions of three variables and not solely on the functions of θ and φ. In Complement F_{VII}, we shall justify the point of view we are adopting here, namely, ignoring the radial part of the wave functions and considering r to be a fixed parameter that is equal to r_e [*cf.* formulas (14) and (16)].

3-b. Eigenstates and eigenvalues of the Hamiltonian

We determined the eigenvalues of the operator $\mathbf{L}^2$ in § D of Chapter VI: they are of the form $l(l+1)\hbar^2$, where l is any non-negative integer. Furthermore, we know an

orthonormal system of eigenfunctions of $\mathbf{L}^2$: the spherical harmonics $Y_l^m(\theta,\varphi)$, which constitute a basis in the space of functions that are square-integrable in θ and φ (§ D-l-c-β of Chapter VI). We shall denote by $|l,m\rangle$ the ket of $\mathcal{E}_\Omega$ associated with $Y_l^m(\theta,\varphi)$:

$$Y_l^m(\theta,\varphi) \leftrightarrow |l,m\rangle \tag{17}$$

We see from (13) that:

$$H|l,m\rangle = \frac{l(l+1)\hbar^2}{2\mu r_e^2}\ |l,m\rangle \tag{18}$$

It is customary to set:

$$B = \frac{\hbar}{4\pi I} = \frac{\hbar}{4\pi\mu r_e^2} \tag{19}$$

B is called the "rotational constant" and has the dimensions of a frequency[2]. The eigenvalues of H are thus of the form:

$$E_l = Bh\, l(l+1) \tag{20}$$

Since, for a given value of l, there exist $(2l+1)$ spherical harmonics $Y_l^m(\theta,\varphi)$, where $m = -l, -l+1, ..., l$, we see that each eigenvalue E_l is $(2l+1)$-fold degenerate. Figure 2 represents the first energy levels of the rotator. The separation of two adjacent levels, l and $l-1$, is equal to:

$$E_l - E_{l-1} = Bh\Big[l(l+1) - l(l-1)\Big] = 2Bhl \tag{21}$$

and increases linearly with l.

The eigenstates of H satisfy the following orthogonality and closure relations (deduced from those satisfied by the spherical harmonics, § D-1-c-β of Chapter VI):

$$\begin{aligned}&\langle l,m|l',m'\rangle = \delta_{ll'}\ \delta_{mm'}\\ &\sum_{l=0}^{\infty}\sum_{m=-l}^{+l} |l,m\rangle\langle l,m| = 1\end{aligned} \tag{22}$$

The most general quantum state of the rotator can be expanded on the states $|l,m\rangle$:

$$|\psi(t)\rangle = \sum_{l=0}^{\infty}\sum_{m=-l}^{+l} c_{l,m}(t)\ |l,m\rangle \tag{23}$$

The component:

$$c_{l,m}(t) = \langle l,m|\psi(t)\rangle = \int \mathrm{d}\Omega\ Y_l^{m*}(\theta,\varphi)\ \psi(\theta,\varphi;t) \tag{24}$$

evolves in time in accordance with the equation:

$$c_{l,m}(t) = c_{l,m}(0)\,\mathrm{e}^{-iE_l t/\hbar} \tag{25}$$

[2]The speed of light c is sometimes placed in the denominator of the right-hand side of (19). B then has the dimensions of an inverse length and is expressed in cm^{-1} (in the CGS system).

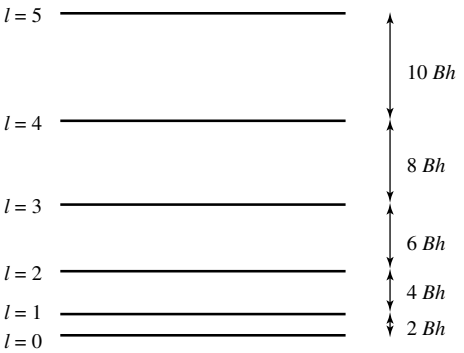


Figure 2: First levels of the rigid rotator, of energies:

$$E_l = Bhl(l+1)$$

(with $l = 0, 1, 2...$). Each level E_l for which $l \geq 1$ is separated from the next lower level by an energy $2Bhl$.

3-c. Study of the observable *Z*

Earlier, we introduced the observables X, Y, Z which correspond to the projections onto the three axes of the segment M_1M_2. In this section, we shall study the evolution of the mean values of these observables and compare the results obtained with those predicted by classical mechanics. We shall confine ourselves to the calculation of $\langle Z \rangle(t)$ since $\langle X \rangle(t)$ and $\langle Y \rangle(t)$ have analogous properties.

A Bohr frequency $(E_l - E_{l'})/h$ can appear in the function $\langle Z \rangle(t)$, if Z has a non-zero matrix element between a state $|l, m\rangle$ of energy E_l and a state $|l', m'\rangle$ of energy $E_{l'}$. The first problem is therefore to find the non-zero matrix elements of Z. To solve this problem, we shall use the following relation, which can be established by using the mathematical properties of spherical harmonics [Complement A_{VI}, formula (35)]:

$$\cos\theta \; Y_l^m(\theta, \varphi) = \sqrt{\frac{l^2 - m^2}{4l^2 - 1}} \; Y_{l-1}^m(\theta, \varphi) + \sqrt{\frac{(l+1)^2 - m^2}{4(l+1)^2 - 1}} \; Y_{l+1}^m(\theta, \varphi) \tag{26}$$

From this, we deduce, using (16), (17) and (22):

$$\langle l', m' | Z | l, m \rangle = r_e \delta_{mm'} \left[\delta_{l',l-1} \sqrt{\frac{l^2 - m^2}{4l^2 - 1}} + \delta_{l',l+1} \sqrt{\frac{(l+1)^2 - m^2}{4(l+1)^2 - 1}} \right] \tag{27}$$

Comment:

According to (27), the selection rules satisfied by Z are: $\Delta l = \pm 1, \Delta m = 0$. It can be shown that for X and Y we have: $\Delta l = \pm 1$, $\Delta m = \pm 1$. Since the energies depend only on l, the Bohr frequencies are the same for $\langle X \rangle, \langle Y \rangle$ and $\langle Z \rangle$.

The operator Z can therefore connect only states belonging to two adjacent levels of Figure 2 (the corresponding transitions are represented by vertical arrows in Figure 2). The only Bohr frequencies which appear in the evolution of $\langle Z \rangle(t)$ are thus of the form:

$$\nu_{l,l-1} = \frac{E_l - E_{l-1}}{h} = 2Bl \tag{28}$$

They form a series of equidistant frequencies, separated by the interval $2B$ (Fig. 3).

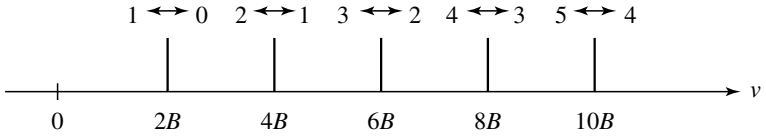


Figure 3: Frequencies appearing in the evolution of the mean value of the observable Z. Because of the selection rule $\Delta l = \pm 1$, only the Bohr frequencies $2Bl$ (with $l \geq 1$), associated with two adjacent levels E_l and E_{l-1} in Figure 2, are observed.

The mean value $\langle Z \rangle(t)$ can evolve only at a well-defined series of frequencies, This is unlike the classical case, in which the frequency of rotation ν_R of the rotator can take on any value.

According to (27), if the system is in a stationary state $|l, m\rangle$, $\langle Z \rangle(t)$ is always zero, even for large l. To obtain a quantum mechanical state in which $\langle Z \rangle$ behaves like the corresponding classical variable z, one must superpose a large number of states $|l, m\rangle$. If we assume that the state of the system is given by formula (23), and that the numbers $|c_{l,m}(0)|^2$ have values which vary with l as is shown in Figure 4, the most probable value of l, $l_{\mathbf{M}}$, is very large; the spread Δl of the values of l is also very large in absolute value, but very small in relative value:

$$l_M, \Delta l \gg 1 \tag{29a}$$

$$\frac{\Delta l}{l_M} \ll 1 \tag{29b}$$

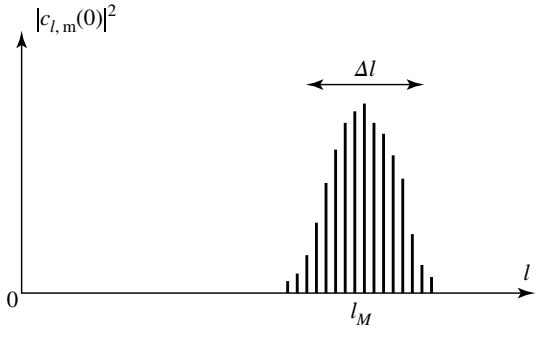


Figure 4: Square of the moduli of the expansion coefficients of a "quasi-classical" state on the stationary states $|l, m\rangle$ of the rigid rotator. The spread Δl is large; however, since the most probable value of l, $l_{\mathbf{M}}$, is very large, we have $\Delta l / l_{\mathbf{M}} \ll 1$, and the relative accuracy with respect to l is very good.

It can then be shown that, in such a state:

$$\langle \mathbf{L} \rangle^2 \simeq \langle \mathbf{L}^2 \rangle \simeq l_M(l_M + 1)\hbar^2 \simeq l_M^2 \hbar^2 \tag{30}$$

In addition, the Bohr frequencies that appear in $\langle Z \rangle(t)$ are then all very close (in relative value) to:

$$\nu_M = 2Bl_M \tag{31}$$

Eliminating l_M between (30) and (31), we obtain, according to (19):

$$\nu_M \simeq \frac{2B|\langle \mathbf{L} \rangle|}{\hbar} = \frac{|\langle \mathbf{L} \rangle|}{2\pi I} \tag{32}$$

which is the equivalent of the classical relation (6).

Comment:

It is interesting to study in greater detail the motion of the wave packet corresponding to the state of Figure 4. It is represented by a function of θ and φ and can be considered to evolve on the sphere of unit radius. The preceding discussion shows that this wave packet rotates on the sphere with the average frequency ν_M. Because of the spread Δl of l and the corresponding spread $2B\Delta l$ of the Bohr frequencies which enter into $\langle X \rangle$, $\langle Y \rangle$ and $\langle Z \rangle$, the wave packet becomes distorted over time. This distortion becomes appreciable after a time of the order of:

$$\tau \simeq \frac{1}{2B\Delta l} \tag{33}$$

Since the spread of l is small in relative value, we have:

$$\nu_M \tau \simeq \frac{l_M}{\Delta l} \gg 1 \tag{34}$$

The distortion of the wave packet is therefore slow, relative to its rotation.

In fact, the Bohr frequencies of the system form a *discrete* series of *equidistant* frequencies, separated by the interval $2B$. The motion which results from the superposition of these frequencies is therefore periodic, of period:

$$T = \frac{1}{2B} \tag{35}$$

with, according to (29a):

$$T \gg \tau \gg \frac{1}{\nu_M} \tag{36}$$

The distortion of the wave packet is therefore not irreversible; it follows a cycle which is repeated periodically. This is related to the fact that the wave packet evolves on the unit sphere, which is a bounded surface. This behavior should be compared with that of free wave packets (irreversible spreading; Complement G_I) and that of the quasi-classical states of the harmonic oscillator (oscillation without distortion; Complement G_V).

4. Experimental evidence for the rotation of molecules

4-a. Heteropolar molecules. Pure rotational spectrum

α. Description of the spectrum

If the molecule is composed of two different atoms, the electrons are attracted by the more electronegative atom, and the molecule generally has a *permanent* electric dipole moment d_0, directed along the molecular axis. The projection of the electric dipole moment onto Oz becomes, in quantum mechanics, an observable proportional to Z. We have seen that $\langle Z \rangle(t)$ evolves at all the Bohr frequencies $2Bl$ ($l = 1, 2, 3, ...$) shown in

Figure 3. Thus we see how the molecule is coupled to the electromagnetic field and can absorb or emit radiation polarized parallel[3] to Oz, on the condition that the frequency of this radiation is equal to one of the Bohr frequencies $2Bl$.

The corresponding absorption or emission spectrum of the molecule is called the "pure rotational spectrum". It is composed of a series of equidistant lines, the frequency separation between two successive lines being equal to $2B$, as in Figure 3. The absorption (or emission) of the line of frequency $2Bl$ corresponds to the passage of the molecule from the level $l-1$ to the level l (or from the level l to the level $l-1$), at the same time that a photon of frequency $2Bl$ is absorbed (or emitted). Figure 5 represents this process schematically [(5-a) represents the absorption and (5-b), the emission of a photon of frequency $2Bl$].

The pure rotational spectra of diatomic molecules therefore provide direct experimental proof of the quantization of the observable $\mathbf{L}^2$.

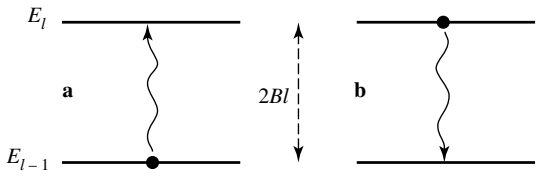


Figure 5: Schematic representation of the passage of the molecule from a given rotational level to the neighboring level with absorption (fig. a) or emission (fig. b) of a photon.

β. Comparison with the "pure vibrational" spectrum

In § 1-c-α of Complement A_V, we studied the "pure vibrational" spectrum of a heteropolar diatomic molecule. It is interesting to compare this spectrum with the pure rotational spectrum we are studying here.

(i) The rotational frequencies of a diatomic molecule are generally much lower than the vibrational frequencies. The separation $2B/c$ of two rotational lines varies between a few tenths of a cm^{-1} and a few dozen cm^{-1}. For small values of l, the rotational frequencies $2Bl$ therefore correspond to wavelengths of the order of a centimeter or a millimeter. Taking HCl as an example, the separation $2B/c$ is equal to 20.8 cm^{-1}, while the vibration frequency, which corresponds to 2 886 cm^{-1}, is more than a hundred times greater.

Pure rotational spectra therefore fall in the very far infrared or the microwave domain.

Comment:

As we shall show in Complement F_{VII}, the rotation of molecules is also responsible for a fine structure of vibrational spectra (vibration-rotation spectra). $2B$ can then be measured in a domain of wavelengths which is no longer that of microwaves. The same comment applies to the Raman rotational effect (§ 4-b below), which appears as a rotational structure of an optical line.

[3] If we study the motion of $\langle X\rangle(t)$ and $\langle Y\rangle(t)$, we see that the molecule can also absorb or emit radiation polarized parallel to Ox or Oy.

(*ii*) The "pure vibrational" spectrum studied in Complement A_V has only *one vibrational line.* This is due to the fact that the various vibrational levels are equidistant (if the anharmonicity of the potential is neglected) and, consequently, only one Bohr frequency appears in the dipole motion (selection rule $\Delta v = \pm 1$). On the other hand, the pure rotational spectrum consists of a *series of equidistant lines.*

(*iii*) We indicated in Complement A_V that the permanent electric dipole moment of the heteropolar molecule can be expanded in powers of $r - r_e$ in the neighborhood of the stable equilibrium position of the molecule:

$$D(r) = d_0 + d_1(r - r_e) + ... \tag{37}$$

For the pure vibrational spectrum to appear, $D(r)$ must vary with r : d_1 must therefore be different from zero. On the other hand, even if r remains fixed and equal to r_e, the rotation of the molecule modulates the projection of the electric dipole onto one of the axes, provided that d_0 is different from zero. Thus we see that the study of the intensity of vibrational and rotational lines permits the separate measurement of the coefficients d_1 and d_0 of (37).

γ. *Applications*

The study of pure rotational spectra has some interesting applications; we shall mention three examples.

(*i*) Measurement of the separation $2B$ of two neighboring lines yields the moment of inertia I of the molecule, according to (19). If we know m_1 and m_2, we can deduce r_e, the separation of the two nuclei in the stable equilibrium position of the molecule [r_e is the abscissa of the minimum of the curve $V(r)$ of Figure 1 in Complement A_V]. Recall that measurement of the vibrational frequency yields the curvature of $V(r)$ at $r = r_e$.

(*ii*) Consider two diatomic molecules $N - M$ and $N - M'$, in which two isotopes M and M' of the same element are bound to the same atom N. Since the distances r_e between the nuclei are equal in the two molecules, measurement of the ratio of the corresponding coefficients B, which can be performed with great accuracy, yields the ratio of the masses of the two isotopes M and M'.

One could also compare the vibrational frequencies of the two molecules, but it is preferable to use the rotational spectrum, since the rotational frequencies vary with $1/\mu$ [formula (19)], while the vibrational frequencies vary with $1/\sqrt{\mu}$ [formula (5) of A_V].

(*iii*) In the study of a sample containing a great number of identical molecules, the relative intensities of the lines (in absorption or emission) of the pure rotational spectrum yields information about the distribution of the molecules among the various levels E_l. Unlike what happens in the case of the vibrational spectrum, transitions between two given adjacent levels (arrows of Figure 2) occur at a particular frequency, which is characteristic of these two levels. Thus, the intensity of the corresponding line depends on the number of molecules that are found in the two levels.

This information can be used to determine the temperature of a medium[4]. If thermodynamic equilibrium has been attained, we know that the probability that a given molecule is in a particular state of energy E_l is proportional to $e^{-E_l/kT}$; since the degeneracy of the rotational level E_l is $(2l + 1)$, the total probability $\mathscr{P}$ of finding the

[4] Actually, the vibration-rotation or Raman rotational spectra are more often used, since they fall into more convenient frequency ranges than does the pure rotational spectrum.

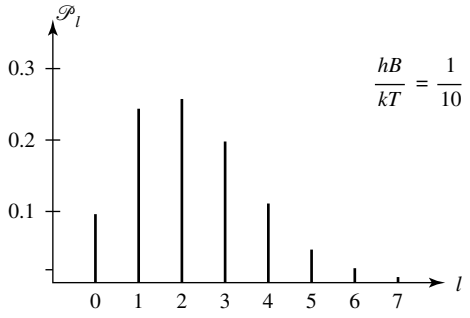


Figure 6: Population $\mathscr{P}_l$ of the various rotational levels E_l at thermodynamic equilibrium. The fact that $\mathscr{P}_l$ begins by increasing with l arises from the $(2l+1)$-fold degeneracy of the levels E_l. When l becomes sufficiently large, the Boltzmann factor $\mathrm{e}^{-E_l/kT}$ prevails and is responsible for the decrease in $\mathscr{P}_l$.

molecule being considered in one of the states of the level E_l (the "population" of the level E_l) is:

$$\begin{aligned}\mathscr{P}_l &= \frac{1}{Z}(2l+1)\mathrm{e}^{-E_l/kT} \\ &= \frac{1}{Z}(2l+1)\mathrm{e}^{-l(l+1)hB/kT}\end{aligned} \tag{38}$$

where:

$$Z = \sum_{l=0}^{\infty}(2l+1)\mathrm{e}^{-l(l+1)hB/kT} \tag{39}$$

is the associated partition function. If we are studying a system containing a large number of molecules whose interactions can be neglected, $\mathscr{P}_l$ gives the fraction of them whose energy is E_l.

At ordinary temperatures, hB is much smaller than kT, so that several rotational levels are populated. Note that the presence of the factor $(2l+1)$ means that it is not the lowest levels that are the most populated: Figure 6 indicates the shape of $\mathscr{P}_l$ as a function of l for a temperature T such that hB/kT is of the order of 1/10. Recall that the vibrational levels, on the other hand, are non-degenerate, and that their separation is much greater than hB; consequently, when the distribution of the molecules between the two rotational levels is that of Figure 6, they are practically all in the vibrational ground state ($v = 0$).

4-b. Homopolar molecules. Raman rotational spectra

As we pointed out in § 1-c-β of Complement $\mathrm{A_V}$, a homopolar molecule (that is, a molecule composed of two identical atoms) has no permanent electric dipole moment: in formula (37), we have $d_0 = d_1 = ... = 0$. The vibration and rotation of the molecule

induce no coupling with the electromagnetic field, and the molecule is consequently "inactive" in the near infrared (vibration) and the microwave (rotation) domains. Like the vibration (*cf.* § 1-c-β of A_V), the rotation of the molecule can, however, be observed via the inelastic scattering of light (the Raman effect).

α. The Raman rotational effect. Classical treatment

We have already introduced, in Complement A_V, the susceptibility χ of a molecule in the optical domain: an incident light wave, whose electric field is $\mathbf{E}\,e^{i\Omega t}$, sets the electrons of the molecule in forced motion and causes an electric dipole $\mathbf{D}\,e^{i\Omega t}$, oscillating at the same frequency as the incident wave, to appear. χ is the coefficient of proportionality between $\mathbf{D}$ and $\mathbf{E}$. If $\mathbf{E}$ is parallel to the axis of the molecule, χ depends on the distance r between the two nuclei: when the molecule vibrates, χ vibrates at the same frequency. This is the origin of the Raman vibrational effect described in § 1-c-β of A_V.

Actually, a diatomic molecule is an anisotropic system. When the angle between the molecular axis and $\mathbf{E}$ is arbitrary, $\mathbf{D}$ is not generally parallel to $\mathbf{E}$: the relation between $\mathbf{D}$ and $\mathbf{E}$ is tensorial (χ is the "susceptibility tensor"). $\mathbf{D}$ is parallel to $\mathbf{E}$ in the two following simple cases: $\mathbf{E}$ parallel to the molecular axis (we then have $\chi = \chi_{/\!/}$), and $\mathbf{E}$ perpendicular to this axis ($\chi = \chi_\perp$). In the general case, we choose the Oz axis along the electric field $\mathbf{E}$ of the light wave (assumed to be polarized); we consider a molecule whose axis points in the direction of polar angles θ and φ and calculate the component along Oz of the dipole induced by $\mathbf{E}$ on this molecule. $\mathbf{E}$ can be decomposed into a component $\mathbf{E}_{/\!/}$, parallel to the molecular axis, and a component $\mathbf{E}_\perp$, perpendicular to M_1M_2 and contained in the plane formed by Oz and M_1M_2 (Fig. 7). The dipole induced on the molecule by the field $\mathbf{E}\cos\Omega t$ is then equal to:

$$\mathbf{D} = (\chi_{/\!/}\mathbf{E}_{/\!/} + \chi_\perp\mathbf{E}_\perp)\cos\Omega t \tag{40}$$

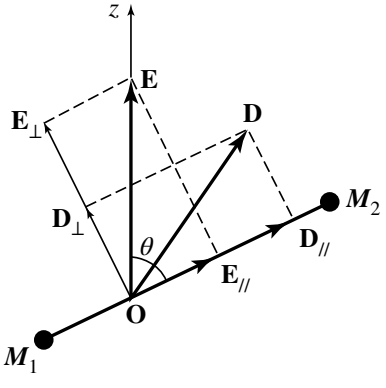


Figure 7: Decomposition of the electric field $\mathbf{E}$ into a component $\mathbf{E}_{/\!/}$ parallel to the molecular axis and a component $\mathbf{E}_\perp$ perpendicular to this axis. These fields induce electric dipoles $\chi_{/\!/}\mathbf{E}_{/\!/}$ and $\chi_\perp\mathbf{E}_\perp$ collinear with the corresponding fields. However, since $\chi_{/\!/}$ and $\chi_\perp$ have different values (the molecule is anisotropic), the induced electrical dipole $\mathbf{D} = \chi_{/\!/}\mathbf{E}_{/\!/} + \chi_\perp\mathbf{E}_\perp$ is not collinear with $\mathbf{E}$.

Its projection onto Oz can be calculated immediately:

$$\begin{aligned} D_z &= (\cos\theta\,\chi_{/\!/}|\mathbf{E}_{/\!/}| + \sin\theta\,\chi_\perp|\mathbf{E}_\perp|)\cos\,\Omega t \\ &= (\cos^2\theta\,\chi_{/\!/} + \sin^2\theta\,\chi_\perp)E\cos\,\Omega t \\ &= [\chi_\perp + (\chi_{/\!/} - \chi_\perp)\cos^2\theta]E\cos\,\Omega t \end{aligned} \tag{41}$$

We see that D_z depends on θ, since $\chi_{/\!/}$ and $\chi_\perp$ are not equal (the molecule is anisotropic).

To see what happens when the molecule rotates, we shall begin by reasoning classically. The fact that the molecule rotates at the frequency $\omega_R/2\pi$ means that $\cos\theta$ oscillates at the same frequency:

$$\cos\theta = \alpha\cos(\omega_R t - \beta) \tag{42}$$

where α and β depend on the initial conditions and the orientation of the angular momentum $\boldsymbol{\mathcal{L}}$ (which is fixed). Thus we see that the term in $\cos^2\theta$ of (41) gives rise to components of D_z that oscillate at frequencies of $(\Omega \pm 2\omega_R)/2\pi$, in addition to the component that varies at the frequency $\Omega/2\pi$. The fact that the rotation of the molecule at the frequency $\omega_R/2\pi$ modulates its polarizability at twice its frequency is easy to understand: after half a rotation, performed in half a period, the molecule returns to the same geometrical position with respect to the incident light wave. The light re-emitted with a polarization parallel to Oz is that which is radiated by D_z. It has an unshifted line of frequency $\Omega/2\pi$ (Rayleigh line), as well as two shifted lines, one on each side of the Rayleigh line, of frequencies $(\Omega - 2\omega_R)/2\pi$ (Raman-Stokes line) and $(\Omega + 2\omega_R)/2\pi$ (Raman-anti-Stokes line).

β. Quantum mechanical selection rules. Form of the Raman spectrum

Quantum mechanically, Raman scattering corresponds to an inelastic scattering process in which the molecule goes from level E_l to level $E_{l'}$, while the energy $\hbar\Omega$ of the photon becomes $\hbar\Omega + E_l - E_{l'}$ (the total energy of the system is conserved during this process).

The quantum theory of the Raman effect (which we shall not discuss here) indicates that the probability of such a process involves the matrix elements of $(\chi_{/\!/} - \chi_\perp)\cos^2\theta + \chi_\perp$ between the initial state $Y_l^m(\theta,\varphi)$ and the final state $Y_{l'}^{m'}(\theta,\varphi)$ of the molecule:

$$\int \mathrm{d}\Omega\; Y_{l'}^{m'*}(\theta,\varphi)\left[(\chi_{/\!/} - \chi_\perp)\cos^2\theta + \chi_\perp\right] Y_l^m(\theta,\varphi) \tag{43}$$

It can be shown, using the properties of the spherical harmonics, that such a matrix element is different from zero only if[5]:

$$l' - l = 0, +2, -2 \tag{44}$$

There is only one Rayleigh line (which corresponds to $l = l'$). Since the rotational levels are not equidistant, there are several Raman-anti-Stokes lines (which correspond to $l' = l - 2$), of frequencies:

$$\begin{cases} \dfrac{\Omega}{2\pi} + \dfrac{E_{l'+2} - E_{l'}}{h} = \dfrac{\Omega}{2\pi} + 4B\left(l' + \dfrac{3}{2}\right) \\ \text{with} \quad l' = 0, 1, 2, \ldots \end{cases} \tag{45}$$

and several Raman-Stokes lines (which correspond to $l' = l + 2$), of frequencies:

$$\begin{cases} \dfrac{\Omega}{2\pi} + \dfrac{E_l - E_{l+2}}{h} = \dfrac{\Omega}{2\pi} - 4B\left(l + \dfrac{3}{2}\right) \\ \text{with} \quad l = 0, 1, 2, \ldots \end{cases} \tag{46}$$

[5]The integral (43) is also zero if $m \neq m'$. If we consider light re-emitted with a different polarization state from that of the incident light wave, we obtain the following selection rules for m: $\Delta m = 0, \pm 1, \pm 2$.

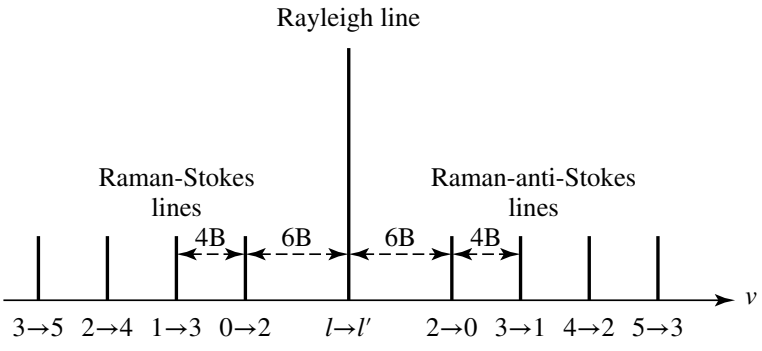


Figure 8: Raman rotational spectrum of a molecule. This molecule, initially in the rotational level E_l, inelastically scatters an incident photon of energy $\hbar\Omega$. After the scattering, the molecule has moved to the rotational level $E_{l'}$, and the energy of the photon is $\hbar\Omega + E_l - E_{l'}$ (conservation of energy).
If $l = l'$, the scattered photon has the same frequency $\nu = \Omega/2\pi$ as the incident photon; this yields the Rayleigh line. But it is also possible to have $l' - l = \pm 2$; if $l' = l + 2$, the frequency of the scattered photon is lower (Stokes scattering); if $l' = l - 2$, it is higher (anti-Stokes scattering). Since the rotational levels E_l are not equidistant (cf. Fig. 2), there are as many Stokes or anti-Stokes lines as there are values of l. These lines are labeled in the figure by $l \longrightarrow l'$ (with $l' = l \pm 2$).

The form of the Raman rotational spectrum is shown in Figure 8. The Stokes and anti-Stokes lines occur symmetrically with respect to the Rayleigh line. The separation between two adjacent Stokes (or anti-Stokes) lines is equal to $4\,B$, that is, to *twice* the separation which would be found between two adjacent lines of the pure rotational spectrum if it existed. Moreover, since the vibrational frequency is much larger than B, the Raman-Stokes and anti-Stokes vibrational lines are situated much further to the left and to the right of the Rayleigh line than the rotational Raman lines and hence do not appear in the figure (these lines also have rotational structures similar to that of Figure 8).

Comments:

(i) Consider a wave packet like those studied in § 3-c, that is, one for which the values of l are grouped about a very large integer l_M (Fig. 4). According to (45) and (46), the frequencies of the various Stokes and anti-Stokes lines will be very close (in relative value) to:

$$\frac{\Omega}{2\pi} \pm 4Bl_M \tag{47}$$

that is, according to (31):

$$\frac{\Omega}{2\pi} \pm 2\nu_M \tag{48}$$

where ν_M is the average rotational frequency of the molecule. Thus, the quantum mechanical treatment, at the classical limit, yields the results of § 4-b-α.

(*ii*) In Raman rotational spectra, the Stokes and anti-Stokes lines appear with comparable intensities since levels of large l have large populations, as hB is much smaller than kT. This is necessary for the observation of anti-Stokes lines, for which the initial state of the molecule must be at least $l = 2$. However, the anti-Stokes vibrational line has a much smaller intensity than the Stokes line. The vibrational energy is much larger than kT; the population of the vibrational ground state $v = 0$ is much larger than the others, and Stokes processes $v = 0 \longrightarrow v = 1$ are much more frequent than anti-Stokes processes $v = 1 \longrightarrow v = 0$.

(*iii*) The Raman rotational effect also exists for heteropolar molecules.

References and suggestions for further reading:

Karplus and Porter (12.1), § 7.4; Herzberg (12.4), Vol. I, Chap. III, §§ 1 and 2; Landau and Lifshitz (1.19), Chaps. XI and XIII; Townes and Schawlow (12.10), Chaps. 1 to 4.

Complement D_{VI}

Angular momentum of stationary states of a two-dimensional harmonic oscillator

1 Introduction 755
1-a Review of the classical problem 755
1-b The problem in quantum mechanics 757
2 Classification of the stationary states by the quantum numbers n_x and n_y 759
2-a Energies; stationary states 759
2-b H_{xy} does not constitue a C.S.C.O. in $\mathcal{E}_{xy}$ 760
3 Classification of the stationary states in terms of their angular momenta 761
3-a Significance and properties of the operator L_z 761
3-b Right and left circular quanta 761
3-c Stationary states of well-defined angular momentum 762
3-d Wave functions associated with the eigenstates common to H_{xy} and L_z 763
4 Quasi-classical states 765
4-a Definition of the states $|\alpha_x, \alpha_y\rangle$ and $|\alpha_r, \alpha_l\rangle$ 766
4-b Mean values and root mean square deviations of the various observables 767

In this complement, we shall be concerned with the quantum mechanical properties of a two-dimensional harmonic oscillator. The quantum mechanical problem is exactly soluble and does not involve complicated calculations. Furthermore, this subject provides an opportunity to study a simple application of the properties of the orbital angular momentum $\mathbf{L}$, since, as we shall see, the stationary states of such an oscillator can be classified with respect to the possible values of the observable L_z. In addition, the results obtained will be useful in the subsequent Complement E_{VI}.

1. Introduction

1-a. Review of the classical problem

A physical particle always moves in three-dimensional space. However, if its potential energy depends only on x and y, the problem can be treated in two dimensions. We shall assume here that this potential energy can be written:

$$V(x, y) = \frac{\mu}{2}\omega^2(x^2 + y^2) \tag{1}$$

where μ is the mass of the particle and ω is a constant. The classical Hamiltonian of the system is then:

$$\mathcal{H} = \mathcal{H}_{xy} + \mathcal{H}_z \tag{2}$$

with:

$$\mathcal{H}_{xy} = \frac{1}{2\mu}(p_x^2 + p_y^2) + \frac{1}{2}\mu\omega^2(x^2 + y^2)$$
$$\mathcal{H}_z = \frac{1}{2\mu}p_z^2 \tag{3}$$

where p_x, p_y, p_z are the three components of the momentum $\mathbf{p}$ of the particle. $\mathcal{H}_{xy}$ is a two-dimensional harmonic oscillator Hamiltonian.

The equations of motion can easily be integrated to yield:

$$\begin{cases} p_z(t) = p_0 \\ z(t) = \dfrac{p_0}{\mu}\, t + z_0 \end{cases} \tag{4}$$

$$\begin{cases} x(t) = x_M\, \cos(\omega t - \varphi_x) \\ p_x(t) = -\mu\omega\, x_M\, \sin(\omega t - \varphi_x) \end{cases} \tag{5}$$

$$\begin{cases} y(t) = y_M \cos(\omega t - \varphi_y) \\ p_y(t) = -\mu\omega\ y_M \sin(\omega t - \varphi_y) \end{cases} \tag{6}$$

where p_0, z_0, x_M, φ_x, y_M, φ_y are constants which depend on the initial conditions (we assume x_M and y_M to be positive).

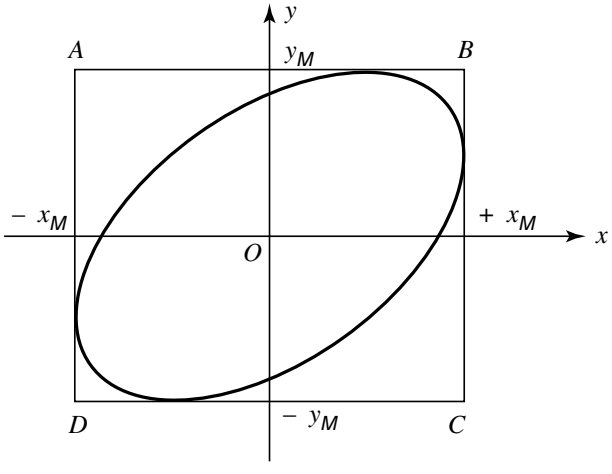


Figure 1: Projection of the classical trajectory of a particle in a two-dimensional harmonic potential onto the xOy plane; we obtain an ellipse inscribed in the rectangle ABCD.

We see that the projection of the particle onto Oz describes a uniform motion with a velocity of p_0/μ. The projection onto the xOy plane describes an ellipse inscribed in the rectangle $ABCD$ of Figure 1. The direction the particle takes on this ellipse depends on the phase difference $\varphi_y - \varphi_x$. When $\varphi_y - \varphi_x = \pm\pi$, the ellipse reduces to the line AC. When $\varphi_y - \varphi_x$ is between $-\pi$ and 0, the particle moves clockwise on the ellipse ("left-handed" motion), with the axes of the ellipse parallel to Ox and Oy for $\varphi_y - \varphi_x = -\pi/2$. When $\varphi_y - \varphi_x = 0$, the ellipse reduces to the line BD. Finally, when $\varphi_y - \varphi_x$ is between 0

and π, the particle moves counterclockwise on the ellipse ("right-handed" motion), with the axes parallel to Ox and Oy for $\varphi_y - \varphi_x = \pm\pi/2$. Note that the ellipse reduces to a circle if $\varphi_y - \varphi_x = \pm\pi/2$ and $x_M = y_M$.

It is easy to determine several constants of the motion related to the projection of the motion onto the xOy plane:

– the total energy $\mathcal{H}_{xy}$, which, according to (3), (5), (6), is equal to:

$$\mathcal{H}_{xy} = \frac{1}{2}\mu\omega^2(x_M^2 + y_M^2) \tag{7}$$

– the energies:

$$\mathcal{H}_x = \frac{1}{2}\mu\omega^2 x_M^2 \tag{8a}$$

$$\mathcal{H}_y = \frac{1}{2}\mu\omega^2 y_M^2 \tag{8b}$$

of the projections of the motion onto Ox and Oy;

– the component of the orbital angular momentum $\boldsymbol{\mathcal{L}}$ of the particle along Oz:

$$\mathcal{L}_z = xp_y - yp_x \tag{9}$$

which, according to (5) and (6), is equal to:

$$\mathcal{L}_z = \mu\omega\, x_M y_M\, \sin(\varphi_y - \varphi_x) \tag{10}$$

We see that $\mathcal{L}_z$ is positive or negative depending on whether the motion is counterclockwise ($0 < \varphi_y - \varphi_x < \pi$) or clockwise ($-\pi < \varphi_y - \varphi_x < 0$). $\mathcal{L}_z$ is zero for the two rectilinear motions ($\varphi_y - \varphi_x = \pm\pi$ and $\varphi_y - \varphi_x = 0$). Finally, for a motion at a given energy, that is, according to (7), for a fixed value of $x_M^2 + y_M^2$, $|\mathcal{L}_z|$ is maximal when $\varphi_y - \varphi_x = \pm\pi/2$ and the product $x_M y_M$ is maximal, which implies $x_M = y_M$. Of all motions at a given energy, it is the counterclockwise (clockwise) motion which corresponds to the maximal (minimal) algebraic value of $\mathcal{L}_z$.

1-b. The problem in quantum mechanics

The quantization rules of Chapter III enable us to obtain H, H_{xy}, H_z from $\mathcal{H}$, $\mathcal{H}_{xy}$, $\mathcal{H}_z$. The stationary states $|\varphi\rangle$ of the particle are given by:

$$H|\varphi\rangle = (H_{xy} + H_z)|\varphi\rangle = E|\varphi\rangle \tag{11}$$

with:

$$H_{xy} = \frac{P_x^2 + P_y^2}{2\mu} + \frac{1}{2}\mu\omega^2(X^2 + Y^2) \tag{12a}$$

$$H_z = \frac{P_z^2}{2\mu} \tag{12b}$$

According to the results of Complement $\mathrm{F_I}$, we know that we can choose a basis of eigenstates of H composed of vectors of the form:

$$|\varphi\rangle = |\varphi_{xy}\rangle \otimes |\varphi_z\rangle \tag{13}$$

where $|\varphi_{xy}\rangle$ is an eigenvector of H_{xy} in the state space $\mathcal{E}_{xy}$ associated with the variables x and y:

$$H_{xy}|\varphi_{xy}\rangle = E_{xy}|\varphi_{xy}\rangle \tag{14}$$

and $|\varphi_z\rangle$ is an eigenvector of H_z in the space $\mathcal{E}_z$ associated with the variable z:

$$H_z|\varphi_z\rangle = E_z|\varphi_z\rangle \tag{15}$$

The total energy associated with the state (13) is then:

$$E = E_{xy} + E_z \tag{16}$$

Now, equation (15), which in fact describes the stationary states of a free particle in a one-dimensional problem, can be solved immediately; it yields:

$$\langle z|\varphi_z\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, \mathrm{e}^{ip_z z/\hbar} \tag{17}$$

(where p_z is an arbitrary real constant), with:

$$E_z = \frac{p_z^2}{2\mu} \tag{18}$$

The problem therefore reduces to the determination of the solutions of equation (14), that is, the energies and stationary states of a two-dimensional harmonic oscillator. This is the problem we shall now try to solve.

We shall see that the eigenvalues E_{xy} of H_{xy} are degenerate: H_{xy} alone does not constitute a C.S.C.O. in $\mathcal{E}_{xy}$. We must therefore add one or several other observables to H_{xy} in order to construct a C.S.C.O. In fact, we find in quantum mechanics the same constants of the motion as in classical mechanics: H_x and H_y, the energies of the projection of the motion onto Ox and Oy; and L_z, the component along Oz of the orbital angular momentum $\mathbf{L}$. Since L_z commutes with neither H_x nor H_y, we shall see that a C.S.C.O. can be formed of H_{xy}, H_x and H_y (§ 2) or of H_{xy} and L_z (§ 3).

Comments:

(*i*) Formula (18) indicates that the eigenvalues E_z of H_z are all two-fold degenerate in the space $\mathcal{E}_z$. Furthermore, the degeneracy in $\mathcal{E} = \mathcal{E}_{xy} \otimes \mathcal{E}_z$ of the eigenvalues (16) of the total Hamiltonian H is not due solely to the degeneracy of E_{xy} in $\mathcal{E}_{xy}$ and of E_z in $\mathcal{E}_z$: two eigenvectors of H of the form (13) can have the same total energy E without their corresponding values of E_{xy} (and of E_z) being equal.

(*ii*) H commutes with the component L_z of $\mathbf{L}$, but not with L_x and L_y. This results from the fact that the potential energy written in (1) is rotation-invariant only about Oz. Moreover, of the three operators L_x, L_y and L_z, only one, L_z, acts only in $\mathcal{E}_{xy}$. In the study of the two-dimensional harmonic oscillator, therefore, we shall use only the observable L_z. In Complement B_{VII}, we shall study the isotropic three-dimensional harmonic oscillator, whose potential energy is invariant with respect to any rotation about an axis which passes through the origin; we shall see that all the components of $\mathbf{L}$ then commute with the Hamiltonian.

2. Classification of the stationary states by the quantum numbers n_x and n_y

2-a. Energies; stationary states

To obtain the solutions of the eigenvalue equation (14), note that H_{xy}, can be written:

$$H_{xy} = H_x + H_y \tag{19}$$

where H_x and H_y are both Hamiltonians of one-dimensional harmonic oscillators:

$$\begin{aligned} H_x &= \frac{P_x^2}{2\mu} + \frac{1}{2}\mu\omega^2 X^2 \\ H_y &= \frac{P_y^2}{2\mu} + \frac{1}{2}\mu\omega^2 Y^2 \end{aligned} \tag{20}$$

We know the eigenstates $|\varphi_{n_x}\rangle$ of H_x in $\mathcal{E}_x$ and the eigenstates $|\varphi_{n_y}\rangle$ of H_y in $\mathcal{E}_y$. Their energies are, respectively, $E_x = (n_x + 1/2)\hbar\omega$ and $E_y = (n_y + 1/2)\hbar\omega$ (where n_x and n_y are positive integers or zero). The eigenstates of H_{xy} can thus be chosen in the form:

$$|\varphi_{n_x,n_y}\rangle = |\varphi_{n_x}\rangle \otimes |\varphi_{n_y}\rangle \tag{21}$$

where the corresponding energy E_{xy} is given by:

$$\begin{aligned} E_{xy} &= \left(n_x + \frac{1}{2}\right)\hbar\omega + \left(n_y + \frac{1}{2}\right)\hbar\omega \\ &= (n_x + n_y + 1)\hbar\omega \end{aligned} \tag{22}$$

According to the properties of the one-dimensional harmonic oscillator, E_x is non-degenerate in $\mathcal{E}_x$, and E_y in $\mathcal{E}_y$. Consequently, a vector $|\varphi_{n_x,n_y}\rangle$ of $\mathcal{E}_{xy}$, which is unique to within a constant factor, corresponds to a pair $\{n_x, n_y\}$: H_x and H_y form a C.S.C.O. in $\mathcal{E}_{xy}$.

It will prove convenient to use the operators a_x and a_y (destruction operators of a quantum, relative to Ox and Oy respectively), defined by:

$$\begin{aligned} a_x &= \frac{1}{\sqrt{2}}\left(\beta X + i\frac{P_x}{\beta\hbar}\right) \\ a_y &= \frac{1}{\sqrt{2}}\left(\beta Y + i\frac{P_y}{\beta\hbar}\right) \end{aligned} \tag{23}$$

with:

$$\beta = \sqrt{\frac{\mu\omega}{\hbar}} \tag{24}$$

Since a_x and a_y act in different spaces, $\mathcal{E}_x$ and $\mathcal{E}_y$, the only non-zero commutators between the four operators a_x, a_y, $a_x^\dagger$, $a_y^\dagger$, are:

$$\left[a_x, a_x^\dagger\right] = \left[a_y, a_y^\dagger\right] = 1 \tag{25}$$

The operators N_x (the number of quanta relative to the Ox axis) and N_y (the number of quanta relative to the Oy axis) are given by:

$$N_x = a_x^\dagger a_x$$
$$N_y = a_y^\dagger a_y \tag{26}$$

which enables us to write H_{xy} in the form:

$$H_{xy} = H_x + H_y = (N_x + N_y + 1)\hbar\omega \tag{27}$$

We have, obviously:

$$N_x|\varphi_{n_x,n_y}\rangle = n_x|\varphi_{n_x,n_y}\rangle$$
$$N_y|\varphi_{n_x,n_y}\rangle = n_y|\varphi_{n_x,n_y}\rangle \tag{28}$$

The ground state $|\varphi_{0,0}\rangle$ is given by:

$$|\varphi_{0,0}\rangle = |\varphi_{n_x=0}\rangle \otimes |\varphi_{n_y=0}\rangle \tag{29}$$

The state $|\varphi_{n_x,n_y}\rangle$ defined by (21) can be obtained from $|\varphi_{0,0}\rangle$ by the successive application of the operators $a_x^\dagger$ and $a_y^\dagger$:

$$|\varphi_{n_x,n_y}\rangle = \frac{1}{\sqrt{n_x!n_y!}}(a_x^\dagger)^{n_x}(a_y^\dagger)^{n_y}|\varphi_{0,0}\rangle \tag{30}$$

The corresponding wave function is the product of $\varphi_{n_x}(x)$ and $\varphi_{n_y}(y)$ [*cf.* Complement B_V, formula (35)]:

$$\varphi_{n_x,n_y}(x,y) = \frac{\beta}{\sqrt{\pi(2)^{n_x+n_y}(n_x)!(n_y)!}}\ \mathrm{e}^{-\beta^2(x^2+y^2)/2}H_{n_x}(\beta x)\,H_{n_y}(\beta y) \tag{31}$$

where the H_n are the Hermite polynomials (Complement B_V, § 1).

2-b. H_{xy} does not constitue a C.S.C.O. in $\mathcal{E}_{xy}$

We see from (22) that the eigenvalues of H_{xy} are of the form:

$$E_{xy} = E_n = (n+1)\hbar\omega \tag{32}$$

where:

$$n = n_x + n_y \tag{33}$$

is a positive integer or zero. To each value of the energy correspond the various orthogonal eigenvectors:

$$|\varphi_{n_x=n,\,n_y=0}\rangle,\ |\varphi_{n_x=n-1,\,n_y=1}\rangle,\ \ldots,|\varphi_{n_x=0,\,n_y=n}\rangle \tag{34}$$

Since there are $(n+1)$ of these vectors, the eigenvalue E_n is $(n+1)$-fold degenerate in $\mathcal{E}_{xy}$. H_{xy} alone does not, therefore, constitute a C.S.C.O. On the other hand, we have seen that $\{H_x,\, H_y\}$ is a C.S.C.O.; this is also, obviously, true of $\{H_{xy},\, H_x\}$ et $\{H_{xy},\, H_y\}$.

3. Classification of the stationary states in terms of their angular momenta

3-a. Significance and properties of the operator L_z

In the preceding section, we identified the stationary states by the quantum numbers n_x and n_y. But the Ox and Oy axes do not enjoy a privileged position in this problem. Since the potential energy is invariant under rotation about Oz, we could just as well have chosen another system of orthogonal axes Ox' and Oy' in the xOy plane; we would have then obtained stationary states different from the preceding ones.

Therefore, in order to take better advantage of the symmetry of the problem, we shall now consider the component L_z of the angular momentum, defined by:

$$L_z = XP_y - YP_x \tag{35}$$

Expressing X and P_x in terms of a_x and $a_x^\dagger$, and Y and P_y in terms of a_y and $a_y^\dagger$, we get:

$$L_z = i\hbar \left(a_x a_y^\dagger - a_x^\dagger a_y\right) \tag{36}$$

Now, the expression for H_{xy} in terms of the same operators is:

$$H_{xy} = \left(a_x^\dagger a_x + a_y^\dagger a_y + 1\right) \hbar\omega \tag{37}$$

Since:

$$\begin{aligned} \left[a_x a_y^\dagger,\, a_x^\dagger a_x + a_y^\dagger a_y\right] &= a_x a_y^\dagger - a_x a_y^\dagger = 0 \\ \left[a_x^\dagger a_y,\, a_x^\dagger a_x + a_y^\dagger a_y\right] &= -a_x^\dagger a_y + a_x^\dagger a_y = 0 \end{aligned} \tag{38}$$

we find that:

$$[H_{xy},\, L_z] = 0 \tag{39}$$

We shall therefore look for a basis of eigenvectors common to H_{xy} and L_z.

3-b. Right and left circular quanta

We introduce the operators a_r and a_l defined by:

$$\begin{aligned} a_r &= \frac{1}{\sqrt{2}}(a_x - i a_y) \\ a_l &= \frac{1}{\sqrt{2}}(a_x + i a_y) \end{aligned} \tag{40}$$

We see from this definition that the action of a_r (or a_l) on $|\varphi_{n_x,\, n_y}\rangle$ yields a state which is a linear combination of $|\varphi_{n_x-1,\, n_y}\rangle$ and $|\varphi_{n_x,\, n_y-1}\rangle$, that is, a stationary state which has one less energy quantum $\hbar\omega$. Similarly, the action of $a_r^\dagger$ (or $a_l^\dagger$) on $|\varphi_{n_x,\, n_y}\rangle$ yields another stationary state which has one more energy quantum. In fact, we shall see that a_r (or a_l) is similar to a_x (or a_y), and that a_r and a_l can be interpreted as being destruction operators of a right and left "circular quantum" respectively.

First of all, using (40) and (25), it is simple to verify that the only non-zero commutators between the four operators a_r, a_l, $a_r^\dagger$, $a_l^\dagger$ are:

$$[a_r,\, a_r^\dagger] = [a_l,\, a_l^\dagger] = 1 \tag{41}$$

These relations are indeed analogous to (25). Moreover, H_{xy} can be written, in terms of these operators, in a way that is similar to (37); since:

$$\begin{aligned} a_r^\dagger a_r &= \frac{1}{2}\left(a_x^\dagger a_x + a_y^\dagger a_y - i a_x^\dagger a_y + i a_x a_y^\dagger\right) \\ a_l^\dagger a_l &= \frac{1}{2}\left(a_x^\dagger a_x + a_y^\dagger a_y + i a_x^\dagger a_y - i a_x a_y^\dagger\right) \end{aligned} \tag{42}$$

we have:

$$H_{xy} = \left(a_r^\dagger a_r + a_l^\dagger a_l + 1\right)\hbar\omega \tag{43}$$

In addition, using (36), we see that:

$$L_z = \hbar\left(a_r^\dagger a_r - a_l^\dagger a_l\right) \tag{44}$$

If we introduce the operators N_r and N_l (the number of right and left "circular quanta"):

$$\begin{aligned} N_r &= a_r^\dagger a_r \\ N_l &= a_l^\dagger a_l \end{aligned} \tag{45}$$

formulas (43) and (44) become:

$$\begin{aligned} H_{xy} &= (N_r + N_l + 1)\hbar\omega \\ L_z &= \hbar(N_r - N_l) \end{aligned} \tag{46}$$

Thus, while maintaining H in a form as simple as (27), we have simplified that of L_z.

3-c. Stationary states of well-defined angular momentum

Using the operators a_r and a_l, we can now go through the same arguments we used for a_x and a_y. It follows that the spectra of N_r and N_l are composed of all positive integers and zero. In addition, specifying a pair $\{n_r,\, n_l\}$ of such integers determines uniquely (to within a constant factor) the eigenvector common to N_r and N_l, associated with these eigenvalues, which is written:

$$|\chi_{n_r,\,n_l}\rangle = \frac{1}{\sqrt{(n_r)!(n_l)!}}(a_r^\dagger)^{n_r}(a_l^\dagger)^{n_l}|\varphi_{0,\,0}\rangle \tag{47}$$

N_r and N_l therefore form a C.S.C.O. in $\mathcal{E}_{xy}$. Thus we see, by using (46), that $|\chi_{n_r,\,n_l}\rangle$ is also an eigenvector of H_{xy} and of L_z, with the eigenvalues $(n+1)\hbar\omega$ and $m\hbar$, where n and m are given by:

$$\begin{aligned} n &= n_r + n_l \\ m &= n_r - n_l \end{aligned} \tag{48}$$

Equations (48) enable us to understand the origin of the name of right or left "circular quanta". The action of the operator $a_r^\dagger$ on $|\chi_{n_r,\,n_l}\rangle$ yields a state with one more quantum, to which, since m has increased by one, an additional angular momentum $+\hbar$ must be attributed (this corresponds to a counterclockwise rotation about Oz). Similarly, $a_l^\dagger$ yields a state with one more quantum, of angular momentum $-\hbar$ (clockwise rotation).

Since n_r and n_l are positive integers (or zero), our results are in agreement with those of the preceding section: the eigenvalues of H_{xy} are of the form $(n+1)\hbar\omega$, where n is a positive integer or zero; their degree of degeneracy is $(n+1)$ since, for fixed n, we can have:

$$\begin{array}{ll} n_r = n & ;\ n_l = 0 \\ n_r = n-1 & ;\ n_l = 1 \\ \quad\vdots & \\ n_r = 0 & ;\ n_l = n \end{array} \tag{49}$$

Furthermore, we see that the eigenvalues of L_z are of the form $m\hbar$, where m is a positive or negative integer or zero, which is the result that was established for the general case in Chapter VI. In addition, table (49) tells us which values of m are associated with a given value of n. For example, for the ground state, we have $n_r - n_l = 0$, and therefore, necessarily, $m = 0$; for the first excited state, we can have $n_r = 1$ and $n_l = 0$, or $n_r = 0$ and $n_l = 1$, which yields either $m = +1$ or $m = -1$. In general, formulas (48) and (49) show that, for a given energy level $(n+1)\hbar\omega$, the possible values of m are:

$$m = n,\ n-2,\ n-4,\ ...,\ -n+2,\ -n \tag{50}$$

It follows that, to a pair of values of n and m, there corresponds a single vector (to within a constant factor):

$$|\chi_{n_r=\frac{n+m}{2},\,n_l=\frac{n-m}{2}}\rangle$$

H_{xy} and L_z therefore form a C.S.C.O. in $\mathcal{E}_{xy}$.

Comment:

For a given value of the total energy (labeled by n), the states $|\chi_{n_r=n,\,n_l=0}\rangle$ and $|\chi_{n_r=0,\,n_l=n}\rangle$ correspond to the maximal $(n\hbar)$ and minimal $(-n\hbar)$ values of L_z. These states therefore recall the classical right and left circular motions associated with a given value of the total energy, for which $\mathcal{L}_z$ takes on its maximal and minimal values (see § 1-a).

3-d. Wave functions associated with the eigenstates common to H_{xy} and L_z

To conserve the symmetry of the problem with respect to rotation about Oz, we shall use polar coordinates, setting:

$$\begin{array}{ll} x = \rho\cos\varphi & \rho \geq 0 \\ y = \rho\sin\varphi & 0 \leq \varphi < 2\pi \end{array} \tag{51}$$

Now, what is the action of the operators a_r and a_l on a function of ρ and φ? We shall begin by determining their action on a function of x and y. Knowing that of X and P_x and therefore that of a_x (and, by analogy, that of a_y), we can use (40), which yields:

$$a_r \Longrightarrow \frac{1}{2}\left[\beta(x - iy) + \frac{1}{\beta}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\right] \tag{52}$$

According to the rules for differentiating functions of several variables, we then obtain:

$$a_r \Longrightarrow \frac{\mathrm{e}^{-i\varphi}}{2}\left[\beta\rho + \frac{1}{\beta}\frac{\partial}{\partial\rho} - \frac{i}{\beta\rho}\frac{\partial}{\partial\varphi}\right] \tag{53}$$

Similarly:

$$a_r^\dagger \Longrightarrow \frac{\mathrm{e}^{i\varphi}}{2}\left[\beta\rho - \frac{1}{\beta}\frac{\partial}{\partial\rho} - \frac{i}{\beta\rho}\frac{\partial}{\partial\varphi}\right] \tag{54}$$

and:

$$\begin{aligned} a_l &\Longrightarrow \frac{\mathrm{e}^{i\varphi}}{2}\left[\beta\rho + \frac{1}{\beta}\frac{\partial}{\partial\rho} + \frac{i}{\beta\rho}\frac{\partial}{\partial\varphi}\right] \\ a_l^\dagger &\Longrightarrow \frac{\mathrm{e}^{-i\varphi}}{2}\left[\beta\rho - \frac{1}{\beta}\frac{\partial}{\partial\rho} + \frac{i}{\beta\rho}\frac{\partial}{\partial\varphi}\right] \end{aligned} \tag{55}$$

To calculate the wave functions $\chi_{n_r,\,n_l}(\rho,\,\varphi)$, simply apply the differential operators representing $a_r^\dagger$ and $a_l^\dagger$ to the function $\chi_{0,0}(\rho,\,\varphi)$, which is, according to (31):

$$\chi_{0,0}(\rho,\,\varphi) = \frac{\beta}{\sqrt{\pi}}\,\mathrm{e}^{-\beta^2\rho^2/2} \tag{56}$$

Now it can be seen from (54) and (55) that the action of $a_r^\dagger$ (or of $a_l^\dagger$) on a function of the form $\mathrm{e}^{im\varphi}F(\rho)$ is given by:

$$\begin{aligned} a_r^\dagger\left[\mathrm{e}^{im\varphi}F(\rho)\right] &= \frac{\mathrm{e}^{i(m+1)\varphi}}{2}\left[\left(\beta\rho + \frac{m}{\beta\rho}\right)F(\rho) - \frac{1}{\beta}\frac{\mathrm{d}F}{\mathrm{d}\rho}\right] \\ a_l^\dagger\left[\mathrm{e}^{im\varphi}F(\rho)\right] &= \frac{\mathrm{e}^{i(m-1)\varphi}}{2}\left[\left(\beta\rho - \frac{m}{\beta\rho}\right)F(\rho) - \frac{1}{\beta}\frac{\mathrm{d}F}{\mathrm{d}\rho}\right] \end{aligned} \tag{57}$$

Through the repeated application of these relations to (56), we see that the φ-dependence of $\chi_{n_r,\,n_l}(\rho,\,\varphi)$ is simply given by: $\mathrm{e}^{i(n_r - n_l)\varphi}$. This is a general result, established in Chapter VI: the φ-dependence of an eigenfunction of L_z of eigenvalue $m\hbar$ is $\mathrm{e}^{im\varphi}$.

If, in (57), we choose $F(\rho) = \rho^m\mathrm{e}^{-\beta^2\rho^2/2}$, then:

$$a_r^\dagger\left[\mathrm{e}^{im\varphi}\rho^m\,\mathrm{e}^{-\beta^2\rho^2/2}\right] = \beta\,\mathrm{e}^{i(m+1)\varphi}\rho^{m+1}\,\mathrm{e}^{-\beta^2\rho^2/2} \tag{58}$$

Applying the operator $a_r^\dagger$ to the function $\chi_{0,0}(\rho)$ n_r times, we obtain:

$$\chi_{n_r,0}(\rho,\,\varphi) = \frac{\beta}{\sqrt{\pi(n_r)!}}\,\mathrm{e}^{in_r\,\varphi}(\beta\rho)^{n_r}\,\mathrm{e}^{-\beta^2\rho^2/2} \tag{59}$$

An analogous calculation yields:

$$\chi_{0,\,n_l}(\rho,\,\varphi) = \frac{\beta}{\sqrt{\pi(n_l)!}}\,\mathrm{e}^{-in_l\,\varphi}(\beta\rho)^{n_l}\,\mathrm{e}^{-\beta^2\rho^2/2} \tag{60}$$

These wave functions are normalized. For a given energy level $(n+1)\hbar\omega$, the wave functions (59) and (60) correspond to the limiting values $+n$ and $-n$ of the quantum number m. Their ρ-dependence is particularly simple: their modulus reaches a maximum for $\rho = \sqrt{n}/\beta$. Therefore (as in the case of a one-dimensional harmonic oscillator), the spatial spread of these wave functions increases with the energy $(n+1)\hbar\omega$ with which they are associated.

In the same way, application of the operators $a_r^\dagger$ (or $a_l^\dagger$) to (59) and (60) permits the construction of the functions $\chi_{n_r,\,n_l}(\rho,\,\varphi)$ for any n_r and n_l. The results obtained for the first excited levels are given in Table I.

$n=0$	$m=0$	$\chi_{0,0}(\rho) = \dfrac{\beta}{\sqrt{\pi}}\,\mathrm{e}^{-\beta^2\rho^2/2}$
$n=1$	$m=1$	$\chi_{1,0}(\rho,\,\varphi) = \dfrac{\beta}{\sqrt{\pi}}\,\beta\rho\,\mathrm{e}^{-\beta^2\rho^2/2}\,\mathrm{e}^{i\varphi}$
	$m=-1$	$\chi_{0,1}(\rho,\,\varphi) = \dfrac{\beta}{\sqrt{\pi}}\,\beta\rho\,\mathrm{e}^{-\beta^2\rho^2/2}\,\mathrm{e}^{-i\varphi}$
$n=2$	$m=2$	$\chi_{2,0}(\rho,\,\varphi) = \dfrac{\beta}{\sqrt{2\pi}}\,(\beta\rho)^2\,\mathrm{e}^{-\beta^2\rho^2/2}\,\mathrm{e}^{2i\varphi}$
	$m=0$	$\chi_{1,1}(\rho,\,\varphi) = \dfrac{\beta}{\sqrt{\pi}}\left[(\beta\rho)^2-1\right]\,\mathrm{e}^{-\beta^2\rho^2/2}$
	$m=-2$	$\chi_{0,2}(\rho,\,\varphi) = \dfrac{\beta}{\sqrt{2\pi}}\,(\beta\rho)^2\,\mathrm{e}^{-\beta^2\rho^2/2}\,\mathrm{e}^{-2i\varphi}$

Table I: *Eigenfunctions common to the Hamiltonian H_{xy} and the observable L_z, for the first levels of the two-dimensional harmonic oscillator.*

Comment:

The functions $\chi_{n_r,\,0}(\rho,\,\varphi)$ given in (59) are proportional to $\mathrm{e}^{-\beta^2\rho^2/2}(\beta\rho\,\mathrm{e}^{i\varphi})^{n_r}$. More generally, all their linear combinations are of the form:

$$F(\rho,\,\varphi) = \mathrm{e}^{-\beta^2\rho^2/2}\,f(\beta\rho\,\mathrm{e}^{i\varphi}) \tag{61}$$

(where f is an arbitrary function of one variable) and are eigenfunctions of N_l with the eigenvalue zero. As expected, it can easily be shown from (55) that:

$$a_l\,F(\rho,\,\varphi) = 0 \tag{62}$$

Similarly, the subspace of eigenfunctions of N_r of eigenvalue zero is composed of functions of the form:

$$G(\rho,\,\varphi) = \mathrm{e}^{-\beta^2\rho^2/2}g(\beta\rho\,\mathrm{e}^{-i\varphi}) \tag{63}$$

4. Quasi-classical states

Using the properties of the one-dimensional harmonic oscillator, we can easily calculate the time evolution of the state vector and the mean values of the various observables

of the two-dimensional oscillator. For example, it is not difficult to show that in the mean values $\langle X\rangle(t)$ and $\langle Y\rangle(t)$, as well as $\langle P_x\rangle(t)$ and $\langle P_y\rangle(t)$, only the Bohr frequency ω appears. Moreover, it can be shown that these mean values exactly obey the classical equations of motion. In this section, we shall be concerned with the properties and evolution of the quasi-classical states of the two-dimensional harmonic oscillator.

4-a. Definition of the states $|\alpha_x, \alpha_y\rangle$ and $|\alpha_r, \alpha_l\rangle$

To construct a quasi-classical state of the two-dimensional harmonic oscillator, we can base our reasoning on the one-dimensional oscillator (*cf.* Complement G_V). Recall that, in a quasi-classical state associated with a given classical motion, the mean values $\langle X\rangle(t)$ and $\langle P\rangle(t)$ coincide at each instant with $x(t)$ and $p(t)$. Similarly, the mean value of the Hamiltonian H is equal (to within a half-quantum $\hbar\omega/2$) to the classical energy. We showed in Complement G_V that, at any time, the quasi-classical states are eigenstates of the destruction operator a and can be written:

$$|\alpha\rangle = \sum_n c_n(\alpha)|\varphi_n\rangle \tag{64}$$

where α is the eigenvalue of a, and:

$$c_n(\alpha) = \frac{\alpha^n}{\sqrt{n!}}\, \mathrm{e}^{-|\alpha|^2/2} \tag{65}$$

In the case which concerns us here, we can use the rules of the tensor product to obtain the quasi-classical states in the form:

$$|\alpha_x, \alpha_y\rangle = |\alpha_x\rangle \otimes |\alpha_y\rangle = \sum_{n_x=0}^{\infty} \sum_{n_y=0}^{\infty} c_{n_x}(\alpha_x)\, c_{n_y}(\alpha_y)|\varphi_{n_x, n_y}\rangle \tag{66}$$

with:

$$\begin{aligned} a_x|\alpha_x,\, \alpha_y\rangle &= \alpha_x|\alpha_x,\, \alpha_y\rangle \\ a_y|\alpha_x,\, \alpha_y\rangle &= \alpha_y|\alpha_x,\, \alpha_y\rangle \end{aligned} \tag{67}$$

We are then sure that $\langle X\rangle$, $\langle P_x\rangle$, $\langle H_x\rangle$, $\langle Y\rangle$, $\langle P_y\rangle$, $\langle H_y\rangle$ are the same as the corresponding classical quantities. Now, returning to definition (40) and using (67), we see that:

$$\begin{aligned} a_r|\alpha_x,\, \alpha_y\rangle &= \alpha_r|\alpha_x,\, \alpha_y\rangle \\ a_l|\alpha_x,\, \alpha_y\rangle &= \alpha_l|\alpha_x,\, \alpha_y\rangle \end{aligned} \tag{68}$$

with:

$$\begin{aligned} \alpha_r &= \frac{1}{\sqrt{2}}(\alpha_x - i\alpha_y) \\ a_l &= \frac{1}{\sqrt{2}}(\alpha_x + i\alpha_y) \end{aligned} \tag{69}$$

Therefore, the state $|\alpha_x,\, \alpha_y\rangle$ is also an eigenvector of a_r and a_l with the eigenvalues given in (69). We shall denote by $|\alpha_r,\, \alpha_l\rangle$ the eigenvector common to a_r and a_l associated

with the eigenvalues α_r and α_l. It is easy to show that the expansion of $|\alpha_r,\, \alpha_l\rangle$ on the $\{|\chi_{n_r,\, n_l}\rangle\}$ basis has the same form as that of $|\alpha_x,\, \alpha_y\rangle$ on the $\{|\varphi_{n_x,\, n_y}\rangle\}$ basis:

$$|\alpha_r, \alpha_l\rangle = \sum_{n_r=0}^{\infty} \sum_{n_l=0}^{\infty} c_{n_r}(\alpha_r)\, c_{n_l}(\alpha_l) |\chi_{n_r,\, n_l}\rangle \tag{70}$$

where the coefficients c_n are given by (65). It follows from (68) and (69) that:

$$|\alpha_x, \alpha_y\rangle = \left|\alpha_r = \frac{\alpha_x - i\alpha_y}{\sqrt{2}},\ \alpha_l = \frac{\alpha_x + i\alpha_y}{\sqrt{2}}\right\rangle \tag{71}$$

Because of the properties of the states $|\alpha\rangle$ (*cf.* Complement G_V, § 3-a), we see that if:

$$|\psi(0)\rangle = |\alpha_x,\, \alpha_y\rangle = |\alpha_r,\, \alpha_l\rangle \tag{72}$$

the state vector at the instant t will be:

$$\begin{aligned} |\psi(t)\rangle &= \mathrm{e}^{-i\omega t} |\mathrm{e}^{-i\omega t}\alpha_x,\, \mathrm{e}^{-i\omega t}\alpha_y\rangle \\ &= \mathrm{e}^{-i\omega t} |\mathrm{e}^{-i\omega t}\alpha_r,\, \mathrm{e}^{-i\omega t}\alpha_l\rangle \end{aligned} \tag{73}$$

4-b. Mean values and root mean square deviations of the various observables

We set:

$$\begin{aligned} \alpha_x &= |\alpha_x|\, \mathrm{e}^{i\varphi_x} \\ \alpha_y &= |\alpha_y|\, \mathrm{e}^{i\varphi_y} \end{aligned} \tag{74}$$

Using formulas (93) of Complement G_V, we obtain:

$$\begin{cases} \langle X\rangle(t) = \dfrac{\sqrt{2}}{\beta} |\alpha_x| \cos(\omega t - \varphi_x) \\ \langle Y\rangle(t) = \dfrac{\sqrt{2}}{\beta} |\alpha_y| \cos(\omega t - \varphi_y) \end{cases} \tag{75}$$

$$\begin{cases} \langle P_x\rangle(t) = -\mu\omega \dfrac{\sqrt{2}}{\beta} |\alpha_x| \sin(\omega t - \varphi_x) \\ \langle P_y\rangle(t) = -\mu\omega \dfrac{\sqrt{2}}{\beta} |\alpha_y| \sin(\omega t - \varphi_y) \end{cases} \tag{76}$$

Comparing (75) and (76) with (5) and (6), we see that:

$$\begin{aligned} \alpha_x &= \frac{\beta x_M}{\sqrt{2}} \mathrm{e}^{i\varphi_x} \\ \alpha_y &= \frac{\beta y_M}{\sqrt{2}} \mathrm{e}^{i\varphi_y} \end{aligned} \tag{77}$$

where x_M, φ_x, y_M, φ_y are the parameters defining the classical motion which the state $|\alpha_x,\, \alpha_y\rangle$ best reproduces.

Also:

$$\langle N_x \rangle = |\alpha_x|^2$$
$$\langle N_y \rangle = |\alpha_y|^2 \tag{78}$$

and:

$$\langle N_r \rangle = |\alpha_r|^2 = \frac{1}{2}\left[|\alpha_x|^2 + |\alpha_y|^2 + i\left(\alpha_x \alpha_y^* - \alpha_x^* \alpha_y\right)\right]$$
$$\langle N_l \rangle = |\alpha_l|^2 = \frac{1}{2}\left[|\alpha_x|^2 + |\alpha_y|^2 - i\left(\alpha_x \alpha_y^* - \alpha_x^* \alpha_y\right)\right] \tag{79}$$

that is, according to (46):

$$\langle H_{xy} \rangle = \hbar\omega\left(|\alpha_x|^2 + |\alpha_y|^2 + 1\right) = \hbar\omega\left(|\alpha_r|^2 + |\alpha_l|^2 + 1\right) \tag{80}$$

and:

$$\langle L_z \rangle = 2\hbar|\alpha_x||\alpha_y|\sin(\varphi_y - \varphi_x) = \hbar\left(|\alpha_r|^2 - |\alpha_l|^2\right) \tag{81}$$

According to (77), $\langle L_z \rangle$ is the same as the classical value of $\mathcal{L}_z$ [formula (10)].

Now let us consider the root mean square deviations of the position and momentum and then of the energy and angular momentum in a state $|\alpha_x, \alpha_y\rangle$. Directly applying the results of Complement G_V, we obtain:

$$\Delta X = \Delta Y = \frac{1}{\beta\sqrt{2}}$$
$$\Delta P_x = \Delta P_y = \frac{\mu\omega}{\beta\sqrt{2}} \tag{82}$$

The root mean square deviations of the position and momentum are independent of α_x and α_y; if $|\alpha_x|$ and $|\alpha_y|$ are much greater than 1, the position and momentum of the oscillator have a very small spread about $\langle X \rangle$, $\langle Y \rangle$ and $\langle P_x \rangle$, $\langle P_y \rangle$.

Finally, let us calculate the root mean square deviations ΔH_{xy} for the energy and ΔL_z for the angular momentum. As in Complement G_V:

$$\Delta N_x = |\alpha_x|$$
$$\Delta N_y = |\alpha_y|$$
$$\Delta N_r = |\alpha_r|$$
$$\Delta N_l = |\alpha_l| \tag{83}$$

But the Hamiltonian H_{xy} involves $N = N_x + N_y$, and L_z is proportional to $N_r - N_l$. We must now calculate, for example:

$$\begin{aligned}(\Delta N)^2 &= \langle (N_x + N_y)^2 \rangle - \langle (N_x + N_y) \rangle^2 \\ &= (\Delta N_x)^2 + (\Delta N_y)^2 + 2[\langle N_x N_y \rangle - \langle N_x \rangle \langle N_y \rangle]\end{aligned} \tag{84}$$

According to (66), the state of the system is a tensor product, which means that the observables N_x and N_y are not correlated:

$$\langle N_x N_y \rangle = \langle N_x \rangle \langle N_y \rangle \tag{85}$$

It follows that:

$$(\Delta N)^2 = (\Delta N_x)^2 + (\Delta N_y)^2 \tag{86}$$

that is:

$$\Delta H_{xy} = \hbar\omega\sqrt{|\alpha_x|^2 + |\alpha_y|^2} = \hbar\omega\sqrt{|\alpha_r|^2 + |\alpha_l|^2} \tag{87}$$

Similarly:

$$\Delta L_z = \hbar\sqrt{|\alpha_r|^2 + |\alpha_l|^2} = \hbar\sqrt{|\alpha_x|^2 + |\alpha_y|^2} \tag{88}$$

Complement E_{VI}

A charged particle in a magnetic field: Landau levels

1 Review of the classical problem 771
1-a Motion of the particle 771
1-b The vector potential. The classical Lagrangian and Hamiltonian 773
1-c Constants of the motion in a uniform field 774
2 General quantum mechanical properties of a particle in a magnetic field 775
2-a Quantization. Hamiltonian 775
2-b Commutation relations 776
2-c Physical consequences 777
3 Case of a uniform magnetic field 779
3-a Eigenvalues of the Hamiltonian 779
3-b The observables in a particular gauge 782
3-c The stationary states 785
3-d Time evolution 791

Thus far, we have been considering, for various special cases, the properties of a particle subjected to a scalar potential $V(\mathbf{r})$ (representing, for example, the effect of an electric field on a charged particle). Chapter V (the harmonic oscillator) and Chapter VII (particle subjected to a central potential) treat other examples of scalar potentials. Here we shall be concerned with a complementary problem, that of the properties of a particle subjected to a vector potential $\mathbf{A}(\mathbf{r})$ (a charged particle placed in a magnetic field). We shall encounter a number of purely quantum mechanical effects, such as equally spaced energy levels in a uniform magnetic field (Landau levels)[1]. Before studying the problem from a quantum mechanical point of view, we shall rapidly review some classical results.

1. Review of the classical problem

1-a. Motion of the particle

When a particle of position $\mathbf{r}$ and charge q is subjected to a magnetic field $\mathbf{B}(\mathbf{r})$, the force $\mathbf{f}$ exerted on it is the Lorentz force:

$$\mathbf{f} = q\,\mathbf{v} \times \mathbf{B}(\mathbf{r}) \tag{1}$$

where:

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \tag{2}$$

[1] This equal spacing is, as we shall show, a consequence of the properties of the harmonic oscillator, and it could have been treated in Chapter V. However, we shall also see that the properties of angular momentum are useful in the study and classification of the stationary states of the particle. This is why this complement follows Chapter VI.

is the velocity of the particle. Its motion obeys the fundamental law of dynamics:

$$\mu \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \mathbf{f} \tag{3}$$

(where μ is the mass of the particle).

In the rest of this complement, we shall often be considering the case in which the magnetic field is uniform; we shall choose an Oz axis parallel to this field. By solving the equation of motion (3), one can show that in this case the three coordinates $x(t)$, $y(t)$ and $z(t)$ of the particle are given by:

$$\begin{aligned} x(t) &= x_0 + \sigma \cos(\omega_c t - \varphi_0) \\ y(t) &= y_0 + \sigma \sin(\omega_c t - \varphi_0) \\ z(t) &= v_{0z} t + z_0 \end{aligned} \tag{4}$$

where x_0, y_0, z_0, σ, φ_0 and v_{0z} are six constant parameters which depend on the initial conditions; the *cyclotron frequency* ω_c is given by:

$$\omega_c = -q \frac{B}{\mu} \tag{5}$$

Equations (4) show that the projection of the position M of the particle onto the xOy plane performs a uniform circular motion, of angular velocity ω_c and initial phase φ_0, on a circle of radius σ whose center is the point C_0, with coordinates x_0 and y_0. The motion of the projection of M onto Oz is simply rectilinear and uniform. It follows that the particle moves in space along a circular helix (*cf.* Fig. 1), whose axis is parallel to Oz and passes through C_0.

If we are concerned only with the motion of the point Q, the projection of M onto the xOy plane, we study the behavior of the vector:

$$\boldsymbol{\rho} = x\,\mathbf{e}_x + y\,\mathbf{e}_y \tag{6}$$

(where $\mathbf{e}_x$ and $\mathbf{e}_y$ are the unit vectors of the Ox and Oy axes). The velocity of Q is:

$$\mathbf{v}_\perp = \frac{\mathrm{d}\boldsymbol{\rho}}{\mathrm{d}t} \tag{7}$$

It is therefore convenient to introduce the components x' and y' of the vector $\mathbf{C_0Q}$:

$$\begin{aligned} x' &= x - x_0 \\ y' &= y - y_0 \end{aligned} \tag{8}$$

Since Q performs a uniform circular motion about C_0, we have:

$$\mathbf{v}_\perp = \omega_c\,\mathbf{e}_z \times \mathbf{C_0Q} \tag{9}$$

(where $\mathbf{e}_z$ is the unit vector of Oz). This implies that the coordinates x_0 and y_0 of C_0 are related to the coordinates of Q and to the components of $\mathbf{v}_\perp$ by:

$$\begin{aligned} x_0 &= x - \frac{1}{\omega_c} v_y \\ y_0 &= y + \frac{1}{\omega_c} v_x \end{aligned} \tag{10}$$

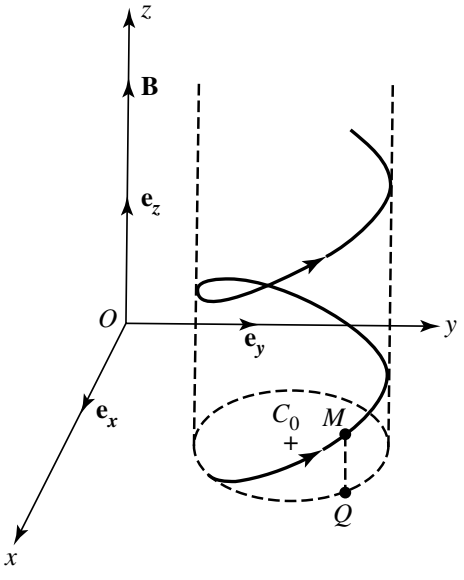


Figure 1: Classical trajectory of a charged particle in a uniform magnetic field parallel to Oz. The particle moves at constant velocity along a circular helix whose axis, parallel to Oz, passes through the point C_0. The figure is drawn for $q < 0$ (the case of the electron), that is, $\omega_c > 0$.

1-b. The vector potential. The classical Lagrangian and Hamiltonian

To describe the magnetic field $\mathbf{B}(\mathbf{r})$, one can use a vector potential $\mathbf{A}(\mathbf{r})$ which is, by definition, related to $\mathbf{B}(\mathbf{r})$ by:

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}) \tag{11}$$

For example, if the field $\mathbf{B}$ is uniform, one can choose:

$$\mathbf{A}(\mathbf{r}) = -\frac{1}{2}\mathbf{r} \times \mathbf{B} \tag{12}$$

We know, furthermore, that when $\mathbf{B}(\mathbf{r})$ is given, condition (11) does not determine $\mathbf{A}(\mathbf{r})$ uniquely: a gradient of an arbitrary function of $\mathbf{r}$ can be added to $\mathbf{A}(\mathbf{r})$ without changing[2] $\mathbf{B}(\mathbf{r})$.

It can be shown (*cf.* Appendix III, § 4-b) that the Lagrange function $\mathcal{L}(\mathbf{r}, \mathbf{v})$ of the particle is given by:

$$\mathcal{L}(\mathbf{r}, \mathbf{v}) = \frac{1}{2}\mu\,\mathbf{v}^2 + q\,\mathbf{v}.\mathbf{A}(\mathbf{r}) \tag{13}$$

[2]For example, for a uniform field parallel to Oz, one could choose, instead of the vector $\mathbf{A}(\mathbf{r})$ given by (12), the vector whose components are $A_x = 0$, $A_y = xB$, $A_z = 0$.

It follows that $\mathbf{p}$, the conjugate momentum of the position $\mathbf{r}$, is related to $\mathbf{v}$ and $\mathbf{A}(\mathbf{r})$ by:

$$\mathbf{p} = \nabla_{\mathbf{v}} \mathcal{L}(\mathbf{r}, \mathbf{v}) = \mu\,\mathbf{v} + q\mathbf{A}(\mathbf{r}) \tag{14}$$

The classical Hamiltonian $\mathcal{H}(\mathbf{r}, \mathbf{p})$ is then:

$$\mathcal{H}(\mathbf{r}, \mathbf{p}) = \frac{1}{2\mu}[\mathbf{p} - q\mathbf{A}(\mathbf{r})]^2 \tag{15}$$

It will prove convenient to set:

$$\mathcal{H}(\mathbf{r}, \mathbf{p}) = \mathcal{H}_\perp(\mathbf{r}, \mathbf{p}) + \mathcal{H}_{/\!/}(\mathbf{r}, \mathbf{p}) \tag{16}$$

with:

$$\begin{aligned} \mathcal{H}_\perp(\mathbf{r}, \mathbf{p}) &= \frac{1}{2\mu}\left\{[p_x - qA_x(\mathbf{r})]^2 + [p_y - qA_y(\mathbf{r})]^2\right\} \\ \mathcal{H}_{/\!/}(\mathbf{r}, \mathbf{p}) &= \frac{1}{2\mu}\,[p_z - qA_z(\mathbf{r})]^2 \end{aligned} \tag{17}$$

Comment:

In this case, unlike that of a scalar potential $V(\mathbf{r})$, relation (14) shows that the momentum $\mathbf{p}$ is not equal to the mechanical momentum $\mu\,\mathbf{v}$. Also, comparing (14) with (15), we see that $\mathcal{H}$ is equal to the kinetic energy $\mu\,\mathbf{v}^2/2$ of the particle; this results from the fact that since the Lorentz force written in (1) is always perpendicular to $\mathbf{v}$, it does no work during the motion. Similarly, it must be noted that the angular momentum:

$$\boldsymbol{\mathcal{L}} = \mathbf{r} \times \mathbf{p} \tag{18}$$

is different from the moment of the mechanical momentum $\mu\,\mathbf{v}$:

$$\boldsymbol{\lambda} = \mathbf{r} \times \mu\mathbf{v} \tag{19}$$

1-c. Constants of the motion in a uniform field

Consider the special case in which the field $\mathbf{B}$ is uniform. The motion of the particle (§ 1-a) is such that $\mathcal{H}_{/\!/}$ and $\mathcal{H}_\perp$, defined in (17), are constants of the motion[3].

If we substitute (14) into (10), we obtain:

$$\begin{aligned} x_0 &= x - \frac{1}{\mu\omega_c}[p_y - qA_y(\mathbf{r})] \\ y_0 &= y + \frac{1}{\mu\omega_c}[p_x - qA_x(\mathbf{r})] \end{aligned} \tag{20}$$

It follows that the radius σ of the helical trajectory satisfies:

$$\begin{aligned} \sigma^2 &= (x - x_0)^2 + (y - y_0)^2 = \left(\frac{1}{\mu\omega_c}\right)^2 \left\{[p_y - qA_y(\mathbf{r})]^2 + [p_x - qA_x(\mathbf{r})]^2\right\} \\ &= \frac{2}{\mu\omega_c^2}\mathcal{H}_\perp \end{aligned} \tag{21}$$

[3]This follows from the fact that, according to (14) and (17), $\mathcal{H}_\perp$ and $\mathcal{H}_{/\!/}$ are equal, respectively, to the kinetic energies $\mu\mathbf{v}_\perp^2/2$ and $\mu\mathbf{v}_z^2/2$ associated with the motions perpendicular and parallel to Oz.

σ^2 is therefore proportional to the Hamiltonian $\mathcal{H}_\perp$.

Similarly, let $\boldsymbol{\theta}$ be the moment of the mechanical momentum $\mu\,\mathbf{v}$ with respect to the center C_0 of the circle:

$$\boldsymbol{\theta} = \mathbf{C_0M} \times \mu\,\mathbf{v} \tag{22}$$

The component θ_z of this moment can then be written, with (20) taken into account:

$$\begin{aligned}\theta_z &= \mu\,[(x-x_0)v_y - (y-y_0)v_x] \\ &= \frac{1}{\mu\omega_c}\left\{[p_y - qA_y(\mathbf{r})]^2 + [p_x - qA_x(\mathbf{r})]^2\right\} = \frac{2}{\omega_c}\mathcal{H}_\perp\end{aligned} \tag{23}$$

θ_z is therefore a constant of the motion, as might have been expected. On the other hand, the component λ_z of the moment of the mechanical momentum $\mu\,\mathbf{v}$ with respect to O is not, in general, constant, since:

$$\lambda_z = \theta_z + \mu\,[x_0 v_y(t) - y_0 v_x(t)] \tag{24}$$

Therefore, according to (4), λ_z varies sinusoidally in time.

Finally, consider the projection $\mathcal{L}_z$ onto Oz of the angular momentum $\boldsymbol{\mathcal{L}}$:

$$\mathcal{L}_z = xp_y - yp_x \tag{25}$$

According to (14), it can be written:

$$\mathcal{L}_z = x\,[\mu\,v_y + qA_y(\mathbf{r})] - y\,[\mu\,v_x + qA_x(\mathbf{r})] \tag{26}$$

It therefore depends explicitly on the gauge chosen, that is, on the vector potential $\mathbf{A}$ picked to describe the magnetic field. In most cases, $\mathcal{L}_z$ is not a constant of the motion. Nevertheless, if one chooses the gauge given in (12), one obtains from (4):

$$\mathcal{L}_z = \frac{qB}{2}\left(x_0^2 + y_0^2 - \sigma^2\right) \tag{27}$$

$\mathcal{L}_z$ is then a constant of the motion.

Relation (27) does not have a simple physical interpretation, since it is valid only in a particular gauge. However, it will prove useful to us in the following sections for the quantum mechanical study of the problem.

2. General quantum mechanical properties of a particle in a magnetic field

2-a. Quantization. Hamiltonian

Consider a particle placed in an arbitrary magnetic field described by the vector potential $\mathbf{A}(x, y, z)$. In quantum mechanics, the vector potential becomes an operator, a function of three observables, X, Y and Z. The operator H, the Hamiltonian of the particle, can be obtained from (15):

$$H = \frac{1}{2\mu}\left[\mathbf{P} - q\mathbf{A}(X, Y, Z)\right]^2 \tag{28}$$

According to (14), the operator $\mathbf{V}$ associated with the velocity of the particle is given by:

$$\mathbf{V} = \frac{1}{\mu}\left[\mathbf{P} - q\mathbf{A}(X, Y, Z)\right] \tag{29}$$

which enables us to write H in the form:

$$H = \frac{\mu}{2}\mathbf{V}^2 \tag{30}$$

2-b. Commutation relations

The observables $\mathbf{R}$ and $\mathbf{P}$ satisfy the canonical commutation relations:

$$[X,\, P_x] = [Y,\, P_y] = [Z,\, P_z] = i\hbar \tag{31}$$

The other commutators between components of $\mathbf{R}$ and $\mathbf{P}$ are zero. Two components of $\mathbf{P}$ therefore commute. However, we see from (29) that the same is not true for $\mathbf{V}$; for example:

$$[V_x, V_y] = -\frac{q}{\mu^2}\left\{[P_x,\, A_y(\mathbf{R})] + [A_x(\mathbf{R}),\, P_y]\right\} \tag{32}$$

This expression is easy to calculate, using the rule given in Complement B_{II} [*cf.* formula (48)]:

$$[V_x,\, V_y] = \frac{iq\hbar}{\mu^2}\left\{\frac{\partial A_y}{\partial X} - \frac{\partial A_x}{\partial Y}\right\} = \frac{iq\hbar}{\mu^2}B_z(\mathbf{R}) \tag{33a}$$

Similarly, it can be shown that:

$$[V_y,\, V_z] = \frac{iq\hbar}{\mu^2}B_x(\mathbf{R}) \tag{33b}$$

$$[V_z,\, V_x] = \frac{iq\hbar}{\mu^2}B_y(\mathbf{R}) \tag{33c}$$

The magnetic field therefore enters explicitly into the commutation relations for the velocity.

However, since $\mathbf{A}(\mathbf{R})$ commutes with X, Y and Z, relation (29) implies that:

$$[X,\, V_x] = \frac{1}{\mu}[X,\, P_x] = \frac{i\hbar}{\mu} \tag{34a}$$

and, similarly:

$$[Y,\, V_y] = [Z,\, V_z] = \frac{i\hbar}{\mu} \tag{34b}$$

(the other commutators between a component of $\mathbf{R}$ and a component of $\mathbf{V}$ are zero). From these relations, it can be deduced (*cf.* Complement C_{III}) that:

$$\Delta X.\Delta V_x \geq \frac{\hbar}{2\mu} \tag{35}$$

(with analogous inequalities for the components along Oy and Oz). The physical consequences of the Heisenberg uncertainty relations are therefore not modified by the presence of a magnetic field.

Finally, let us calculate the commutation relations between the components of the operator:

$$\boldsymbol{\Lambda} = \mu\, \mathbf{R} \times \mathbf{V} \tag{36}$$

associated with the moment with respect to O of the mechanical momentum[4]. We obtain:

$$\begin{aligned}[\Lambda_x,\, \Lambda_y] &= \mu^2\, [YV_z - ZV_y,\, ZV_x - XV_z] \\ &= \mu^2 Y\, \{[V_z,\, Z]\, V_x + Z\, [V_z,\, V_x]\} - \mu^2 Z^2\, [V_y,\, V_x] \\ &\qquad + \mu^2 X\, \{Z\, [V_y,\, V_z] + [Z,\, V_z]\, V_y\}\end{aligned} \tag{37}$$

that is, with (33) and (34) taken into account:

$$[\Lambda_x,\, \Lambda_y] = i\hbar \left\{-\mu Y V_x + qYZB_y + qZ^2 B_z + qXZB_x + \mu X V_y\right\} \tag{38}$$

It follows that:

$$[\Lambda_x,\, \Lambda_y] = i\hbar\, \{\Lambda_z + qZ\, \mathbf{R} \cdot \mathbf{B}(\mathbf{R})\} \tag{39}$$

(the other commutators can be obtained by cyclic permutation of the indices x, y and z). When the field $\mathbf{B}$ is not zero, the commutation relations of $\boldsymbol{\Lambda}$ are completely different from those of $\mathbf{L}$. The operator $\boldsymbol{\Lambda}$ therefore does not, *a priori*, possess the properties of angular momenta proved in Chapter VI.

2-c. Physical consequences

α. Evolution of $\langle \mathbf{R} \rangle$

The time variation of the mean position of the particle is given by Ehrenfest's theorem:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbf{R} \rangle = \langle [\mathbf{R},\, H] \rangle = \left\langle \left[\mathbf{R},\, \frac{\mu}{2} \mathbf{V}^2\right] \right\rangle \tag{40}$$

[according to formula (30)]. Equations (34) are not difficult to interpret, since, substituted into (40), they yield:

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbf{R} \rangle = \langle \mathbf{V} \rangle \tag{41}$$

As in the case in which the magnetic field is zero, the mean velocity is therefore equal to the derivative of $\langle \mathbf{R} \rangle$. Equation (41) is the quantum mechanical analogue of (2).

β. Evolution of $\langle \mathbf{V} \rangle$*. The Lorentz force*

Let us calculate the time derivative of the mean value $\langle \mathbf{V} \rangle$ of the velocity:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbf{V} \rangle = \left\langle \left[\mathbf{V},\, \frac{\mu}{2} \mathbf{V}^2\right] \right\rangle \tag{42}$$

[4] Of course, the components of the angular momentum $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ always satisfy the usual commutation relations.

Since, according to relations (33):

$$\begin{aligned}\left[\mathbf{V}^2,\, V_x\right] &= \left[V_x^2 + V_y^2 + V_z^2,\, V_x\right] \\ &= V_y\left[V_y,\, V_x\right] + \left[V_y,\, V_x\right]V_y + V_z\left[V_z,\, V_x\right] + \left[V_z,\, V_x\right]V_z \\ &= \frac{iq\hbar}{\mu^2}\left\{-V_yB_z(\mathbf{R}) - B_z(\mathbf{R})V_y + V_zB_y(\mathbf{R}) + B_y(\mathbf{R})V_z\right\} \end{aligned} \tag{43}$$

it is easy to see that:

$$\mu\frac{\mathrm{d}}{\mathrm{d}t}\langle\mathbf{V}\rangle = \langle\mathbf{F}(\mathbf{R},\, \mathbf{V})\rangle \tag{44}$$

where the operator $\mathbf{F}(\mathbf{R},\, \mathbf{V})$ is defined by:

$$\mathbf{F}(\mathbf{R},\, \mathbf{V}) = \frac{q}{2}\left\{\mathbf{V}\times\mathbf{B}(\mathbf{R}) - \mathbf{B}(\mathbf{R})\times\mathbf{V}\right\} \tag{45}$$

The last two relations are simply the analogues of the classical relations (1) and (3). Here, we obtain a symmetrized expression for $\mathbf{F}(\mathbf{R},\, \mathbf{V})$ (*cf.* Chap. III, § B-5), since $\mathbf{R}$ and $\mathbf{V}$ do not commute; a minus sign appears since the vector product is antisymmetric.

γ. *Evolution of* $\langle\mathbf{\Lambda}\rangle$

Now let us evaluate:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle\mathbf{\Lambda}\rangle = \langle[\mathbf{\Lambda},\, H]\rangle \tag{46}$$

To do so, let us calculate, for example, the commutator $[XV_y - YV_x,\, H]$:

$$\begin{aligned}[XV_y - YV_x,\, H] &= X[V_y,\, H] + [X,\, H]V_y - Y[V_x,\, H] - [Y,\, H]V_x \\ &= \frac{i\hbar}{\mu}\left(XF_y - YF_x\right) + i\hbar\left(V_xV_y - V_yV_x\right) \end{aligned} \tag{47}$$

But X and V_y commute, as do Y and V_x. The commutator we are calculating is therefore equal to:

$$\begin{aligned}[V_yX - V_xY,\, H] &= V_y[X,\, H] + [V_y,\, H]X - V_x[Y,\, H] - [V_x,\, H]Y \\ &= \frac{i\hbar}{\mu}\left(F_yX - F_xY\right) + i\hbar\left(V_yV_x - V_xV_y\right) \end{aligned} \tag{48}$$

Taking half the sum of these two expressions, we find $\dfrac{\mathrm{d}}{\mathrm{d}t}\langle\Lambda_z\rangle$ in the form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\Lambda_z\rangle = \frac{1}{2}\left\langle XF_y - YF_x - F_xY + F_yX\right\rangle \tag{49}$$

Analogous arguments give the derivative of $\langle\Lambda_x\rangle$ and $\langle\Lambda_y\rangle$; finally:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\mathbf{\Lambda}\rangle = \frac{1}{2}\left\langle\mathbf{R}\times\mathbf{F}(\mathbf{R},\, \mathbf{V}) - \mathbf{F}(\mathbf{R},\, \mathbf{V})\times\mathbf{R}\right\rangle \tag{50}$$

The classical analogue of this relation is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\lambda} = \mathbf{r}\times\mathbf{f}(\mathbf{r},\, \mathbf{v}) \tag{51}$$

which expresses a well-known theorem: the time derivative of the moment of the mechanical momentum with respect to a fixed point O is equal to the moment with respect to O of the force exerted on the particle.

3. Case of a uniform magnetic field

When the magnetic field is uniform, the preceding general study can easily be pursued further. We choose the direction of the field $\mathbf{B}$ as the Oz axis. The commutation relations (33) then become, using definition (5):

$$[V_x, V_y] = -i\frac{\hbar\omega_c}{\mu} \tag{52a}$$

$$[V_y, V_z] = [V_z, V_x] = 0 \tag{52b}$$

Comment:

Applying the results of Complement C_{III} to V_x and V_y, we can see from (52a) that their root mean square deviations satisfy:

$$\Delta V_x . \Delta V_y \geq \frac{\hbar|\omega_c|}{2\mu} \tag{53}$$

The components of the velocity $\mathbf{V}_\perp$ are therefore incompatible physical quantities.

3-a. Eigenvalues of the Hamiltonian

By analogy with (16), H can be written in the form:

$$H = H_\perp + H_{/\!/} \tag{54}$$

with:

$$H_\perp = \frac{\mu}{2}\left(V_x^2 + V_y^2\right) \tag{55a}$$

$$H_{/\!/} = \frac{\mu}{2}V_z^2 \tag{55b}$$

According to (52b):

$$[H_\perp, H_{/\!/}] = 0 \tag{56}$$

We can now look for a basis of eigenvectors common to $H_\perp$ (eigenvalues $E_\perp$) and $H_{/\!/}$ (eigenvalues $E_{/\!/}$); they will automatically be eigenvectors of H, with the eigenvalues:

$$E = E_\perp + E_{/\!/} \tag{57}$$

α. Eigenvalues of $H_{/\!/}$

The eigenvectors of the operator V_z are also eigenvectors of $H_{/\!/}$. Now, Z and V_z are two Hermitian operators which satisfy the relation:

$$[Z, V_z] = \frac{i\hbar}{\mu} \tag{58}$$

We can therefore apply to them the results of Complement E_{II}; in particular, the spectrum of V_z includes all the real numbers.

Consequently, the eigenvalues of $H_{/\!/}$ are of the form:

$$E_{/\!/} = \frac{\mu}{2} v_z^2 \tag{59}$$

where v_z is a real arbitrary constant. The spectrum of $H_{/\!/}$ is therefore continuous: the energy $E_{/\!/}$ can take on any positive value or zero.

The interpretation of this result is obvious: $H_{/\!/}$ describes the kinetic energy of a free particle moving along Oz (as in classical mechanics; § 1-a).

β. Eigenvalues of $H_\perp$

We shall assume, as an example, that the particle under consideration has a negative charge q; the cyclotron frequency ω_c is then positive [formula (5)][5].

We set:

$$\hat{Q} = \sqrt{\frac{\mu}{\hbar\omega_c}} V_y$$
$$\hat{S} = \sqrt{\frac{\mu}{\hbar\omega_c}} V_x \tag{60}$$

Relation (52a) can then be written:

$$[\hat{Q}, \hat{S}] = i \tag{61}$$

and $H_\perp$ becomes:

$$H_\perp = \frac{\hbar\omega_c}{2} \left(\hat{Q}^2 + \hat{S}^2 \right) \tag{62}$$

$H_\perp$ then takes on the form of the Hamiltonian of a one-dimensional harmonic oscillator [*cf.* Chap. V, relation (B-4)]. $\hat{Q}$ and $\hat{S}$, which satisfy (61), play the roles of the position $\hat{X}$ and momentum $\hat{P}$ of this oscillator.

The arguments set forth in § B-2 of Chapter V for the operators $\hat{X}$ and $\hat{P}$ can be repeated here for $\hat{Q}$ and $\hat{S}$. For example, it can easily be shown that if $|\varphi_\perp\rangle$ is an eigenvector of $H_\perp$:

$$H_\perp |\varphi_\perp\rangle = E_\perp |\varphi_\perp\rangle \tag{63}$$

the kets:

$$|\varphi'_\perp\rangle = \frac{1}{\sqrt{2}} \left(\hat{Q} + i\hat{S} \right) |\varphi_\perp\rangle \tag{64a}$$
$$|\varphi''_\perp\rangle = \frac{1}{\sqrt{2}} \left(\hat{Q} - i\hat{S} \right) |\varphi_\perp\rangle \tag{64b}$$

[5] For a positive charge q, one can keep the convention of positive ω_c by choosing the direction of the Oz axis opposite to the magnetic field.

are also eigenvectors of $H_\perp$:

$$H_\perp|\varphi'_\perp\rangle = (E_\perp - \hbar\omega_c)\,|\varphi'_\perp\rangle \tag{65a}$$
$$H_\perp|\varphi''_\perp\rangle = (E_\perp + \hbar\omega_c)\,|\varphi''_\perp\rangle \tag{65b}$$

From this we deduce that the possible values of $E_\perp$ are given by:

$$E_\perp = \left(n + \frac{1}{2}\right)\hbar\omega_c \tag{66}$$

where n is a positive integer or zero.

γ. *Eigenvalues of H*

According to the preceding results, the eigenvalues of the total Hamiltonian H are of the form:

$$E(n, v_z) = \left(n + \frac{1}{2}\right)\hbar\omega_c + \frac{1}{2}\mu v_z^2 \tag{67}$$

The corresponding levels are called *Landau levels*.

For a given value of v_z, all possible values of n (positive integers or zero) are actually found. From (64) and 65 we see that the repeated action of the operators $\frac{1}{\sqrt{2}}\left(\hat{Q} \pm i\hat{S}\right)$ on an eigenvector of H of eigenvalue $E(n,\, v_z)$ provides an energy state $E(n',\, v_z)$, where n' is any integer but where v_z has not changed (since $\hat{Q}$ and $\hat{S}$ commute with $H_{/\!/}$). Therefore, although the energy of the motion along Oz is not quantized, that of the motion projected onto xOy is.

Comment:

We showed in Chapter V (§ B-3) that the energy levels of the one-dimensional harmonic oscillator are non-degenerate in $\mathcal{E}_x$. The situation is different here, since the particle under study is moving in three-dimensional space. Since the destruction operator of a quantum $\hbar\omega_c$ is $\frac{1}{\sqrt{2}}\left(\hat{Q} + i\hat{S}\right) = \sqrt{\frac{\mu}{2\hbar\omega_c}}(V_y + iV_x)$, the eigenvectors of $H_\perp$ corresponding to $n = 0$ are solutions of the equation:

$$(V_y + iV_x)|\varphi\rangle = 0 \tag{68}$$

On the one hand, vectors that are solutions of (68) can be eigenvectors of $H_{/\!/}$ with an arbitrary (positive) eigenvalue. On the other, even for a fixed value of v_z, equation (68) is a partial differential equation with respect to x and y, and has an infinite number of solutions. The energies $E(n = 0,\, v_z)$ are therefore infinitely degenerate. By using the creation operator for a quantum, it can easily be shown that this is true for all the levels $E(n, v_z)$, for any n (non-negative integer).

3-b. The observables in a particular gauge

In order to state the above results more precisely, we shall calculate the stationary states of the system. This will enable us to study their physical properties. It is now necessary to choose a gauge; we shall choose the one given by (12). The components of the velocity are then:

$$V_x = \frac{P_x}{\mu} - \frac{\omega_c}{2} Y$$
$$V_y = \frac{P_y}{\mu} + \frac{\omega_c}{2} X$$
$$V_z = \frac{P_z}{\mu} \tag{69}$$

α. The Hamiltonians $H_\perp$ and $H_{/\!/}$. Relation with the two-dimensional harmonic oscillator

Substituting (69) into (55), we obtain:

$$H_\perp = \frac{P_x^2 + P_y^2}{2\mu} + \frac{\omega_c}{2} L_z + \frac{\mu\omega_c^2}{8}\left(X^2 + Y^2\right) \tag{70a}$$

$$H_{/\!/} = \frac{P_z^2}{2\mu} \tag{70b}$$

where L_z is the component along Oz of the angular momentum $\mathbf{L} = \mathbf{R} \times \mathbf{P}$.

In the $\{|\mathbf{r}\rangle\}$ representation, $H_{/\!/}$ is an operator that acts only on the variable z, while $H_\perp$ acts only on the variables x and y. We can therefore find a basis of eigenvectors of H by solving in $\mathcal{E}_z$ the eigenvalue equation of $H_{/\!/}$, and then, in $\mathcal{E}_{xy}$, that of $H_\perp$. All we must then do is take the tensor products of the vectors obtained.

Actually, the eigenvalue equation of $H_{/\!/}$ simply leads to the wave functions:

$$\varphi(z) = \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{ip_z z/\hbar} \tag{71}$$

with:

$$E_{/\!/} = \frac{p_z^2}{2\mu} \tag{72}$$

[we again find (59)]. Therefore, we shall concentrate on solving the eigenvalue equation of $H_\perp$ in $\mathcal{E}_{xy}$; the wave functions we shall be considering now depend on x and y, and not on z.

Comparing (70a) with expression (12a) of Complement D_{VI}, we see that $H_\perp$ can be expressed simply in terms of the Hamiltonian H_{xy} of a two-dimensional harmonic oscillator:

$$H_\perp = H_{xy} + \frac{\omega_c}{2} L_z \tag{73}$$

if we choose for the value of the constant that enters into H_{xy}:

$$\omega = \frac{\omega_c}{2} \tag{74}$$

Now, in Complement D_{VI}, we saw that H_{xy} and L_z form a C.S.C.O. in $\mathcal{E}_{xy}$, and we constructed a basis of eigenvectors $|\chi_{n_r,\,n_l}\rangle$ common to these two observables [*cf.* formula (47) of D_V]. The $|\chi_{n_r,\,n_l}\rangle$ are also eigenvectors of $H_\perp$; Complement D_{VI} therefore gives the solutions to the eigenvalue equation of $H_\perp$.

Comments:

(*i*) In § 3-a, we saw that $H_\perp$ can be written in a form that is analogous to that of a Hamiltonian of a one-dimensional harmonic oscillator. Here, we find that, in a particular gauge, this same operator $H_\perp$ is also simply related to the Hamiltonian H_{xy} of a two-dimensional harmonic oscillator. These two results are not contradictory; they simply correspond to two different decompositions of the same Hamiltonian, which must obviously lead to the same physical conclusions.

(*ii*) One must not lose sight of the fact that the Hamiltonian $H_\perp$ involves a physical problem which is completely different from that of the two-dimensional harmonic oscillator: the charged particle is subjected to a vector potential (describing a uniform magnetic field) and not a harmonic scalar potential (which would describe, for example, a non-uniform electric field). It so happens that, in the gauge chosen, the effects of the magnetic field can be likened to those of a fictitious harmonic scalar potential.

β. Expression for the observables in terms of the creation and destruction operators of circular quanta

First of all, we shall express the observables describing the quantities associated with the particle in terms of the operators a_r and a_l [defined by equations (40) of Complement D_{VI}] and their adjoints $a_r^\dagger$ and $a_l^\dagger$ (we shall also use the operators $N_r = a_r^\dagger a_r$ and $N_l = a_l^\dagger a_l$).

Substituting relations (46) of D_{VI} into (73), we obtain[6]:

$$H_\perp = \left(N_r + \frac{1}{2}\right)\hbar\omega_c \tag{75}$$

The energy associated with the state $|\chi_{n_r,\,n_l}\rangle$ is therefore:

$$E_\perp = \left(N_r + \frac{1}{2}\right)\hbar\omega_c \tag{76}$$

as we found in (66). Moreover, since $E_\perp$ is independent of n_l, we see that all the eigenvalues of $H_\perp$ are infinitely degenerate.

Using relations (23) and (40) of D_{VI}, we can see that:

$$\begin{aligned} X &= \frac{1}{2\beta}\left(a_r + a_r^\dagger + a_l + a_l^\dagger\right) \\ Y &= \frac{i}{2\beta}\left(a_r - a_r^\dagger - a_l + a_l^\dagger\right) \end{aligned} \tag{77}$$

[6] Recall that we have assumed ω_c to be positive. If ω_c were negative, the indices r and l would have to be inverted in a certain number of the following formulas; for example, (75) would become: $H_\perp = (N_l + 1/2)\hbar\,|\omega_c|$.

where, using (74), β is defined by:

$$\beta = \sqrt{\frac{\mu\omega_c}{2\hbar}} \tag{78}$$

Similarly:

$$\begin{aligned} P_x &= \frac{i\hbar\beta}{2}\left(-a_r + a_r^\dagger - a_l + a_l^\dagger\right) \\ P_y &= \frac{\hbar\beta}{2}\left(a_r + a_r^\dagger - a_l - a_l^\dagger\right) \end{aligned} \tag{79}$$

These expressions, substituted into (69), yield:

$$\begin{aligned} V_x &= -\frac{i\omega_c}{2\beta}\left(a_r - a_r^\dagger\right) \\ V_y &= \frac{\omega_c}{2\beta}\left(a_r + a_r^\dagger\right) \end{aligned} \tag{80}$$

Since a_r and $a_r^\dagger$ do not commute with N_r, it can be seen by using (75) that, as in classical mechanics, V_x and V_y are not constants of the motion; in addition, using the commutation relations of a_r and $a_r^\dagger$, we indeed obtain (52a).

It is also interesting to study the quantum mechanical operators associated with the various variables introduced in the description of the classical motion (§ 1): the coordinates (x_0, y_0) of the center C_0 of the classical trajectory, the components (x', y') of the vector $\mathbf{C_0Q}$, etc. As above, we shall denote each of these operators by the capital letter corresponding to the small letter which designates the corresponding classical variable. By analogy with (10), we therefore set:

$$X_0 = X - \frac{1}{\omega_c}V_y = \frac{1}{2\beta}\left(a_l + a_l^\dagger\right) \tag{81a}$$

$$Y_0 = Y + \frac{1}{\omega_c}V_x = \frac{i}{2\beta}\left(a_l^\dagger - a_l\right) \tag{81b}$$

The operators a_l and $a_l^\dagger$ commute with N_r; it follows that X_0 and Y_0 are constants of the motion. Formulas (81) also imply that:

$$[X_0, Y_0] = \frac{i}{2\beta^2} = \frac{i\hbar}{\mu\omega_c} \tag{82}$$

Consequently, X_0 and Y_0 are incompatible physical quantities, their root mean square deviations being related by:

$$\Delta X_0 . \Delta Y_0 \geq \frac{\hbar}{2\mu\omega_c} \tag{83}$$

We also define:

$$\begin{aligned} X' &= X - X_0 = \frac{1}{2\beta}\left(a_r + a_r^\dagger\right) \\ Y' &= Y - Y_0 = \frac{i}{2\beta}\left(a_r - a_r^\dagger\right) \end{aligned} \tag{84}$$

We immediately see that X' and Y', as in classical mechanics, are not constants of the motion; moreover, X' and Y' are simply proportional to V_y and V_x respectively:

$$\begin{aligned} V_x &= -\omega_c Y' \\ V_y &= \omega_c X' \end{aligned} \tag{85}$$

like the corresponding classical variables [formula (9)]. According to (53), equations (85) imply:

$$\Delta X'.\Delta Y' \geq \frac{\hbar}{2\mu\omega_c} \tag{86}$$

Let Σ^2 be the operator corresponding to σ^2 (square of the radius of the classical trajectory):

$$\Sigma^2 = (X - X_0)^2 + (Y - Y_0)^2 \tag{87}$$

According to (81), we have:

$$\Sigma^2 = \left(\frac{1}{\omega_c}\right)^2 \left(V_x^2 + V_y^2\right) = \frac{2}{\mu\omega_c^2} H_\perp \tag{88}$$

Σ^2 is therefore a constant of the motion, as is σ^2 in classical mechanics.

Finally, the operator associated with the moment of the mechanical momentum $\mu\,\mathbf{v}$ with respect to O is:

$$\Theta_z = \mu\left[(X - X_0)V_y - (Y - Y_0)V_x\right] \tag{89}$$

and formulas (81) indicate that:

$$\Theta_z = \frac{2}{\omega_c} H_\perp \tag{90}$$

as in (23). Θ_z is therefore a constant of the motion. On the other hand, the operator Λ_z, the component along Oz of $\mu\,\mathbf{R} \times \mathbf{V}$, is:

$$\Lambda_z = \frac{2}{\omega_c} H_\perp + \hbar\left(a_r a_l + a_r^\dagger a_l^\dagger\right) \tag{91}$$

and therefore does not commute with $H_\perp$.

3-c. The stationary states

We indicated above that the eigenvalues of the Hamiltonian $H_\perp$ are all infinitely degenerate in $\mathcal{E}_{xy}$. For each positive or zero integer n, there exists an infinite-dimensional subspace $\mathcal{E}_{xy}^{(n)}$ of $\mathcal{E}_{xy}$, all of whose kets are eigenvectors of $H_\perp$ with the same eigenvalues $(n + 1/2)\hbar\omega_c$. In this section, we shall study different bases which can be chosen in each of these subspaces. First, we shall indicate the general properties of the stationary states, valid for any basis of eigenstates of $H_\perp$.

α. General properties

Relations (88) and (90) show that an arbitrary stationary state is necessarily an eigenvector of Σ^2 and Θ_z; the corresponding physical quantities are therefore always well-defined in such a state and are equal to:

$$\begin{aligned} &(2n+1)\frac{\hbar}{\mu\omega_c} \qquad \text{for } \Sigma^2 \\ &(2n+1)\hbar \qquad \text{for } \Theta_z \end{aligned} \tag{92}$$

The values of Σ^2 and Θ_z are proportional to the energy; this corresponds to the classical description of the motion (*cf.* § 1).

It follows from (80) and (84) that X', Y', V_x and V_y have no matrix elements inside a given subspace $\mathcal{E}_{xy}^{(n)}$; it follows, for a stationary state, that:

$$\begin{aligned} &\langle V_x\rangle = \langle V_y\rangle = 0 \\ &\langle X'\rangle = \langle Y'\rangle = 0 \end{aligned} \tag{93}$$

Nevertheless, since V_x and V_y (and therefore X' and Y') are not constants of the motion, the corresponding physical quantities do not have perfectly well-defined values in a stationary state. In fact, by using (80), (84) and the properties of the one-dimensional harmonic oscillator [*cf.* Chap. V, relation (D-5)], it can be shown that:

$$\begin{aligned} \Delta V_x &= \omega_c \Delta Y' = \sqrt{\left(n+\frac{1}{2}\right)\frac{\hbar\omega_c}{\mu}} \\ \Delta V_y &= \omega_c \Delta X' = \sqrt{\left(n+\frac{1}{2}\right)\frac{\hbar\omega_c}{\mu}} \end{aligned} \tag{94}$$

in agreement with (53). Moreover, we see that the only stationary states in which the product $\Delta V_x.\Delta V_y$ (or $\Delta X'.\Delta Y'$) takes on its minimal value are the ground states ($n = 0$).

Comment:

The various ground states are solutions of the equation:

$$a_r|\varphi\rangle = 0 \tag{95a}$$

that is, using (80):

$$(V_y + iV_x)|\varphi\rangle = 0 \tag{95b}$$

as we found in (68).

β. The states $|\chi_{n_r,\,n_l}\rangle$

As we saw in Complement D_{VI}, the fact that $H_\perp$ and L_z form a C.S.C.O. in $\mathcal{E}_{xy}$ can be used to construct a basis of eigenvectors common to these two observables. This basis is composed of the vectors $|\chi_{n_r,\,n_l}\rangle$, since, according to (75) and formula (46) of Complement D_{VI}:

$$H_\perp|\chi_{n_r,\,n_l}\rangle = \left(n_r+\frac{1}{2}\right)\hbar\omega_c|\chi_{n_r,\,n_l}\rangle \tag{96a}$$

$$L_z|\chi_{n_r,\,n_l}\rangle = (n_r - n_l)\,\hbar|\chi_{n_r,\,n_l}\rangle \tag{96b}$$

The subspace $\mathcal{E}_{xy}^{(n)}$ defined by specifying the (non-negative) integer n is therefore spanned by the set of vectors $|\chi_{n_r,\,n_l}\rangle$ such that $n_r = n$. The eigenvalues of L_z associated with these different vectors are of the form $m\hbar$, and, for fixed n, m is an integer which can vary between $-\infty$ and n (for example, all the ground states correspond to negative values of m; this is related to the hypothesis $\omega_c > 0$ posed above).

The wave functions associated with the states $|\chi_{n_r,\,n_l}\rangle$ were calculated in Complement D_{VI} (§ 3-a).

Note that the states $|\chi_{n_r,\,n_l}\rangle$ are eigenstates of the operator L_z, but not of the operator Λ_z associated with the moment of the mechanical momentum. This can be seen directly from formula (91).

In a state $|\chi_{n_r,\,n_l}\rangle$, the mean values $\langle X_0\rangle$ and $\langle Y_0\rangle$ are zero, according to (81). However, neither X_0 nor Y_0 corresponds to perfectly well-defined physical quantities, since, by using the properties of the one-dimensional harmonic oscillator, it can easily be shown that, in a state $|\chi_{n_r,\,n_l}\rangle$:

$$\begin{aligned}\Delta X_0 &= \sqrt{\left(n_l + \frac{1}{2}\right)\frac{\hbar}{\mu\omega_c}}\\ \Delta Y_0 &= \sqrt{\left(n_l + \frac{1}{2}\right)\frac{\hbar}{\mu\omega_c}}\end{aligned} \tag{97}$$

The minimal value of the product $\Delta X_0.\Delta Y_0$ is therefore attained for the states $|\chi_{n_r,\,n_l=0}\rangle$, that is, the states of each energy level $E_\perp = (n + 1/2)\hbar\omega_c$ for which L_z takes on its maximal value $n\hbar$ [*cf.* (96)].

However, let us define the operator:

$$\Gamma^2 = X_0^2 + Y_0^2 \tag{98}$$

It corresponds to the square of the distance from the center C_0 of the trajectory to the origin. Using (81), we easily find:

$$\begin{aligned}\Gamma^2 &= \frac{\hbar}{\mu\omega_c}\left(a_l a_l^\dagger + a_l^\dagger a_l\right)\\ &= \frac{\hbar}{\mu\omega_c}\left(2N_l + 1\right)\end{aligned} \tag{99}$$

The state $|\chi_{n_r,\,n_l}\rangle$ is therefore an eigenstate of Γ^2 with the eigenvalue $\dfrac{\hbar}{\mu\omega_c}(2n_l + 1)$; the fact that this value can never go to zero is related to the non-commutativity of the operators X_0 and Y_0.

Comment:

The operator L_z, according to (75) and (99), is given by:

$$L_z = \hbar(N_r - N_l) = \hbar\left[\frac{H_\perp}{\hbar\omega_c} - \frac{1}{2} - \frac{\mu\omega_c}{2\hbar}\Gamma^2 + \frac{1}{2}\right] \tag{100}$$

that is, according to (88):

$$L_z = \frac{\mu\omega_c}{2}\left(\Sigma^2 - \Gamma^2\right) = \frac{qB}{2}\left(\Gamma^2 - \Sigma^2\right) \tag{101}$$

which is the equivalent of the classical relation (27).

γ. *Other types of stationary states*

Any linear combination of vectors $|\chi_{n_r,\,n_l}\rangle$ associated with the same value of n_r is an eigenstate of $H_\perp$ and therefore possesses the properties stated in § 3-c-α. By a suitable choice of the coefficients of the linear combination, one can obtain stationary states that possess other interesting properties as well.

We know, for example (§ 3-b-β), that X_0 and Y_0 are constants of the motion. However, since X_0 and Y_0 do not commute, there are no eigenstates common to these two operators. This means that, in quantum mechanics, it is not possible to obtain a state in which the two coordinates of the point C_0 are known.

To construct the eigenstates common to $H_\perp$ and X_0, we can use the properties of the one-dimensional harmonic oscillator; formula (81a) shows that X_0 has the same expression, to within a constant factor, as the position operator X_l of a one-dimensional oscillator whose destruction operator is a_l:

$$X_0 = \frac{1}{\beta\sqrt{2}}\hat{X}_l \tag{102}$$

Since we know the wave functions $\hat{\varphi}_k(\hat{x})$ associated with the stationary states $|\hat{\varphi}_k\rangle$ of a one-dimensional harmonic oscillator (*cf.* Complement B_V, § 2-b), we know how to write the eigenvectors $|\hat{x}\rangle$ of the position operator as linear combinations of the states $|\hat{\varphi}_k\rangle$:

$$\begin{aligned}|\hat{x}\rangle &= \sum_{k=0}^{\infty} |\hat{\varphi}_k\rangle\langle\hat{\varphi}_k|\hat{x}\rangle \\ &= \sum_{k=0}^{\infty} \hat{\varphi}_k^*(\hat{x})|\hat{\varphi}_k\rangle\end{aligned} \tag{103}$$

In order to obtain the eigenstates common to $H_\perp$ and X_0 it suffices to apply this result to the states $|\chi_{n_r,\,n_l=k}\rangle$; the vector:

$$|\xi_{n,\,x_0}\rangle = \sum_{k=0}^{\infty} \hat{\varphi}_k^*(\beta\sqrt{2}\,x_0)\ |\chi_{n_r=n,\,n_l=k}\rangle \tag{104}$$

is a common eigenvector of $H_\perp$ and X_0 with the eigenvalues $(n+1/2)\hbar\omega_c$ and x_0.

The eigenstates $|\eta_{n,\,y_0}\rangle$ common to $H_\perp$ and Y_0 can be found in an analogous fashion. Relation (81b) indicates that Y_0 is proportional to the momentum operator P_l of the ficticious one-dimensional oscillator just used:

$$Y_0 = \frac{1}{\beta\sqrt{2}}\hat{P}_l \tag{105}$$

Consequently [see formula (20) of Complement D_V]:

$$|\eta_{n,\,y_0}\rangle = \sum_{k=0}^{\infty} i^k\hat{\varphi}_k^*(\beta\sqrt{2}\,y_0)\ |\chi_{n_r=n,\,n_l=k}\rangle \tag{106}$$

We have just constructed the states in which either X_0 or Y_0 is perfectly well-defined. We can also determine the stationary states in which the product $\Delta X_0.\Delta Y_0$ reaches its minimal value, given by (83). For a one-dimensional harmonic oscillator, we studied in Complement G_V the states in which the product $\Delta\hat{X}.\Delta\hat{P}$ is minimal; these are the quasi-classical states, given by:

$$|\alpha\rangle = \sum_{k=0}^{\infty} c_k(\alpha)|\varphi_k\rangle \tag{107}$$

with:

$$c_k(\alpha) = \frac{\alpha^k}{\sqrt{k!}} \mathrm{e}^{-|\alpha|^2/2} \tag{108}$$

In these states:

$$\Delta\hat{X} = \Delta\hat{P} = \frac{1}{\sqrt{2}} \tag{109}$$

It follows that, in the case which interests us here, the state:

$$|\theta_{n,\,\alpha_0}\rangle = \sum_{k=0}^{\infty} c_k(\alpha_0) |\chi_{n_r=n,\,n_l=k}\rangle \tag{110}$$

yields, for X_0 and Y_0, root mean square deviations:

$$\Delta X_0 = \Delta Y_0 = \frac{1}{2\beta} \tag{111}$$

The product $\Delta X_0.\Delta Y_0$ is therefore minimal.

Comment:

Since the magnetic field is uniform, the physical problem we are considering is invariant with respect to translation. Thus far, this symmetry has been masked by the choice of the particular gauge (12), which gives the origin O a privileged position with respect to all other points in space. Consequently, neither the Hamiltonian H nor its eigenstates are invariant with respect to translation. We know, however (*cf.* Complement $\mathrm{H_{III}}$) that the physical predictions of quantum mechanics are gauge-invariant. These predictions must remain the same if, by a change of gauge, we give a point other than O a privileged position. Consequently, the translation symmetry must reappear when we study the physical properties of a given state.

To show this more precisely, let us assume that, at a given instant, the state of the particle is characterized in the gauge (12) by the ket $|\psi\rangle$ with which the wave function $\langle\mathbf{r}|\psi\rangle = \psi(\mathbf{r})$ is associated. We then perform a translation $\mathcal{T}$ defined by the vector $\mathbf{a}$, and consider the ket $|\psi_T\rangle$ defined by:

$$|\psi_T\rangle = \mathrm{e}^{-\frac{i}{\hbar}\mathbf{P}.\mathbf{a}} |\psi\rangle \tag{112}$$

with which, according to the results of Complement $\mathrm{E_{II}}$, is associated the wave function:

$$\psi_T(\mathbf{r}) = \langle\mathbf{r}|\psi_T\rangle = \psi(\mathbf{r}-\mathbf{a}) \tag{113}$$

The same translation can be applied to the vector potential, which becomes:

$$\mathbf{A}_T(\mathbf{r}) = \mathbf{A}(\mathbf{r}-\mathbf{a}) = -\frac{1}{2}(\mathbf{r}-\mathbf{a}) \times \mathbf{B} \tag{114}$$

$\mathbf{A}_T(\mathbf{r})$ clearly describes the same magnetic field as $\mathbf{A}(\mathbf{r})$. Since the physical properties attached to a given state vector depend only on this state vector and the potential $\mathbf{A}$ chosen, they must undergo the translation $\mathcal{T}$ when $\psi(\mathbf{r})$ and $\mathbf{A}(\mathbf{r})$ are replaced by expressions (113) and (114). It is simple to use these relations to obtain the expression for the probability density associated with $|\psi_T\rangle$:

$$\rho_T(\mathbf{r}) = |\psi_T(\mathbf{r})|^2 = |\psi(\mathbf{r}-\mathbf{a})|^2 = \rho(\mathbf{r}-\mathbf{a}) \tag{115}$$

and that for the current $\mathbf{J}_T(\mathbf{r})$, calculated with the vector potential $\mathbf{A}_T(\mathbf{r})$:

$$\begin{aligned}\mathbf{J}_T(\mathbf{r}) &= \frac{1}{2\mu}\left\{\psi_T^*(\mathbf{r})\left[\frac{\hbar}{i}\nabla + \frac{q}{2}(\mathbf{r}-\mathbf{a})\times\mathbf{B}\right]\psi_T(\mathbf{r}) + c.c.\right\} \\ &= \frac{1}{2\mu}\left\{\psi^*(\mathbf{r}-\mathbf{a})\left[\frac{\hbar}{i}\nabla + \frac{q}{2}(\mathbf{r}-\mathbf{a})\times\mathbf{B}\right]\psi(\mathbf{r}-\mathbf{a}) + c.c.\right\} \\ &= \mathbf{J}(\mathbf{r}-\mathbf{a}) \end{aligned} \tag{116}$$

[where $\mathbf{J}(\mathbf{r})$ is the probability current associated with $\psi(\mathbf{r})$ in the gauge (12)]. The ket $|\psi_T\rangle$ therefore describes, *in the new gauge* $\mathbf{A}_T(\mathbf{r})$, a state whose physical properties are related by the translation $\mathcal{T}$ to those corresponding to the ket $|\psi\rangle$ in the gauge $\mathbf{A}(\mathbf{r})$.

Let us show, moreover, that the translation of a possible motion yields another possible motion; this will conclude the proof of the translation invariance of the problem. To do so, consider the Schrödinger equation in the $\{|\mathbf{r}\rangle\}$ representation, in the gauge $\mathbf{A}(\mathbf{r})$:

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},\,t) = \frac{1}{2\mu}\left[\frac{\hbar}{i}\nabla - q\mathbf{A}(\mathbf{r})\right]^2\psi(\mathbf{r},\,t) \tag{117}$$

Changing $\mathbf{r}$ to $\mathbf{r}-\mathbf{a}$ in this equation, we obtain, using (113) and (114):

$$i\hbar\frac{\partial}{\partial t}\;\psi_T(\mathbf{r},\,t) = \frac{1}{2\mu}\left[\frac{\hbar}{i}\nabla - q\mathbf{A}_T(\mathbf{r})\right]^2\psi_T(\mathbf{r},\,t) \tag{118}$$

The operator appearing on the right-hand side of (118) is none other than the Hamiltonian in the gauge $\mathbf{A}_T(\mathbf{r})$. Consequently, if $\psi(\mathbf{r},\,t)$ describes, in the gauge $\mathbf{A}(\mathbf{r})$, a possible motion of the system, $\psi_T(\mathbf{r},\,t)$ describes, in the equivalent gauge $\mathbf{A}_T(\mathbf{r})$, another possible motion, which, according to what we have just shown, is nothing more than the result of a translation of the first motion. In particular, if:

$$\psi(\mathbf{r},\,t) = \varphi(\mathbf{r})\,\mathrm{e}^{-iEt/\hbar}$$

is a stationary state [in the gauge $\mathbf{A}(\mathbf{r})$],

$$\psi_T(\mathbf{r},\,t) = \varphi_T(\mathbf{r})\,\mathrm{e}^{-iEt/\hbar}$$

is another stationnary state of the same energy [in the gauge $\mathbf{A}_T(\mathbf{r})$].

If we want to continue to use the gauge (12) after having performed the translation $\mathcal{T}$ on the physical state of the particle, we must describe the translated state by a mathematical ket $|\psi'_T\rangle$ which is different from $|\psi_T\rangle$. According to § 3-b-α of Complement H_{III}, the ket $|\psi'_T\rangle$ can be obtained from $|\psi_T\rangle$ by a unitary transformation:

$$|\psi'_T\rangle = T_\chi|\psi_T\rangle \tag{119}$$

The operator T_χ is given by:

$$T_\chi = \mathrm{e}^{\frac{i}{\hbar}q\chi(\mathbf{R})} \tag{120}$$

where $\chi(\mathbf{r})$ is the function characterizing the gauge transformation performed. Here, the potential after the gauge change is:

$$\mathbf{A}(\mathbf{r}) = -\frac{1}{2}\,\mathbf{r}\times\mathbf{B} = \mathbf{A}_T(\mathbf{r}) - \frac{1}{2}\,\mathbf{a}\times\mathbf{B} \tag{121}$$

so that:

$$\chi(\mathbf{r}) = -\frac{1}{2}\,\mathbf{r}\cdot(\mathbf{a}\times\mathbf{B}) \tag{122}$$

Substituting (112), (120) and (122) into (119), we finally obtain:

$$|\psi'_T\rangle = U(\mathbf{a})|\psi\rangle \tag{123}$$

with:

$$U(\mathbf{a}) = \mathrm{e}^{-i\frac{q}{2\hbar}\mathbf{R}\cdot(\mathbf{a}\times\mathbf{B})}\mathrm{e}^{-\frac{i}{\hbar}\mathbf{P}\cdot\mathbf{a}} \tag{124}$$

Therefore, if we remain in the gauge $\mathbf{A}(\mathbf{r})$, the translation operator is $U(\mathbf{a})$, given by (124).

The components of $\mathbf{R}$ and $\mathbf{P}$ along two perpendicular axes enter into formula (124); they therefore commute, so we can write:

$$U(\mathbf{a}) = \mathrm{e}^{-i\frac{q}{2\hbar}\mathbf{R}\cdot(\mathbf{a}\times\mathbf{B})-\frac{i}{\hbar}\mathbf{P}\cdot\mathbf{a}} \tag{125}$$

When $\mathbf{a}$ is a vector of the xOy plane, a simple calculation, using formulas (10) and (69), yields:

$$U(\mathbf{a}) = \mathrm{e}^{i\frac{q}{\hbar}(\mathbf{a}\times\mathbf{R}_0)\cdot\mathbf{B}} \tag{126}$$

with:

$$\mathbf{R}_0 = X_0\,\mathbf{e}_x + Y_0\,\mathbf{e}_y \tag{127}$$

The operators X_0 and Y_0 (coordinates of the center of the circle) are therefore associated with translations along Oy and Ox respectively.

3-d. Time evolution

α. Mean values of the observables

We have already encountered a certain number of physical quantities that are constants of the motion: X_0, Y_0, Θ_z, Σ^2. Whatever the state of the system, their mean values are time-independent.

Let us examine the time evolution of the mean values $\langle X\rangle$, $\langle Y\rangle$, $\langle V_x\rangle$, $\langle V_y\rangle$ and $\langle X'\rangle$, $\langle Y'\rangle$. We immediately see from the expressions given in § 3-b-β that the corresponding operators have matrix elements only between states $|\chi_{n_r,\,n_l}\rangle$ whose values of n_r differ by ±1 (or 0). The evolution of these mean values therefore involves only one Bohr frequency, which is none other than the cyclotron frequency $\omega_c/2\pi$ defined in (5).

This result is completely analogous to the one given by classical mechanics.

β. Quasi-classical states

Assume that at $t = 0$ the state of the particle is:

$$|\psi_\perp(0)\rangle = |\alpha_r,\,\alpha_l\rangle \tag{128}$$

where the ket $|\alpha_r,\,\alpha_l\rangle$ is defined by expression (70) of Complement $\mathrm{D_{VI}}$. Since expression (75) for $H_\perp$ involves N_r but not N_l, the state vector $|\psi_\perp(t)\rangle$ at the instant t is obtained by changing α_r to $\alpha_r\,\mathrm{e}^{-i\omega_c t}$:

$$|\psi_\perp(t)\rangle = \mathrm{e}^{-i\omega_c t/2}|\alpha_r\mathrm{e}^{-i\omega_c t},\,\alpha_l\rangle \tag{129}$$

[*cf.* expression (92) of Complement $\mathrm{G_V}$].

We set:

$$\begin{aligned}\alpha_r &= |\alpha_r|\mathrm{e}^{i\varphi_r}\\ \alpha_l &= |\alpha_l|\mathrm{e}^{i\varphi_l}\end{aligned} \tag{130}$$

Relations (80), (81) and (84) then show that:

$$\begin{cases}\langle X_0\rangle = \dfrac{1}{2\beta}(\alpha_l + \alpha_l^*) = \dfrac{|\alpha_l|}{\beta}\cos\varphi_l\\ \langle Y_0\rangle = \dfrac{i}{2\beta}(\alpha_l^* - \alpha_l) = \dfrac{|\alpha_l|}{\beta}\sin\varphi_l\end{cases} \tag{131}$$

$$\begin{cases}\langle X'\rangle(t) = \dfrac{1}{2\beta}(\alpha_r\,\mathrm{e}^{-i\omega_c t} + \alpha_r^*\,\mathrm{e}^{i\omega_c t}) = \dfrac{|\alpha_r|}{\beta}\cos(\omega_c t - \varphi_r)\\ \langle Y'\rangle(t) = \dfrac{i}{2\beta}(\alpha_r\,\mathrm{e}^{-i\omega_c t} - \alpha_r^*\,\mathrm{e}^{i\omega_c t}) = \dfrac{|\alpha_r|}{\beta}\sin(\omega_c t - \varphi_r)\end{cases} \tag{132}$$

and:

$$\begin{cases}\langle V_x\rangle(t) = -\dfrac{|\alpha_r|}{\beta}\omega_c\sin(\omega_c t - \varphi_r)\\ \langle V_y\rangle(t) = \dfrac{|\alpha_r|}{\beta}\omega_c\cos(\omega_c t - \varphi_r)\end{cases} \tag{133}$$

Moreover, the properties of the states $|\alpha\rangle$ imply that:

$$\begin{aligned}\langle H_\perp\rangle &= \hbar\omega_c\left(|\alpha_r|^2 + \frac{1}{2}\right)\\ \langle \Theta_z\rangle &= 2\hbar\left(|\alpha_r|^2 + \frac{1}{2}\right)\\ \langle \Sigma^2\rangle &= \frac{1}{\beta^2}\left(|\alpha_r|^2 + \frac{1}{2}\right)\end{aligned} \tag{134}$$

All these results are extremely close to those given by classical mechanics [*cf.* (4)]. We see that $|\alpha_r|$ is related to the radius σ of the classical trajectory, and φ_r to the initial phase φ_0, while $|\alpha_l|$ is related to the distance OC_0, and φ_l corresponds to the polar angle of the vector $\mathbf{OC}_0$.

Furthermore, the properties of the states $|\alpha\rangle$ can be used to show that:

$$\Delta X_0 = \Delta Y_0 = \Delta X' = \Delta Y' = \frac{1}{2\beta} \tag{135a}$$

$$\Delta V_x = \Delta V_y = \frac{\omega_c}{2\beta} \tag{135b}$$

(the products $\Delta X_0.\Delta Y_0$, $\Delta X'.\Delta Y'$ and $\Delta V_x.\Delta V_y$ therefore take on their minimal values), and:

$$\Delta H_\perp = \hbar\omega_c|\alpha_r| \qquad \Delta\Theta_z = 2\hbar|\alpha_r| \qquad \Delta\Sigma^2 = \frac{1}{\beta^2}|\alpha_r| \tag{136}$$

As for the deviations ΔX and ΔY, they can be calculated by using the fact that:

$$|\psi_\perp(t)\rangle = \mathrm{e}^{-i\omega_c t/2}\left|\alpha_x = \frac{\alpha_r\,\mathrm{e}^{-i\omega_c t} + \alpha_l}{\sqrt{2}},\ \alpha_y = \frac{i\alpha_r\,\mathrm{e}^{-i\omega_c t} - i\alpha_l}{\sqrt{2}}\right\rangle \tag{137}$$

[where $|\alpha_x,\ \alpha_y\rangle$ is defined by relation (66) of $\mathrm{D_{VI}}$], which yields:

$$\Delta X = \Delta Y = \sqrt{\frac{\hbar}{\mu\omega_c}} = \frac{1}{\beta\sqrt{2}} \tag{138}$$

(ΔP_x and ΔP_y can easily be obtained in the same way).

If the conditions:

$$|\alpha_r| \geq 1,\ |\alpha_l| \geq 1 \tag{139}$$

are satisfied, we see, therefore, that the various physical quantities (position, velocity, energy, ...) are, in relative value, very well defined. The states (129) therefore represent "quasi-classical" states of the charged particle placed in a uniform magnetic field.

Comment:

If $\alpha_r = 0$, we obtain:

$$\begin{cases} \langle H_\perp\rangle = \dfrac{1}{2}\hbar\omega_c \\ \Delta H_\perp = 0 \end{cases} \tag{140}$$

The states:

$$|\alpha_x,\ \alpha_y = -i\alpha_x\rangle \tag{141}$$

therefore correspond to the ground state.

References and suggestions for further reading:

Landau and Lifshitz (1.19), Chap. XVI, §§ 124 and 125; Ter Haar (1.23), Chap. 6.

Application to solid state physics: Mott and Jones (13.7), Chap. VI, § 6; Kittel (13.2), Chap. 8, p. 239 and Chap. 9, p. 290.

Complement F_{VI}

Exercises

1. Consider a system of angular momentum $j = 1$, whose state space is spanned by the basis $\{|+1\rangle, |0\rangle, |-1\rangle\}$ of three eigenvectors common to $\mathbf{J}^2$ (eigenvalue $2\hbar^2$) and J_z (respective eigenvalues $+\hbar$, 0 and $-\hbar$). The state of the system is:

$$|\psi\rangle = \alpha|+1\rangle + \beta|0\rangle + \gamma|-1\rangle$$

where α, β, γ are three given complex parameters.

a. Calculate the mean value $\langle \mathbf{J} \rangle$ of the angular momentum in terms of α, β and γ.

b. Give the expression for the three mean values $\langle J_x^2 \rangle$, $\langle J_y^2 \rangle$ and $\langle J_z^2 \rangle$ in terms of the same quantities.

2. Consider an arbitrary physical system whose four-dimensional state space is spanned by a basis of four eigenvectors $|j, m_z\rangle$ common to $\mathbf{J}^2$ and J_z ($j = 0$ or 1; $-j \leq m_z \leq +j$), of eigenvalues $j(j+1)\hbar^2$ and $m_z\hbar$, such that:

$$J_\pm|j, m_z\rangle = \hbar\sqrt{j(j+1) - m_z(m_z \pm 1)}|j, m_z \pm 1\rangle$$
$$J_+|j, j\rangle = J_-|j, -j\rangle = 0$$

a. Express in terms of the kets $|j, m_z\rangle$, the eigenstates common to $\mathbf{J}^2$ and J_x, to be denoted by $|j, m_x\rangle$.

b. Consider a system in the normalized state:

$$|\psi\rangle = \alpha|j = 1, m_z = 1\rangle + \beta|j = 1, m_z = 0\rangle + \gamma|j = 1, m_z = -1\rangle + \delta|j = 0, m_z = 0\rangle$$

(*i*) What is the probability of finding $2\hbar^2$ and $\hbar$ if $\mathbf{J}^2$ and J_x are measured simultaneously?

(*ii*) Calculate the mean value of J_z when the system is in the state $|\psi\rangle$, and the probabilities of the various possible results of a measurement bearing only on this observable.

(*iii*) Same questions for the observable $\mathbf{J}^2$ and for J_x.

(*iv*) J_z^2 is now measured; what are the possible results, their probabilities, and their mean value?

3. Let $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ be the angular momentum of a system whose state space is $\mathcal{E}_\mathbf{r}$. Prove the commutation relations:

$$[L_i, R_j] = i\hbar\,\varepsilon_{ijk}\,R_k$$
$$[L_i, P_j] = i\hbar\,\varepsilon_{ijk}\,P_k$$
$$[L_i, \mathbf{P}^2] = [L_i, \mathbf{R}^2] = [L_i, \mathbf{R} \cdot \mathbf{P}] = 0$$

where L_i, R_j, P_j denote arbitrary components of $\mathbf{L}$, $\mathbf{R}$, $\mathbf{P}$ in an orthonormal system, and ε_{ijk} is defined by:

$$\varepsilon_{ijk} \begin{cases} = 0 & \text{if two (or three) of the indices } i,\ j,\ k \text{ are equal} \\ = 1 & \text{if these indices are an even permutation of } x,\ y,\ z \\ = -1 & \text{if the permutation is odd.} \end{cases}$$

4. Rotation of a polyatomic molecule

Consider a system composed of N different particles, of positions $\mathbf{R}_1$, ..., $\mathbf{R}_m$, ..., $\mathbf{R}_N$, and momenta $\mathbf{P}_1$, ..., $\mathbf{P}_m$, ..., $\mathbf{P}_N$. We set:

$$\mathbf{J} = \sum_m \mathbf{L}_m$$

with:

$$\mathbf{L}_m = \mathbf{R}_m \times \mathbf{P}_m$$

a. Show that the operator $\mathbf{J}$ satisfies the commutation relations that define an angular momentum. Deduce from this that, if $\mathbf{V}$ and $\mathbf{V}'$ denote two ordinary vectors of three-dimensional space, then:

$$[\mathbf{J} \cdot \mathbf{V}\ ,\ \mathbf{J} \cdot \mathbf{V}'] = i\hbar(\mathbf{V} \times \mathbf{V}') \cdot \mathbf{J}$$

b. Calculate the commutators of $\mathbf{J}$ with the three components of $\mathbf{R}_m$ and with those of $\mathbf{P}_m$. Show that:

$$[\mathbf{J}\ ,\ \mathbf{R}_m \cdot \mathbf{R}_p] = 0$$

c. Prove that:

$$[\mathbf{J}\ ,\ \mathbf{J} \cdot \mathbf{R}_m] = 0$$

and deduce from this the relation:

$$[\mathbf{J} \cdot \mathbf{R}_m\ ,\ \mathbf{J} \cdot \mathbf{R}_{m'}] = i\hbar(\mathbf{R}_{m'} \times \mathbf{R}_m) \cdot \mathbf{J} = i\hbar\, \mathbf{J} \cdot (\mathbf{R}_{m'} \times \mathbf{R}_m)$$

We set:

$$\mathbf{W} = \sum_m a_m \mathbf{R}_m$$
$$\mathbf{W}' = \sum_m a'_m \mathbf{R}_m$$

where the coefficients a_m and a'_m are given. Show that:

$$[\mathbf{J} \cdot \mathbf{W}, \mathbf{J} \cdot \mathbf{W}'] = -i\hbar(\mathbf{W} \times \mathbf{W}') \cdot \mathbf{J}$$

Conclusion: what is the difference between the commutation relations of the components of $\mathbf{J}$ along fixed axes and those of the components of $\mathbf{J}$ along the moving axes of the system being studied?

d. Consider a molecule which is formed by N unaligned atoms whose relative distances are assumed to be invariant (a rigid rotator). $\mathbf{J}$ is the sum of the angular momenta of the atoms with respect to the center of mass of the molecule, situated at a fixed point O; the $Oxyz$ axes constitute a fixed orthonormal frame. The three principal inertial axes of the system are denoted by $O\alpha$, $O\beta$ and $O\gamma$, with the ellipsoid of inertia assumed to be an ellipsoid of revolution about $O\gamma$ (a symmetrical rotator). The rotational energy of the molecule is then:

$$H = \frac{1}{2}\left[\frac{J_\gamma^2}{I_{/\!/}} + \frac{J_\alpha^2 + J_\beta^2}{I_\perp}\right]$$

where J_α, J_β and J_γ are the components of $\mathbf{J}$ along the unit vectors $\mathbf{w}_\alpha$, $\mathbf{w}_\beta$ and $\mathbf{w}_\gamma$ of the moving axes $O\alpha$, $O\beta$, $O\gamma$ attached to the molecule, and $I_{/\!/}$ and $I_\perp$ are the corresponding moments of inertia. We grant that:

$$J_\alpha^2 + J_\beta^2 + J_\gamma^2 = J_x^2 + J_y^2 + J_z^2 = \mathbf{J}^2$$

(*i*) Derive the commutation relations of J_α, J_β, J_γ from the results of *c*.

(*ii*) We introduce the operators $N_\pm = J_\alpha \pm iJ_\beta$. Using the general arguments of Chapter VI, show that one can find eigenvectors common to $\mathbf{J}^2$ and J_γ, of eigenvalues $J(J+1)\hbar^2$ and $K\hbar$, with $K = -J, -J+1, \ldots J-1, J$.

(*iii*) Express the Hamiltonian H of the rotator in terms of $\mathbf{J}^2$ and J_γ^2. Find its eigenvalues.

(*iv*) Show that one can find eigenstates common to $\mathbf{J}^2$, J_z and J_γ, to be denoted by $|J, M, K\rangle$ [the respective eigenvalues are $J(J+1)\hbar^2$, $M\hbar$, $K\hbar$]. Show that these states are also eigenstates of H.

(*v*) Calculate the commutators of $J_\pm$ and $N_\pm$ with $\mathbf{J}^2$, J_z, J_γ. Derive from them the action of $J_\pm$ and $N_\pm$ on $|J, M, K\rangle$. Show that the eigenvalues of H are at least $2(2J+1)$-fold degenerate if $K \neq 0$, and $(2J+1)$-fold degenerate if $K = 0$.

(*vi*) Draw the energy diagram of the rigid rotator (J is an integer since $\mathbf{J}$ is a sum of orbital angular momenta; *cf.* Chapter X). What happens to this diagram when $I_{/\!/} = I_\perp$ (spherical rotator)?

5. A system whose state space is $\mathcal{E}_\mathbf{r}$ has for its wave function:

$$\psi(x, y, z) = N(x + y + z)\mathrm{e}^{-\mathbf{r}^2/\alpha^2}$$

where α, which is real, is given and N is a normalization constant.

a. The observables L_z and $\mathbf{L}^2$ are measured; what are the probabilities of finding 0 and $2\hbar^2$? Recall that:

$$Y_1^0(\theta, \varphi) = \sqrt{\frac{3}{4\pi}}\cos\theta$$

b. If one also uses the fact that:

$$Y_1^{\pm 1}(\theta, \varphi) = \mp\sqrt{\frac{3}{8\pi}}\,\sin\theta\,\mathrm{e}^{\pm i\varphi}$$

is it possible to predict directly the probabilities of all possible results of measurements of $\mathbf{L}^2$ and L_z in the system of wave function $\psi(x, y, z)$?

6. Consider a system of angular momentum $l = 1$. A basis of its state space is formed by the three eigenvectors of L_z: $|+1\rangle$, $|0\rangle$, $|-1\rangle$, whose eigenvalues are, respectively, $+\hbar$, 0, and $-\hbar$, and which satisfy:

$$L_\pm|m\rangle = \hbar\sqrt{2}\,|m \pm 1\rangle$$
$$L_+|1\rangle = L_-|-1\rangle = 0$$

This system, which possesses an electric quadrupole moment, is placed in an electric field gradient, so that its Hamiltonian can be written:

$$H = \frac{\omega_0}{\hbar}(L_u^2 - L_v^2)$$

where L_u and L_v are the components of $\mathbf{L}$ along the two directions Ou and Ov of the xOz plane that form angles of 45° with Ox and Oz; ω_0 is a real constant.

a. Write the matrix representing H in the $\{|+1\rangle, |0\rangle, |-1\rangle\}$ basis. What are the stationary states of the system, and what are their energies? (These states are to be written $|E_1\rangle$, $|E_2\rangle$, $|E_3\rangle$, in order of decreasing energies.)

b. At time $t = 0$, the system is in the state:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}[|+1\rangle - |-1\rangle]$$

What is the state vector $|\psi(t)\rangle$ at time t? At t, L_z is measured; what are the probabilities of the various possible results?

c. Calculate the mean values $\langle L_x\rangle(t)$, $\langle L_y\rangle(t)$ and $\langle L_z\rangle(t)$ at t. What is the motion performed by the vector $\langle\mathbf{L}\rangle$?

d. At t, a measurement of L_z^2 is performed.

(*i*) Do times exist when only one result is possible?

(*ii*) Assume that this measurement has yielded the result $\hbar^2$. What is the state of the system immediately after the measurement? Indicate, without calculation, its subsequent evolution.

7. Consider rotations in ordinary three-dimensional space, to be denoted by $\mathscr{R}_{\mathbf{u}}(\alpha)$, where $\mathbf{u}$ is the unit vector which defines the axis of rotation and α is the angle of rotation.

a. Show that, if M' is the transform of M under an infinitesimal rotation of angle ε, then:

$$\mathbf{OM}' = \mathbf{OM} + \varepsilon\mathbf{u} \times \mathbf{OM}$$

b. If **OM** is represented by the column vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, what is the matrix associated with $\mathscr{R}_{\mathbf{u}}(\varepsilon)$? Derive from it the matrices representing the components of the operator $\mathcal{M}$ defined by:

$$\mathscr{R}_{\mathbf{u}}(\varepsilon) = 1 + \varepsilon \mathcal{M}.\mathbf{u}$$

c. Calculate the commutators:

$$[\mathcal{M}_x\ ,\ \mathcal{M}_y] \quad ; \quad [\mathcal{M}_y\ ,\ \mathcal{M}_z] \quad ; \quad [\mathcal{M}_z\ ,\ \mathcal{M}_x]$$

What are the quantum mechanical analogues of the purely geometrical relations obtained?

d. Starting with the matrix representing $\mathcal{M}_z$, calculate the one that represents $\mathrm{e}^{\alpha \mathcal{M}_z}$; show that $\mathscr{R}_z(\alpha) = \mathrm{e}^{\alpha \mathcal{M}_z}$; what is the analogue of this relation in quantum mechanics?

8. Consider a particle in three-dimensional space, whose state vector is $|\psi\rangle$, and whose wave function is $\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle$. Let A be an observable that commutes with $\mathbf{L} = \mathbf{R} \times \mathbf{P}$, the orbital angular momentum of the particle. Assuming that A, $\mathbf{L}^2$ and L_z form a C.S.C.O. in $\mathcal{E}_{\mathbf{r}}$, call $|n, l, m\rangle$ their common eigenkets, whose eigenvalues are, respectively, a_n (the index n is assumed to be discrete), $l(l+1)\hbar^2$ and $m\hbar$.

Let $U(\varphi)$ be the unitary operator defined by:

$$U(\varphi) = \mathrm{e}^{-i\varphi L_z/\hbar}$$

where φ is a real dimensionless parameter. For an arbitrary operator K, we call $\tilde{K}$ the transform of K by the unitary operator $U(\varphi)$:

$$\tilde{K} = U(\varphi)\, K\, U^{\dagger}(\varphi)$$

a. We set $L_+ = L_x + iL_y$, $L_- = L_x - iL_y$. Calculate $\tilde{L}_+|n, l, m\rangle$ and show that L_+ and $\tilde{L}_+$ are proportional; calculate the proportionality constant. Same question for L_- and $\tilde{L}_-$.

b. Express $\tilde{L}_x$, $\tilde{L}_y$ and $\tilde{L}_z$ in terms of L_x, L_y and L_z. What geometrical transformation can be associated with the transformation of $\mathbf{L}$ into $\tilde{\mathbf{L}}$?

c. Calculate the commutators $[X \pm iY, L_z]$ and $[Z, L_z]$. Show that the kets $(X \pm iY)|n, l, m\rangle$ and $Z|n, l, m\rangle$ are eigenvectors of L_z and calculate their eigenvalues. What relation must exist between m and m' for the matrix element $\langle n', l', m'|X \pm iY|n, l, m\rangle$ to be non-zero? Same question for $\langle n', l', m'|Z|n, l, m\rangle$.

d. By comparing the matrix elements of $\widetilde{X \pm iY}$ and $\tilde{Z}$ with those of $X \pm iY$ and Z, calculate $\tilde{X}$, $\tilde{Y}$, $\tilde{Z}$ in terms of X, Y, Z. Give a geometrical interpretation.

9. Consider a physical system of fixed angular momentum l, whose state space is $\mathcal{E}_l$, and whose state vector is $|\psi\rangle$; its orbital angular momentum operator is denoted by $\mathbf{L}$. We assume that a basis of $\mathcal{E}_l$ is composed of $2l+1$ eigenvectors $|l, m\rangle$ of L_z ($-l \leq m \leq +l$), associated with the wave functions $f(r)Y_l^m(\theta, \varphi)$. We call $\langle \mathbf{L} \rangle = \langle \psi | \mathbf{L} | \psi \rangle$ the mean value of $\mathbf{L}$.

a. We begin by assuming that:

$$\langle L_x \rangle = \langle L_y \rangle = 0$$

Out of all the possible states of the system, what are those for which the sum $(\Delta L_x)^2 + (\Delta L_y)^2 + (\Delta L_z)^2$ is minimal? Show that, for these states, the root mean square deviation ΔL_α of the component of $\mathbf{L}$ along an axis making an angle α with Oz is given by:

$$\Delta L_\alpha = \hbar \sqrt{\frac{l}{2}} \sin \alpha$$

b. We now assume that $\langle \mathbf{L} \rangle$ has an arbitrary direction with respect to the $Oxyz$ axes. We denote by $OXYZ$ a frame whose OZ axis is directed along $\langle \mathbf{L} \rangle$, with the OY axis in the xOy plane.

(*i*) Show that the state $|\psi_0\rangle$ of the system for which $(\Delta L_x)^2 + (\Delta L_y)^2 + (\Delta L_z)^2$ is minimal is such that:

$$(L_x + iL_y)|\psi_0\rangle = 0$$
$$L_z|\psi_0\rangle = l\hbar\ |\psi_0\rangle$$

(*ii*) Let θ_0 be the angle between Oz and OZ, and φ_0, the angle between Oy and OY; prove the relations:

$$L_X + iL_Y = \cos^2 \frac{\theta_0}{2}\ \mathrm{e}^{-i\varphi_0} L_+ - \sin^2 \frac{\theta_0}{2}\ \mathrm{e}^{i\varphi_0} L_- - \sin\theta_0\ L_z$$
$$L_z = \sin \frac{\theta_0}{2} \cos \frac{\theta_0}{2}\ \mathrm{e}^{-i\varphi_0} L_+ + \sin \frac{\theta_0}{2} \cos \frac{\theta_0}{2}\ \mathrm{e}^{i\varphi_0} L_- + \cos\theta_0\ L_z$$

If we set:

$$|\psi_0\rangle = \sum_m d_m |l, m\rangle$$

show that:

$$d_m = \tan\frac{\theta_0}{2} \mathrm{e}^{i\varphi_0} \sqrt{\frac{l+m+1}{l-m}}\, d_{m+1}$$

Express d_m in terms of d_l, θ_0, φ_0 and l.

(*iii*) To calculate d_l, show that the wave function associated with $|\psi_0\rangle$ is $\psi_0(X, Y, Z) = c_l\ \dfrac{(X+iY)^l}{r^l}\ f(r)$ [where c_l is defined by equation (D-20) of Chapter VI], the

one associated with $|l, l\rangle$ being $c_l \frac{(x+iy)^l}{r^l} f(r)$. By replacing X, Y and Z in this expression for $\psi_0(X, Y, Z)$ by their values in terms of x, y, z, find the value of d_l and the relation:

$$d_m = \left(\sin\frac{\theta_0}{2}\right)^{l-m} \left(\cos\frac{\theta_0}{2}\right)^{l+m} \mathrm{e}^{-im\varphi_0} \sqrt{\frac{(2l)!}{(l+m)!(l-m)!}}$$

(*iv*) With the system in the state $|\psi_0\rangle$, L_z is measured. What are the probabilities of the various possible results? What is the most probable result? Show that, if l is much greater than 1, the results correspond to the classical limit.

10. Let $\mathbf{J}$ be the angular momentum operator of an arbitrary physical system whose state vector is $|\psi\rangle$.

a. Can states of the system be found for which the root mean square deviations ΔJ_x, ΔJ_y and ΔJ_z are simultaneously zero?

b. Prove the relation:

$$\Delta J_x . \Delta J_y \geq \frac{\hbar}{2} |\langle J_z \rangle|$$

and those obtained by cyclic permutation of x, y, z.

Let $\langle \mathbf{J} \rangle$ be the mean value of the angular momentum of the system. The $Oxyz$ axes are assumed to be chosen in such a way that $\langle J_x \rangle = \langle J_y \rangle = 0$. Show that:

$$(\Delta J_x)^2 + (\Delta J_y)^2 \geq \hbar |\langle J_z \rangle|$$

c. Show that the two inequalities proven in question *b*. both become equalities if and only if $J_+|\psi\rangle = 0$ or $J_-|\psi\rangle = 0$.

d. The system under consideration is a spinless particle for which $\mathbf{J} = \mathbf{L} = \mathbf{R} \times \mathbf{P}$. Show that it is not possible to have both $\Delta L_x . \Delta L_y = -\frac{\hbar}{2} |\langle L_z \rangle|$ and $(\Delta L_x)^2 + (\Delta L_y)^2 = \hbar |\langle L_z \rangle|$ unless the wave function of the system is of the form:

$$\psi(r, \theta, \varphi) = F(r, \sin\theta \, \mathrm{e}^{\pm i\varphi})$$

11. Consider a three-dimensional harmonic oscillator, whose state vector $|\psi\rangle$ is:

$$|\psi\rangle = |\alpha_x\rangle \otimes |\alpha_y\rangle \otimes |\alpha_z\rangle$$

where $|\alpha_x\rangle$, $|\alpha_y\rangle$ and $|\alpha_z\rangle$ are quasi-classical states (*cf.* Complement $\mathrm{G_V}$) for one-dimensional harmonic oscillators moving along Ox, Oy and Oz, respectively. Let $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ be the orbital angular momentum of the three-dimensional oscillator.

a. Prove:

$$\langle L_z \rangle = i\hbar \left(\alpha_x \alpha_y^* - \alpha_x^* \alpha_y \right)$$
$$\Delta L_z = \hbar \sqrt{|\alpha_x|^2 + |\alpha_y|^2}$$

and the analogous expressions for the components of $\mathbf{L}$ along Ox and Oy.

b. We now assume that:

$$\langle L_x \rangle = \langle L_y \rangle = 0 \quad , \qquad \langle L_z \rangle = \lambda\hbar > 0$$

Show that α_z must be zero. We then fix the value of λ. Show that, in order to minimize $\Delta L_x + \Delta L_y$, we must choose:

$$\alpha_x = -i\alpha_y = \sqrt{\frac{\lambda}{2}} \mathrm{e}^{i\varphi_0}$$

(where φ_0 is an arbitrary real number). Do the expressions $\Delta L_x . \Delta L_y$ and $(\Delta L_x)^2 + (\Delta L_y)^2$ in this case have minimum values compatible with the inequalities obtained in question *b*. of the preceding exercise?

c. Show that the state of a system for which the preceding conditions are satisfied is necessarily of the form:

$$|\psi\rangle = \sum_k c_k(\alpha_r) |\chi_{n_r=k,\, n_l=0,\, n_z=0}\rangle$$

with:

$$|\chi_{n_r=k,\, n_l=0,\, n_z=0}\rangle = \frac{\left(a_x^\dagger + i a_y^\dagger\right)^k}{\sqrt{2^k k!}} \; |\varphi_{n_x=0,\, n_y=0,\, n_z=0}\rangle$$
$$c_k(\alpha) = \frac{\alpha^k}{\sqrt{k!}} \; \mathrm{e}^{-|\alpha|^2/2} \quad ; \quad \alpha_r = \mathrm{e}^{i\varphi_0} \sqrt{\lambda}$$

(the results of Complement G_V and of § 4 of Complement D_{VI} can be used). Show that the angular dependence of $|\chi_{n_r=k,\, n_l=0,\, n_z=0}\rangle$ is $(\sin\theta\, \mathrm{e}^{i\varphi})^k$.

$\mathbf{L}^2$ is measured on a system in the state $|\psi\rangle$. Show that the probabilities of the various possible results are given by a Poisson distribution. What results can be obtained in a measurement of L_z that follows a measurement of $\mathbf{L}^2$ whose result was $l(l+1)\hbar^2$?

Exercise 4 :

Reference: Landau and Lifshitz (1.19), § 101; Ter Haar (1.23), §§ 8.13 and 8.14.

Chapter VII

Particle in a central potential. The hydrogen atom

A Stationary states of a particle in a central potential 804

A-1 Outline of the problem 804

A-2 Separation of variables 807

A-3 Stationary states of a particle in a central potential 810

B Motion of the center of mass and relative motion for a system of two interacting particles 812

B-1 Motion of the center of mass and relative motion in classical mechanics 812

B-2 Separation of variables in quantum mechanics 814

C The hydrogen atom 818

C-1 Introduction 818

C-2 The Bohr model 819

C-3 Quantum mechanical theory of the hydrogen atom 820

C-4 Discussion of the results 824

In this chapter, we shall consider the quantum mechanical properties of a particle placed in a central potential [that is, a potential $V(r)$ which depends only on the distance r from the origin]. This problem is closely related to the study of angular momentum presented in the preceding chapter. As we shall see in § A, the fact that $V(r)$ is invariant under any rotation about the origin means that the Hamiltonian H of the particle commutes with the three components of the orbital angular momentum operator $\mathbf{L}$. This considerably simplifies the determination of the eigenfunctions and eigenvalues of H, since these functions can be required to be eigenfunctions of $\mathbf{L}^2$ and L_z as well. This immediately defines their angular dependence, and the eigenvalue equation of H can be replaced by a differential equation involving only the variable r.

The importance of this problem derives from a property that will be established in § B: a two-particle system in which the interaction is described by a potential energy

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

that depends only on the relative positions of the particles can be reduced to a simpler problem involving only one fictitious particle. In addition, when the interaction potential of the two particles depends only on the distance between them, the fictitious particle's motion is governed by a central potential. This explains why the problem considered in this chapter is of such general interest: it is encountered in quantum mechanics whenever we investigate the behavior of an isolated system composed of two interacting particles.

In § C, we shall apply the general methods already described to a special case: that in which $V(r)$ is a Coulomb potential. The hydrogen atom, composed of an electron and a proton which electrostatically attract each other, supplies the simplest example of a system of this type. It is not the only one: in addition to hydrogen isotopes (deuterium, tritium), there are the hydrogenoid ions, which are systems composed of a single electron and a nucleus, such as the ions He^+, Li^{++}, etc... (other examples will be given in Complement A_{VII}). For these systems, we shall explicitly calculate the energies of the bound states and the corresponding wave functions. We also recall the fact that, historically, quantum mechanics was introduced in order to explain atomic properties (in particular, those of the simplest atom, hydrogen), which could not be accounted for by classical mechanics. The remarkable agreement between the theoretical predictions and the experimental observations constitutes one of the most spectacular successes of this branch of physics. Finally, it should be noted that the exact results concerning the hydrogen atom serve as the basis of all approximate calculations relating to more complex atoms (having several electrons).

A. Stationary states of a particle in a central potential

In this section, we consider a (spinless) particle of mass μ, subjected to a central force derived from the potential $V(r)$ (the center of force is chosen as the origin).

A-1. Outline of the problem

A-1-a. Review of some classical results

The force acting on the classical particle situated at the point M (with $\mathbf{OM} = \mathbf{r}$) is equal to:

$$\mathbf{F} = -\boldsymbol{\nabla} V(r) = -\frac{\mathrm{d}V}{\mathrm{d}r}\frac{\mathbf{r}}{r} \tag{A-1}$$

$\mathbf{F}$ is always directed towards O, and its moment with respect to this point is therefore always zero. If:

$$\boldsymbol{\mathcal{L}} = \mathbf{r} \times \mathbf{p} \tag{A-2}$$

is the angular momentum of the particle with respect to O, the angular momentum theorem implies that:

$$\frac{\mathrm{d}\boldsymbol{\mathcal{L}}}{\mathrm{d}t} = \mathbf{0} \tag{A-3}$$

$\boldsymbol{\mathcal{L}}$ is therefore a *constant of the motion*, so that the particle's trajectory is necessarily situated in the plane passing through O and perpendicular to $\boldsymbol{\mathcal{L}}$.

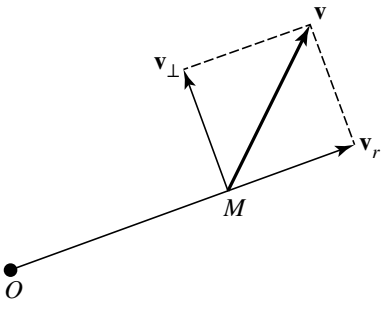


Figure 1: Radial component $\mathbf{v}_r$ and tangential component $\mathbf{v}_\perp$ of a particle's velocity.

Now let us consider (Fig. 1) the position (denoted by $\mathbf{OM} = \mathbf{r}$) and velocity $\mathbf{v}$ of the particle at the instant t. The two vectors $\mathbf{r}$ and $\mathbf{v}$ lie in the plane of the trajectory and the velocity $\mathbf{v}$ can be decomposed into the radial component $\mathbf{v}_r$ (along the axis defined by $\mathbf{r}$) and the tangential component $\mathbf{v}_\perp$ (along the axis perpendicular to $\mathbf{r}$). The radial velocity, the algebraic value of $\mathbf{v}_r$, is the time derivative of the distance of the particle from the point O:

$$v_{\boldsymbol{r}} = \frac{\mathrm{d}r}{\mathrm{d}t} \tag{A-4}$$

The tangential velocity can be expressed in terms of r and the angular momentum $\boldsymbol{\mathcal{L}}$, since:

$$|\mathbf{r} \times \mathbf{v}| = r|\mathbf{v}_\perp| \tag{A-5}$$

so that the modulus of the angular momentum $\boldsymbol{\mathcal{L}}$ is equal to:

$$|\boldsymbol{\mathcal{L}}| = |\mathbf{r} \times \mu\mathbf{v}| = \mu r|\mathbf{v}_\perp| \tag{A-6}$$

The total energy of the particle:

$$E = \frac{1}{2}\mu\mathbf{v}^2 + V(r) = \frac{1}{2}\mu\mathbf{v}_r^2 + \frac{1}{2}\mu\mathbf{v}_\perp^2 + V(r) \tag{A-7}$$

can be written:

$$E = \frac{1}{2}\mu\mathbf{v}_r^2 + \frac{\boldsymbol{\mathcal{L}}^2}{2\mu\, r^2} + V(r) \tag{A-8}$$

The classical Hamiltonian of the system is then:

$$\mathcal{H} = \frac{p_r^2}{2\mu} + \frac{\boldsymbol{\mathcal{L}}^2}{2\mu\, r^2} + V(r) \tag{A-9}$$

where:

$$p_r = \mu\frac{\mathrm{d}r}{\mathrm{d}t} \tag{A-10}$$

is the conjugate momentum of r, and $\boldsymbol{\mathcal{L}}^2$ must be expressed in terms of the variables r, θ, φ and their conjugate momenta p_r, p_θ, p_φ. One finds (*cf.* Appendix III, § 4-a):

$$\boldsymbol{\mathcal{L}}^2 = p_\theta^2 + \frac{1}{\sin^2\theta}\, p_\varphi^2 \tag{A-11}$$

In expression (A-9), the kinetic energy is broken into two terms: the radial kinetic energy and the kinetic energy of rotation about O. The reason is that, since $V(r)$ is independent of θ and φ in this case, the angular variables and their conjugate momenta appear only in the $\boldsymbol{\mathcal{L}}^2$ term. In fact, if we are interested in the evolution of r, we can use the fact that $\boldsymbol{\mathcal{L}}$ is a constant of the motion, and replace $\boldsymbol{\mathcal{L}}^2$ by a constant in expression (A-9). The Hamiltonian $\mathcal{H}$ then appears as a function only of the radial variables r and p_r ($\boldsymbol{\mathcal{L}}^2$ plays the role of a parameter), and the result is a differential equation involving only one variable, r:

$$\frac{\mathrm{d}p_r}{\mathrm{d}t} = \mu\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{\partial \mathcal{H}}{\partial r} \tag{A-12a}$$

that is:

$$\mu\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = \frac{\boldsymbol{\mathcal{L}}^2}{\mu\, r^3} - \frac{\mathrm{d}V}{\mathrm{d}r} \tag{A-12b}$$

It is just as if we had a one-dimensional problem (with r varying only between 0 and $+\infty$), with a particle of mass μ subjected to the "effective potential":

$$V_{\text{eff}}(r) = V(r) + \frac{\boldsymbol{\mathcal{L}}^2}{2\mu\, r^2} \tag{A-13}$$

We shall see that the situation is analogous in quantum mechanics.

A-1-b. The quantum mechanical Hamiltonian

In quantum mechanics, we want to solve the eigenvalue equation of the Hamiltonian H, the observable associated with the total energy. This equation is written, in the $\{|\mathbf{r}\rangle\}$ representation:

$$\left[-\frac{\hbar^2}{2\mu}\Delta + V(r)\right]\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}) \tag{A-14}$$

Since the potential V depends only on the distance r of the particle from the origin, spherical coordinates (*cf.* § D-1-a of Chapter VI) are best adapted to the problem. We therefore express the Laplacian Δ in spherical coordinates[1]:

$$\Delta = \frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2}\left(\frac{\partial^2}{\partial\,\theta^2} + \frac{1}{\tan\theta}\frac{\partial}{\partial\,\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\,\varphi^2}\right) \tag{A-15}$$

and look for eigenfunctions $\varphi(\mathbf{r})$ that are functions of the variables r, θ, φ.

[1] Expression (A-15) gives the Laplacian only for non-zero r. This is because of the privileged position of the origin in spherical coordinates; it can be seen, moreover, that expression (A-15) is not defined for $r = 0$.

If we compare expression (A-15) with the one for the operator $\mathbf{L}^2$ [formula (D-6a) of Chapter VI], we see that the quantum mechanical Hamiltonian H can be put in a form completely analogous to (A-9):

$$\boxed{H = -\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{2\mu\, r^2}\mathbf{L}^2 + V(r)} \tag{A-16}$$

The angular dependence of the Hamiltonian is contained entirely in the $\mathbf{L}^2$ term, which is an operator here. We could, in fact, perfect the analogy by defining an operator $P_{\boldsymbol{r}}$, which would allow us to write the first term of (A-16) like the one in (A-9).

We shall now show how one can solve the eigenvalue equation:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{2\mu\, r^2}\mathbf{L}^2 + V(r)\right]\varphi(r,\theta,\varphi) = E\,\varphi(r,\theta,\varphi) \tag{A-17}$$

A-2. Separation of variables

A-2-a. Angular dependence of the eigenfunctions

We know [*cf.* formulas (D-5) of Chapter VI] that the three components of the angular momentum operator $\mathbf{L}$ act only on the angular variables θ and φ; consequently, they commute with all operators acting only on the r-dependence. In addition, they commute with $\mathbf{L}^2$. Therefore, according to expression (A-16) for the Hamiltonian, *the three components of* $\mathbf{L}$ *are constants of the motion*[2] in the quantum mechanical sense:

$$[H, \mathbf{L}] = \mathbf{0} \tag{A-18}$$

Obviously, H also commutes with $\mathbf{L}^2$.

Although we have at our disposition four constants of the motion (L_x, L_y, L_z and $\mathbf{L}^2$), we cannot use all four of them to solve equation (A-17) because they do not commute with each other; we shall use only $\mathbf{L}^2$ and L_z. Since the three observables H, $\mathbf{L}^2$ and L_z commute, we can find a basis of the state space $\mathcal{E}_{\mathbf{r}}$ of the particle composed of eigenfunctions common to these three observables. We can, therefore, without restricting the generality of the problem outlined in § A-1 above, require the functions $\varphi(r,\theta,\varphi)$, solutions of equation (A-17), to be eigenfunctions of $\mathbf{L}^2$ and L_z as well. We must then solve the system of differential equations:

$$H\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}) \tag{A-19a}$$

$$\mathbf{L}^2\varphi(\mathbf{r}) = l(l+1)\hbar^2\,\varphi(\mathbf{r}) \tag{A-19b}$$

$$L_z\varphi(\mathbf{r}) = m\hbar\,\varphi(\mathbf{r}) \tag{A-19c}$$

But we already know the general form of the common eigenfunctions of $\mathbf{L}^2$ and L_z (Chap. VI, § D-1): the solutions $\varphi(\mathbf{r})$ of equations (A-19), corresponding to fixed values of l and m, are necessarily products of a function of r alone and the spherical harmonic $Y_l^m(\theta,\varphi)$:

$$\varphi(\mathbf{r}) = R(r)\,Y_l^m(\theta,\varphi) \tag{A-20}$$

[2] Equation (A-18) expresses the fact that H is a scalar operator with respect to rotations about the point O (see Complement $\mathrm{B_{VI}}$). This is true because the potential energy is invariant under rotations about O.

Whatever the radial function $R(r)$, $\varphi(\mathbf{r})$ is a solution of equations (A-19b) and (A-19c). The only problem which remains to be solved is therefore how to determine $R(r)$ such that $\varphi(\mathbf{r})$ is also an eigenfunction of H [equation (A-19a)].

A-2-b. The radial equation

We shall now substitute expressions (A-16) and (A-20) into equation (A-19a). Since $\varphi(\mathbf{r})$ is an eigenfunction of $\mathbf{L}^2$ with the eigenvalue $l(l+1)\hbar^2$, we see that $Y_l^m(\theta, \varphi)$ is a common factor on both sides. After simplifying, we obtain the radial equation:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r + \frac{l(l+1)\hbar^2}{2\mu\, r^2} + V(r)\right] R(r) = E\, R(r) \tag{A-21}$$

Actually, a solution of (A-21), substituted into (A-20), does not necessarily yield a solution of the eigenvalue equation (A-14) of the Hamiltonian. As we have already pointed out (*cf.* footnote 1), expression (A-15) for the Laplacian is not necessarily valid at $r = 0$. We must therefore make sure that the behavior of the solutions $R(r)$ of (A-21) at the origin is sufficiently regular for (A-20) to be in fact a solution of (A-14).

Instead of solving the partial differential equation (A-17) involving the three variables r, θ, φ, we must now solve a differential equation involving only the variable r, but dependent on a parameter l: we are looking for eigenvalues and eigenfunctions of an operator H_l which is different for each value of l.

In other words, we consider separately, in the state space $\mathcal{E}_{\mathbf{r}}$, the subspaces $\mathcal{E}(l, m)$ corresponding to fixed values of l and m (*cf.* Chap. VI, § C-3-a), studying the eigenvalue equation of H in each of these subspaces (which is possible because H commutes with $\mathbf{L}^2$ and L_z). The equation to be solved depends on l, but not on m; it is therefore the same in the $(2l+1)$ subspaces $\mathcal{E}(l, m)$ associated with a given value of l. We shall denote by $E_{k,l}$ the eigenvalues of H_l, that is, the eigenvalues of the Hamiltonian H inside a given subspace $\mathcal{E}(l, m)$. The index k, which can be discrete or continuous, represents the various eigenvalues associated with the same value of l. As for the eigenfunctions of H_l, we shall label them with the same two indices as the eigenvalues: $R_{k,l}(r)$. It is not obvious that this is sufficient: several radial functions might exist and be eigenfunctions of the same operator H_l with the same eigenvalue $E_{k,l}$; we shall see in § A-3-b that this is not the case and that, consequently, the two indices k and l are sufficient to characterize the different radial functions. We shall therefore rewrite equation (A-21) in the form:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\, r + \frac{l(l+1)\hbar^2}{2\mu\, r^2} + V(r)\right] R_{k,l}(r) = E_{k,l}\, R_{k,l}(r) \tag{A-22}$$

We can simplify the differential operator to be studied by a change in functions. We set:

$$R_{k,l}(r) = \frac{1}{r} u_{k,l}(r) \tag{A-23}$$

Multiplying both sides of (A-22) by r, we obtain for $u_{k,l}(r)$ the following differential equation:

$$\boxed{\left[-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{l(l+1)\hbar^2}{2\mu\, r^2} + V(r)\right] u_{k,l}(r) = E_{k,l}\, u_{k,l}(r)} \tag{A-24}$$

This equation is analogous to the one we would have to solve if, in a *one-dimensional* problem, a particle of mass μ were moving in an *effective potential* $V_{\text{eff}}(r)$:

$$V_{\text{eff}}(r) = V(r) + \frac{l(l+1)\hbar^2}{2\mu\, r^2} \tag{A-25}$$

Nevertheless, we must not lose sight of the fact that the variable r can take on only *non-negative* real values. The term $l(l+1)\hbar^2/2\mu r^2$ which is added to the potential $V(r)$ is always positive or zero; the corresponding force (equal to minus the gradient of this term) always tends to repel the particle from the force center O; this is why this term is called the *centrifugal potential* (or centrifugal barrier). Figure 2 represents the shape of the effective potential $V_{\text{eff}}(r)$ for various values of l in the case where $V(r)$ is an attractive Coulomb potential [$V(r) = -e^2/r$]: for $l \geq 1$, the presence of the centrifugal term, which predominates for small r values, causes V_{eff} to be repulsive for short distances.

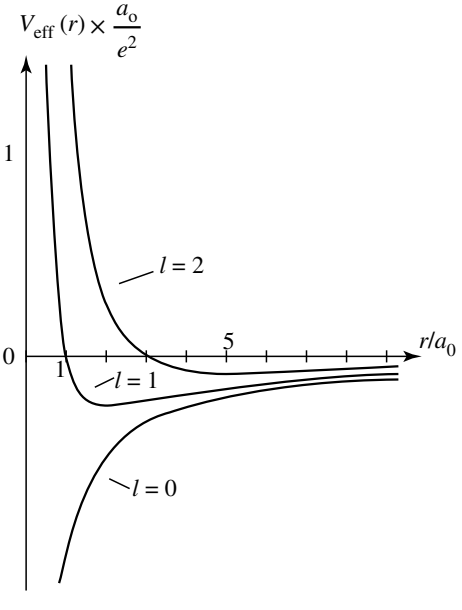


Figure 2: Shape of the effective potential $V_{\text{eff}}(r)$ for the first values of l in the case where $V(r) = -e^2/r$. When $l = 0$, $V_{\text{eff}}(r)$ is simply equal to $V(r)$. When l takes on the values 1, 2, etc., $V_{\text{eff}}(r)$ is obtained by adding to $V(r)$ the centrifugal potential $l(l+1)\hbar^2/2\mu\, r^2$, which approaches $+\infty$ when r approaches zero.

A-2-c. Behavior of the solutions of the radial equation at the origin

We have already pointed out that it is necessary to examine the behavior of the solutions $R(r)$ of the radial equation (A-21) at the origin in order to know if they are really solutions of (A-14).

We shall assume that when r approaches zero, the potential $V(r)$ remains finite, or at least approaches infinity less rapidly than $1/r$ (this hypothesis is true in most cases encountered in physics and, in particular, in the case of the Coulomb potential, to be studied in § C). We shall consider a solution of (A-22) and assume that it behaves at the origin like r^s:

$$R_{k,l}(r) \underset{r\to 0}{\sim} C\, r^s \tag{A-26}$$

Substituting (A-26) into (A-22), and setting the coefficient of the dominant term equal to zero, we obtain the equation:

$$-s(s+1)+l(l+1)=0 \tag{A-27}$$

and, consequently:

$$\begin{cases} \text{either} \quad s=l & \text{(A-28a)} \\ \text{or} \quad s=-(l+1) & \text{(A-28b)} \end{cases}$$

For a given value of $E_{k,l}$, there are therefore two linearly independent solutions of the second-order equation (A-22), behaving at the origin like r^l and $1/r^{l+1}$, respectively. But those which behave like $1/r^{l+1}$ must be rejected, since it can be shown[3] that $\dfrac{1}{r^{l+1}}Y_l^m(\theta,\varphi)$ is not a solution of the eigenvalue equation (A-14) for $r=0$. From this, we see that acceptable solutions of (A-24) go to zero at the origin for all l, since:

$$u_{k,l}(r) \underset{r\to 0}{\sim} C\,r^{l+1} \tag{A-29}$$

Consequently, to equation (A-24) must be added the condition:

$$\boxed{u_{k,l}(0)=0} \tag{A-30}$$

Comment:

In equation (A-24), r, the distance of the particle from the origin, varies only between 0 and $+\infty$. However, thanks to condition (A-30), we can assume that we are actually dealing with a one-dimensional problem, in which the particle can theoretically move along the entire axis, but in which the effective potential is infinite for all negative values of the variable. We know that, in such a case, the wave function must be identically zero on the negative half-axis; condition (A-30) insures the continuity of the wave function at $r=0$.

A-3. Stationary states of a particle in a central potential

A-3-a. Quantum numbers

We can summarize the results of § 2 as follows: the fact that the potential $V(r)$ is independent of θ and φ makes it possible:

(*i*) to require the eigenfunctions of H to be *simultaneous eigenfunctions of* $\mathbf{L}^2$ *and* L_z, *which determines their angular dependence*:

$$\varphi_{k,l,m}(\mathbf{r})=R_{k,l}(r)\,Y_l^m(\theta,\varphi)=\frac{1}{r}u_{k,l}(r)\,Y_l^m(\theta,\varphi) \tag{A-31}$$

(*ii*) *to replace the eigenvalue equation of* H, an equation involving partial derivatives with respect to r, θ, φ, *by a differential equation involving only the variable* r and depending on a parameter l [equation (A-24)], with condition (A-30) imposed.

[3] This is because the Laplacian of $\dfrac{1}{r^{l+1}}Y_l^m(\theta,\varphi)$ involves the lth derivatives of $\delta(\mathbf{r})$ (*cf.* Appendix II, end of § 4).

These results can be compared with those recalled in § 1-a, of which they are the quantum mechanical analogues.

In principle, the functions $\varphi_{k,l,m}(r,\theta,\varphi)$ must be square-integrable, that is, normalizable:

$$\int |\varphi_{k,l,m}(r,\theta,\varphi)|^2 \, r^2 \, \mathrm{d}r \, \mathrm{d}\Omega = 1 \tag{A-32}$$

Their form (A-31) allows us to separate radial and angular integrations:

$$\int |\varphi_{k,l,m}(r,\theta,\varphi)|^2 \, r^2 \, \mathrm{d}r \, \mathrm{d}\Omega = \int_0^\infty r^2 \, \mathrm{d}r \, |R_{k,l}(r)|^2 \int \mathrm{d}\Omega \, |Y_l^m(\theta,\varphi)|^2 \tag{A-33}$$

But the spherical harmonics $Y_l^m(\theta,\varphi)$ are normalized with respect to θ and φ; condition (A-32) therefore reduces to:

$$\int_0^\infty r^2 \, \mathrm{d}r \, |R_{k,l}(r)|^2 = \int_0^\infty \mathrm{d}r \, |u_{k,l}(r)|^2 = 1 \tag{A-34}$$

Actually, we know that it is often convenient to accept eigenfunctions of the Hamiltonian that are not square-integrable. If the spectrum of H has a continuous part, we shall require only that the corresponding eigenfunctions be orthonormalized in the extended sense, that is, that they satisfy a condition of the form:

$$\int_0^\infty r^2 \, \mathrm{d}r \, R^*_{k',l}(r) \, R_{k,l}(r) = \int_0^\infty \mathrm{d}r \; u^*_{k',l}(r) \, u_{k,l}(r) = \delta(k'-k) \tag{A-35}$$

where k is a continuous index.

In (A-34) and (A-35), the integrals converge at their lower limit, $r = 0$ [condition (A-30)]. This is physically satisfying since the probability of finding the particle in any volume of finite dimensions is then always finite. It is therefore only because of the behavior of the wave functions for $r \longrightarrow \infty$ that, in the case of a continuous spectrum, the normalization integrals (A-35) diverge if $k = k'$.

Finally, the eigenfunctions of the Hamiltonian H of a particle placed in a central potential $V(r)$ depend on at least three indices [formula (A-31)]: $\varphi_{k,l,m}(r,\theta,\varphi) = R_{k,l}(r)\, Y_l^m(\theta,\varphi)$ is a simultaneous eigenfunction of H, $\mathbf{L}^2$ and L_z with the respective eigenvalues $E_{k,l}$, $l(l+1)\hbar^2$ and $m\hbar$. k is called the *radial* quantum number; l, the *azimuthal* quantum number; and m, the *magnetic* quantum number. The radial part $R_{k,l}(r) = \dfrac{1}{r} u_{k,l}(r)$ of the eigenfunction and the eigenvalue $E_{k,l}$ of H are independent of the magnetic quantum number and are given by the radial equation (A-24). The angular part of the eigenfunction depends only on l and m and not on k; it does not depend on the form of the potential $V(r)$.

A-3-b. Degeneracy of the energy levels

Finally, we shall consider the degeneracy of the energy levels, that is, of the eigenvalues of the Hamiltonian H. The $(2l+1)$ functions $\varphi_{k,l,m}(r,\theta,\varphi)$ with k and l fixed and m varying from $-l$ to $+l$ are eigenfunctions of H with the same eigenvalue $E_{k,l}$ [these $(2l+1)$ functions are clearly orthogonal, since they correspond to different eigenvalues of L_z]. The level $E_{k,l}$ is therefore at least $(2l+1)$-fold degenerate. This degeneracy,

which exists for all potentials $V(r)$, is called an *essential degeneracy*: it is due to the fact that the Hamiltonian H contains $\mathbf{L}^2$, but not L_z, which means that m does not appear in the radial equation[4]. It is also possible for one of the eigenvalues $E_{k,l}$ of the radial equation corresponding to a given value of l to be the same as an eigenvalue $E_{k',l'}$, associated with another radial equation, characterized by $l' \neq l$. This occurs only for certain potentials $V(r)$. The resulting degeneracies are called *accidental* (we shall see in § C that the energy states of the hydrogen atom present accidental degeneracies).

We must now show that, for a fixed value of l, the radial equation has at most one physically acceptable solution for each eigenvalue $E_{k,l}$. This actually results from condition (A-30). The radial equation, since it is a second-order differential equation, has *a priori* two linearly independent solutions for each value of $E_{k,l}$. Condition (A-30) eliminates one of them, so there is at most one acceptable solution for each value of $E_{k,l}$. We must also consider the behavior of the solutions for r approaching infinity; if $V(r) \longrightarrow 0$ when $r \longrightarrow \infty$, the negative values of $E_{k,l}$ for which the solution we have just chosen is also acceptable at infinity (that is, bounded) form a discrete set (see example of § C below and Complement $\mathrm{B_{VII}}$).

It follows from the preceding considerations that H, $\mathbf{L}^2$ *and* L_z *constitute a C.S.C.O.*[5]. If we fix three eigenvalues $E_{k,l}$, $l(l+1)\hbar^2$ and $m\hbar$, there corresponds to them a single function $\varphi_{k,l,m}(\mathbf{r})$. The eigenvalue of $\mathbf{L}^2$ indicates which equation yields the radial function; the eigenvalue of H determines this radial function $R_{k,l}(r)$ uniquely, as we have just seen; finally, there exists only one spherical harmonic $Y_l^m(\theta, \varphi)$ for a given l and m.

B. Motion of the center of mass and relative motion for a system of two interacting particles

Consider a system of two spinless particles, of masses m_1 and m_2 and positions $\mathbf{r}_1$ and $\mathbf{r}_2$. We assume that the forces exerted on these particles are derived from a *potential energy* $V(\mathbf{r}_1 - \mathbf{r}_2)$ *which depends only on* $\mathbf{r}_1 - \mathbf{r}_2$. This is true if there are no forces originating outside the system (that is, the system is isolated), and if the interactions between the two particles are derived from a potential. This potential must depend only on $\mathbf{r}_1 - \mathbf{r}_2$, since only the relative positions of the two particles are involved. We shall show that the study of such a system can be reduced to that of a single particle placed in the potential $V(\mathbf{r})$.

B-1. Motion of the center of mass and relative motion in classical mechanics

In classical mechanics, the two-particle system is described by the Lagrangian (*cf.* Appendix III):

$$\mathcal{L}(\mathbf{r}_1, \dot{\mathbf{r}}_1; \mathbf{r}_2, \dot{\mathbf{r}}_2) = T - V = \frac{1}{2} m_1 \dot{\mathbf{r}}_1^2 + \frac{1}{2} m_2 \dot{\mathbf{r}}_2^2 - V(\mathbf{r}_1 - \mathbf{r}_2) \tag{B-1}$$

[4] This essential degeneracy appears whenever the Hamiltonian is rotation-invariant (*cf.* Complement $\mathrm{B_{VI}}$). This is why it is encountered in numerous physical problems.

[5] Actually, we have not proven that these operators are observables, that is, that the set of $\varphi_{k,l,m}(\mathbf{r})$ form a basis in the state space $\mathcal{E}_{\mathbf{r}}$.

and the conjugate momenta of the six coordinates of the two particles are the components of the mechanical momenta:

$$\begin{aligned}\mathbf{p}_1 &= m_1\dot{\mathbf{r}}_1\\ \mathbf{p}_2 &= m_2\dot{\mathbf{r}}_2\end{aligned} \tag{B-2}$$

The study of the motion of the two particles is simplified by replacing the positions $\mathbf{r}_i$ by the three *coordinates of the center of mass* (or center of gravity):

$$\mathbf{r}_G = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \tag{B-3}$$

and the three *relative coordinates*[6]:

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{B-4}$$

Formulas (B-3) and (B-4) can be inverted to yield:

$$\begin{aligned}\mathbf{r}_1 &= \mathbf{r}_G + \frac{m_2}{m_1 + m_2}\mathbf{r}\\ \mathbf{r}_2 &= \mathbf{r}_G - \frac{m_1}{m_1 + m_2}\mathbf{r}\end{aligned} \tag{B-5}$$

The Lagrangian can then be written, in terms of the new variables $\mathbf{r}_G$ and $\mathbf{r}$:

$$\begin{aligned}\mathcal{L}(\mathbf{r}_G, \dot{\mathbf{r}}_G; \mathbf{r}, \dot{\mathbf{r}}) &= \frac{1}{2}m_1\left[\dot{\mathbf{r}}_G + \frac{m_2}{m_1 + m_2}\dot{\mathbf{r}}\right]^2 + \frac{1}{2}m_2\left[\dot{\mathbf{r}}_G - \frac{m_1}{m_1 + m_2}\dot{\mathbf{r}}\right]^2 - V(\mathbf{r})\\ &= \frac{1}{2}M\,\dot{\mathbf{r}}_G^2 + \frac{1}{2}\mu\,\dot{\mathbf{r}}^2 - V(\mathbf{r})\end{aligned} \tag{B-6}$$

where:

$$M = m_1 + m_2 \tag{B-7}$$

is the *total mass* of the system, and:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{B-8a}$$

is its *reduced mass* (the geometrical mean of the two masses m_1 and m_2), which is also given by:

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \tag{B-8b}$$

The conjugate momenta of the variables $\mathbf{r}_G$ and $\mathbf{r}$ are obtained by differentiating expression (B-6) with respect to the components of $\dot{\mathbf{r}}_G$ and $\dot{\mathbf{r}}$. Using (B-3), (B-4) and (B-2), we find:

$$\mathbf{p}_G = M\dot{\mathbf{r}}_G = m_1\dot{\mathbf{r}}_1 + m_2\dot{\mathbf{r}}_2 = \mathbf{p}_1 + \mathbf{p}_2 \tag{B-9a}$$

$$\mathbf{p} = \mu\dot{\mathbf{r}} = \frac{m_2\mathbf{p}_1 - m_1\mathbf{p}_2}{m_1 + m_2} \tag{B-9b}$$

[6]Definition (B-4) introduces a slight asymmetry between the two particles.

or:

$$\frac{\mathbf{p}}{\mu} = \frac{\mathbf{p}_1}{m_1} - \frac{\mathbf{p}_2}{m_2} \tag{B-9c}$$

$\mathbf{p}_G$ is the *total momentum* of the system, and $\mathbf{p}$ is called the *relative momentum* of the two particles.

We can express the classical Hamiltonian of the system in terms of the new dynamical variables we have just introduced:

$$\mathcal{H}(\mathbf{r}_G, \mathbf{p}_G; \mathbf{r}, \mathbf{p}) = \frac{\mathbf{p}_G^2}{2M} + \frac{\mathbf{p}^2}{2\mu} + V(\mathbf{r}) \tag{B-10}$$

This leads to the following equations of motion [formulas (27) of Appendix III]:

$$\dot{\mathbf{p}}_G = \mathbf{0} \tag{B-11}$$

$$\dot{\mathbf{p}} = -\boldsymbol{\nabla} V(\mathbf{r}) \tag{B-12}$$

The first term of expression (B-10) represents the kinetic energy of a fictitious particle whose mass M would be the sum $m_1 + m_2$ of the masses of the two real particles, whose position would be that of the center of mass of the system [formula (B-3)], and whose momentum $\mathbf{p}_G$ would be the total momentum $\mathbf{p}_1 + \mathbf{p}_2$ of the system. Equation (B-11) indicates that this fictitious particle is in uniform rectilinear motion (free particle). This result is well known in classical mechanics: the center of mass of a system of particles moves like a single particle whose mass is the total mass of this system, subjected to the resultant of all the forces exerted on the various particles. Here, this resultant is zero since the only forces present are internal ones obeying the principle of action and reaction.

Since the center of mass is in uniform rectilinear motion with respect to the initially chosen frame, the frame in which it is at rest ($\mathbf{p}_G = \mathbf{0}$) is also an inertial frame. In this *center of mass frame*, the first term of (B-10) is zero. The classical Hamiltonian, that is, the total energy of the system, then reduces to:

$$\mathcal{H}_r = \frac{\mathbf{p}^2}{2\mu} + V(\mathbf{r}) \tag{B-13}$$

$\mathcal{H}_r$ is the energy associated with the *relative motion* of the two particles. It is obviously this relative motion that is the most interesting in the study of the two interacting particles. It can be described by introducing a fictitious particle, called the *relative particle*: its mass is the reduced mass μ of the two real particles, its position is characterized by the relative coordinates $\mathbf{r}$, and its momentum is the relative momentum $\mathbf{p}$. Since its motion obeys equation (B-12), it behaves as if it were subjected to a potential $V(\mathbf{r})$ equal to the potential energy of interaction between the two real particles.

The study of the relative motion of two interacting particles therefore reduces to that of the motion of a single fictitious particle, characterized by formulas (B-4), (B-8) and (B-9c). This last equation expresses the fact that the velocity $\mathbf{p}/\mu$ of the relative particle is indeed the difference between the velocities of the two particles, that is, their relative velocity.

B-2. Separation of variables in quantum mechanics

The considerations of the preceding section can easily be transposed to quantum mechanics, as we shall now show.

B-2-a. Observables associated with the center of mass and the relative particle

The operators $\mathbf{R}_1$, $\mathbf{P}_1$ and $\mathbf{R}_2$, $\mathbf{P}_2$, which describe the positions and momenta of the two particles of the system, satisfy the canonical commutation relations:

$$[X_1, P_{1x}] = i\hbar$$
$$[X_2, P_{2x}] = i\hbar \tag{B-14}$$

with analogous expressions for the components along Oy and Oz. All the observables labeled by the index 1 commute with all those of index 2, and all the observables relating to one of the axes Ox, Oy or Oz commute with those corresponding to another one of these axes.

Now let us define the observables $\mathbf{R}_G$ and $\mathbf{R}$ by formulas similar to (B-3) and (B-4):

$$\mathbf{R}_G = \frac{m_1\mathbf{R}_1 + m_2\mathbf{R}_2}{m_1 + m_2} \tag{B-15a}$$
$$\mathbf{R} = \mathbf{R}_1 - \mathbf{R}_2 \tag{B-15b}$$

and the observables $\mathbf{P}_G$ and $\mathbf{P}$ by formulas similar to (B-9):

$$\mathbf{P}_G = \mathbf{P}_1 + \mathbf{P}_2 \tag{B-16a}$$
$$\mathbf{P} = \frac{m_2\mathbf{P}_1 - m_1\mathbf{P}_2}{m_1 + m_2} \tag{B-16b}$$

It is easy to calculate the various commutators of these new observables. The results are as follows:

$$[X_G, P_{Gx}] = i\hbar \tag{B-17a}$$
$$[X, P_x] = i\hbar \tag{B-17b}$$

with analogous expressions for the components along Oy and Oz; all the other commutators are zero. Consequently, $\mathbf{R}$ and $\mathbf{P}$, like $\mathbf{R}_G$ and $\mathbf{P}_G$, satisfy canonical commutation relations. Moreover, every observable of the set $\{\mathbf{R}, \mathbf{P}\}$ commutes with every observable of the set $\{\mathbf{R}_G, \mathbf{P}_G\}$.

We can also interpret $\mathbf{R}$ *and* $\mathbf{P}$, *on the one hand, and* $\mathbf{R}_G$ *and* $\mathbf{P}_G$, *on the other, as being the position and momentum observables of two distinct fictitious particles.*

B-2-b. Eigenvalues and eigenfunctions of the Hamiltonian

The Hamiltonian operator of the system is obtained from formulas (B-1) and (B-2) and the quantization rules of Chapter III:

$$H = \frac{\mathbf{P}_1^2}{2m_1} + \frac{\mathbf{P}_2^2}{2m_2} + V(\mathbf{R}_1 - \mathbf{R}_2) \tag{B-18}$$

Since definitions (B-15) and (B-16) are formally identical to (B-3), (B-4) and (B-9), and since all the momentum operators commute, a simple algebraic calculation yields the equivalent of expression (B-10).

$$H = \frac{\mathbf{P}_G^2}{2M} + \frac{\mathbf{P}^2}{2\mu} + V(\mathbf{R}) \tag{B-19}$$

The Hamiltonian H then appears as the sum of two terms:

$$H = H_G + H_r \tag{B-20}$$

with:

$$H_G = \frac{\mathbf{P}_G^2}{2M} \tag{B-21a}$$

$$H_r = \frac{\mathbf{P}^2}{2\mu} + V(\mathbf{R}) \tag{B-21b}$$

which commute, according to the results of § B-2-a:

$$[H_G, H_r] = 0 \tag{B-22}$$

H_G and H_r therefore commute with H. It follows that there exists a basis of eigenvectors of H that are also eigenvectors of H_G and H_r; we shall therefore look for solutions of the system:

$$\begin{aligned} H_G|\varphi\rangle &= E_G|\varphi\rangle \\ H_r|\varphi\rangle &= E_r|\varphi\rangle \end{aligned} \tag{B-23}$$

which immediately implies, according to (B-20):

$$H|\varphi\rangle = E|\varphi\rangle \tag{B-24}$$

with:

$$E = E_G + E_r \tag{B-25}$$

Consider the $\{|\mathbf{r}_G, \mathbf{r}\rangle\}$ representation, whose basis vectors are the eigenvectors common to the observables $\mathbf{R}_G$ and $\mathbf{R}$. In this representation, a state is characterized by a wave function $\varphi(\mathbf{r}_G, \mathbf{r})$ which is a function of six variables. The action of the operators $\mathbf{R}_G$ and $\mathbf{R}$ is expressed by the multiplication of the wave functions by the variables $\mathbf{r}_G$ and $\mathbf{r}$ respectively. $\mathbf{P}_G$ and $\mathbf{P}$ become the differential operators $\frac{\hbar}{i}\boldsymbol{\nabla}_G$ and $\frac{\hbar}{i}\boldsymbol{\nabla}$ (where $\boldsymbol{\nabla}_G$ denotes the set of three operators $\partial/\partial x_G$, $\partial/\partial y_G$ and $\partial/\partial z_G$). The state space $\mathcal{E}$ of the system can then be considered to be the tensor product $\mathcal{E}_{\mathbf{r}_G} \otimes \mathcal{E}_{\mathbf{r}}$ of the state space $\mathcal{E}_{\mathbf{r}_G}$ associated with the observable $\mathbf{R}_G$ and the space $\mathcal{E}_{\mathbf{r}}$ associated with $\mathbf{R}$. H_G and H_r then appear as the extensions into $\mathcal{E}$ of operators actually acting only in $\mathcal{E}_{\mathbf{r}_G}$ and $\mathcal{E}_{\mathbf{r}}$, respectively. We can therefore, as we saw in § F of Chapter II, find a basis of eigenvectors $|\varphi\rangle$ satisfying (B-23), in the form:

$$|\varphi\rangle = |\chi_G\rangle \otimes |\omega_{\boldsymbol{r}}\rangle \tag{B-26}$$

with:

$$\begin{cases} H_G|\chi_G\rangle = E_G|\chi_G\rangle \\ |\chi_G\rangle \in \mathcal{E}_{\mathbf{r}_G} \end{cases} \tag{B-27a}$$

$$\begin{cases} H_{\boldsymbol{r}}|\omega_{\boldsymbol{r}}\rangle = E_{\boldsymbol{r}}|\omega_{\boldsymbol{r}}\rangle \\ |\omega_{\boldsymbol{r}}\rangle \in \mathcal{E}_{\mathbf{r}} \end{cases} \tag{B-27b}$$

Writing these equations in the $\{|\mathbf{r}_G\rangle\}$ and $\{|\mathbf{r}\rangle\}$ representations respectively, we obtain:

$$-\frac{\hbar^2}{2M}\Delta_G\ \chi_G(\mathbf{r}_G) = E_G\ \chi_G(\mathbf{r}_G) \tag{B-28a}$$

$$\left[-\frac{\hbar^2}{2\mu}\Delta + V(\mathbf{r})\right]\omega_r(\mathbf{r}) = E_r\ \omega_r(\mathbf{r}) \tag{B-28b}$$

The first of these equations, (B-28a), shows that the particle associated with the center of mass of the system is free, as in classical mechanics. We know its solutions: they are, for example, the plane waves:

$$\chi_G(\mathbf{r}_G) = \frac{1}{(2\pi\hbar)^{3/2}}\ \mathrm{e}^{\frac{i}{\hbar}\mathbf{p}_G\cdot\mathbf{r}_G} \tag{B-29}$$

whose energy is equal to:

$$E_G = \frac{\mathbf{p}_G^2}{2M} \tag{B-30}$$

E_G can take on any positive value or zero; it is the kinetic energy corresponding to a translation of the system as a whole.

The more interesting equation from a physical point of view is the second one, (B-28b), which concerns the relative particle. It describes the behavior of the system of the two interacting particles in the center of mass frame. If the interaction potential of the two real particles depends only on the distance between them, $|\mathbf{r}_1 - \mathbf{r}_2|$, and not on the direction of the vector $\mathbf{r}_1 - \mathbf{r}_2$, the relative particle is subjected to a central potential $V(r)$; the problem is then reduced to the one treated in § A.

Comment:

The total angular momentum of the system of the two real particles is:

$$\mathbf{J} = \mathbf{L}_1 + \mathbf{L}_2 \tag{B-31}$$

with:

$$\begin{aligned}\mathbf{L}_1 &= \mathbf{R}_1 \times \mathbf{P}_1\\ \mathbf{L}_2 &= \mathbf{R}_2 \times \mathbf{P}_2\end{aligned} \tag{B-32}$$

It can easily be shown that it can also be written:

$$\mathbf{J} = \mathbf{L}_G + \mathbf{L} \tag{B-33}$$

where:

$$\begin{aligned}\mathbf{L}_G &= \mathbf{R}_G \times \mathbf{P}_G\\ \mathbf{L} &= \mathbf{R} \times \mathbf{P}\end{aligned} \tag{B-34}$$

are the angular momenta of the fictitious particles (according to the results of § B-2-a, $\mathbf{L}_G$ and $\mathbf{L}$ satisfy the commutation relations that characterize angular momenta, and the components of $\mathbf{L}$ commute with those of $\mathbf{L}_G$).

C. The hydrogen atom

C-1. Introduction

The hydrogen atom consists of a proton, of mass:

$$m_p \simeq 1.7 \times 10^{-27} \text{ kg} \tag{C-1}$$

and charge:

$$q \simeq 1.6 \times 10^{-19} \text{ Coulomb} \tag{C-2}$$

and of an electron, of mass:

$$m_e \simeq 0.91 \times 10^{-30} \text{ kg} \tag{C-3}$$

and charge $-q$. The interaction between these two particles is essentially electrostatic. The corresponding potential energy is:

$$V(r) \simeq -\frac{q^2}{4\pi\varepsilon_0}\frac{1}{r} = -\frac{e^2}{r} \tag{C-4}$$

where r denotes the distance between the two particles, and:

$$\frac{q^2}{4\pi\varepsilon_0} = e^2 \tag{C-5}$$

Using the results of § B, we confine ourselves to the study of this system in the center of mass frame. The classical Hamiltonian that describes the relative motion of the two particles is then[7]:

$$\mathcal{H}(\mathbf{r},\mathbf{p}) = \frac{\mathbf{p}^2}{2\mu} - \frac{e^2}{r} \tag{C-6}$$

Since $m_p \gg m_e$ [formulas (C-1) and (C-3)], the reduced mass μ of the system is very close to m_e:

$$\mu = \frac{m_e m_p}{m_e + m_p} \simeq m_e\left(1 - \frac{m_e}{m_p}\right) \tag{C-7}$$

(the correction term m_e/m_p is on the order of 1/1 800). This means that the center of mass of the system is practically at the position of the proton, and that the relative particle can be identified, to a very good approximation, with the electron. This is why we shall adopt the slightly inaccurate convention of calling the relative particle the electron and the center of mass the proton.

[7] Henceforth, we shall omit the index r which was used to label in § B the quantities corresponding to the relative motion.

C-2. The Bohr model

We shall briefly review the results of the Bohr model, which relate to the hydrogen atom. This model, which is based on the concept of a trajectory, is incompatible with the ideas of quantum mechanics. However, it allows us to introduce, in a very simple way, fundamental quantities such as the ionization energy E_I of the hydrogen atom and a parameter which characterizes atomic dimensions (the Bohr radius a_0). In addition, it so happens that the energies E_n given by the Bohr theory are the same as the eigenvalues of the Hamiltonian we shall calculate in § C-3. Finally, quantum mechanical theory is in agreement with some of the intuitive images of the Bohr model (§ C-4-c-β).

This semi-classical model is based on the hypothesis that the electron describes a *circular orbit* of radius r about the proton, obeying the following equations:

$$E = \frac{1}{2}\mu v^2 - \frac{e^2}{r} \tag{C-8}$$

$$\frac{\mu v^2}{r} = \frac{e^2}{r^2} \tag{C-9}$$

$$\mu v r = n\hbar \quad ; \quad n \text{ a positive integer} \tag{C-10}$$

The first two equations are classical ones. (C-8) expresses the fact that the total energy E of the electron is the sum of its kinetic energy $\mu v^2/2$ and its potential energy $-e^2/r$. (C-9) is none other than the fundamental equation of Newtonian dynamics (e^2/r^2 is the Coulomb force exerted on the electron, and v^2/r is the acceleration of its uniform circular motion). The third equation expresses the quantization condition, introduced empirically by Bohr in order to explain the existence of discrete energy levels: he postulated that only circular orbits satisfying this condition are possible trajectories for the electron. The different orbits, as well as the corresponding values of the various physical quantities, are labeled by the integer n associated with them.

A very simple algebraic calculation then yields the expressions for E_n, r_n and v_n:

$$E_n = -\frac{1}{n^2} E_I \tag{C-11a}$$

$$r_n = n^2 a_0 \tag{C-11b}$$

$$v_n = \frac{1}{n} v_0 \tag{C-11c}$$

with:

$$E_I = \frac{\mu e^4}{2\hbar^2} \tag{C-12a}$$

$$a_0 = \frac{\hbar^2}{\mu e^2} \tag{C-12b}$$

$$v_0 = \frac{e^2}{\hbar} \tag{C-12c}$$

When this model was proposed by Bohr, it marked an important step towards the understanding of atomic phenomena, since it yielded the correct values for the energy levels of the hydrogen atom. These values indeed follow the $1/n^2$ (the Balmer formula)

law indicated by expression (C-11a). Moreover, the experimentally measured *ionization energy* (the energy which must be supplied to the hydrogen atom in its ground state in order to remove the electron) is equal to the numerical value of E_I:

$$E_I \simeq 13.6 \text{ eV} \tag{C-13}$$

Finally, the *Bohr radius* a_0 indeed characterizes atomic dimensions:

$$a_0 \simeq 0.52 \text{ Å} \tag{C-14}$$

Comment:

Complement $\mathrm{C_I}$ shows how the uncertainty principle, applied to the hydrogen atom, explains the existence of a stable ground state and permits the evaluation of the order of magnitude of its energy and its spatial extension.

C-3. Quantum mechanical theory of the hydrogen atom

We shall now take up the question of the determination of the eigenvalues and eigenfunctions of the Hamiltonian H describing the relative motion of the proton and the electron in the center of mass frame [formula (C-6)]. In the $\{|\mathbf{r}\rangle\}$ representation, the eigenvalue equation of the Hamiltonian H is written:

$$\left[-\frac{\hbar^2}{2\mu}\Delta - \frac{e^2}{r}\right]\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}) \tag{C-15}$$

Since the potential $-e^2/r$ is central, we can apply the results of § A: the eigenfunctions $\varphi(\mathbf{r})$ are of the form:

$$\varphi_{k,l,m}(\mathbf{r}) = \frac{1}{r}u_{k,l}(r)\,Y_l^m(\theta,\varphi) \tag{C-16}$$

$u_{k,l}(r)$ is given by the radial equation, (A-24), that is:

$$\left[-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{l(l+1)\hbar^2}{2\mu\,r^2} - \frac{e^2}{r}\right]u_{k,l}(r) = E_{k,l}\,u_{k,l}(r) \tag{C-17}$$

We add to this equation condition (A-30):

$$u_{k,l}(0) = 0 \tag{C-18}$$

It can be shown that the spectrum of H includes a discrete part (negative eigenvalues) and a continuous part (positive eigenvalues). Consider Figure 3, which shows the effective potential for a given value of l (the figure is drawn for $l \neq 0$, but the reasoning remains valid for $l = 0$).

For a positive value of E, the classical motion is not bounded in space: for the value $E > 0$ chosen in Figure 3, it is limited on the left by the abscissa of point A, but it is not limited on the right[8]. As a result (*cf.* Complement $\mathrm{M_{III}}$) equation (C-17) has acceptable solutions for any $E > 0$. The spectrum of H is therefore continuous for $E > 0$, and the corresponding eigenfunctions are not square-integrable.

[8]For a $-1/r$ potential, the classical trajectories are conic sections; unbounded motion follows a hyperbola or a parabola.

On the other hand, for $E < 0$, the classical motion is bounded: it is confined to the region between the abscissas of the two points B and C[9]. We shall see later that equation (C-17) has acceptable solutions only for certain discrete values of E. The spectrum of H is therefore discrete for $E < 0$, and the corresponding eigenfunctions are square-integrable.

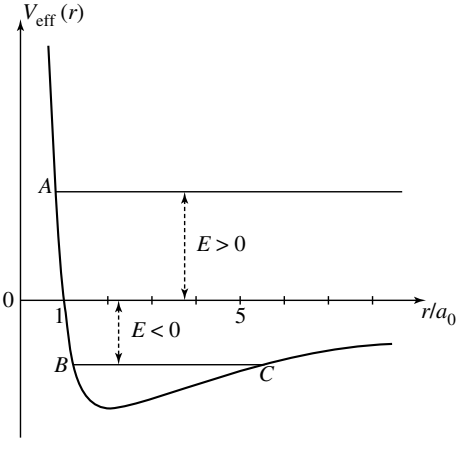


Figure 3: For a positive value of the energy E, the classical motion is not bounded. The spectrum of the quantum mechanical Hamiltonian H is therefore continuous for $E > 0$, and the corresponding eigenfunctions are not normalizable. On the other hand, for negative E, the classical motion is limited to the interval BC. The spectrum of H is therefore discrete for $E < 0$, and the corresponding eigenfunctions are normalizable.

C-3-a. Change of variables

To simplify the reasoning, we shall choose a_0 and E_I [formulas (C-12)] as the units of length and energy. That is, we shall introduce the dimensionless quantities:

$$\rho = r/a_0 \tag{C-19}$$

$$\lambda_{k,l} = \sqrt{-E_{k,l}/E_I} \tag{C-20}$$

(the quantity under the radical sign is positive, since we are looking for the bound states).

With expressions (C-12a) and (C-12b) for E_I and a_0 taken into account, the radial equation (C-17) becomes simply:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - \frac{l(l+1)}{\rho^2} + \frac{2}{\rho} - \lambda_{k,l}^2\right] u_{k,l}(\rho) = 0 \tag{C-21}$$

C-3-b. Solving the radial equation

In order to solve equation (C-21), we shall use the method illustrated in Complement $\mathrm{C_V}$, expanding $u_{k,l}(\rho)$ in a power series.

α. Asymptotic behavior

Let us determine the asymptotic behavior of $u_{k,l}(\rho)$ qualitatively. When ρ approaches infinity, the terms in $1/\rho$ and $1/\rho^2$ become negligible compared to the constant

[9] The classical trajectory is then an ellipse or a circle.

term $\lambda_{k,l}^2$, so that equation (C-21) practically reduces to:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - \lambda_{k,l}^2\right] u_{k,l}(\rho) = 0 \tag{C-22}$$

whose solutions are $\mathrm{e}^{\pm\rho\lambda_{k,l}}$. This argument is not rigorous, since we have completely neglected the terms in $1/\rho$ and $1/\rho^2$; actually, it can be shown that $u_{k,l}(\rho)$ is equal to $\mathrm{e}^{\pm\rho\lambda_{k,l}}$ multiplied by a power of ρ.

We shall later be led by physical considerations to require the function $u_{k,l}(\rho)$ to be bounded at infinity, and hence to reject the solutions of (C-21) whose asymptotic behavior is governed by $\mathrm{e}^{+\rho\lambda_{k,l}}$. This is why we perform the change of function:

$$u_{k,l}(\rho) = \mathrm{e}^{-\rho\lambda_{k,l}} y_{k,l}(\rho) \tag{C-23}$$

Although this change of function singles out $\mathrm{e}^{-\rho\lambda_{k,l}}$, it clearly does not eliminate solutions in $\mathrm{e}^{+\rho\lambda_{k,l}}$, which must be identified and then rejected at the end of the calculation. The differential equation that $y_{k,l}(\rho)$ must satisfy can easily be derived from (C-21):

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - 2\lambda_{k,l}\frac{\mathrm{d}}{\mathrm{d}\rho} + \left[\frac{2}{\rho} - \frac{l(l+1)}{\rho^2}\right]\right\} y_{k,l}(\rho) = 0 \tag{C-24}$$

Condition (C-18) must be associated with this equation, that is:

$$y_{k,l}(0) = 0 \tag{C-25}$$

β. Solutions in the form of power series

Consider the expansion of $y_{k,l}(\rho)$ in powers of ρ:

$$y_{k,l}(\rho) = \rho^s \sum_{q=0}^{\infty} c_q\, \rho^q \tag{C-26}$$

By definition, c_0 is the first non-zero coefficient of this expansion:

$$c_0 \neq 0 \tag{C-27}$$

Condition (C-25) implies that s is strictly positive.

We calculate $\dfrac{\mathrm{d}}{\mathrm{d}\rho}\, y_{k,l}(\rho)$ and $\dfrac{\mathrm{d}^2}{\mathrm{d}\rho^2}\, y_{k,l}(\rho)$ from (C-26):

$$\frac{\mathrm{d}}{\mathrm{d}\rho} y_{k,l}(\rho) = \sum_{q=0}^{\infty} (q+s)\; c_q\, \rho^{q+s-1} \tag{C-28a}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} y_{k,l}(\rho) = \sum_{q=0}^{\infty} (q+s)(q+s-1)\; c_q\, \rho^{q+s-2} \tag{C-28b}$$

To obtain the left-hand side of (C-24), we multiply expressions (C-26), (C-28a) and (C-28b) respectively by the factors $\left[\dfrac{2}{\rho} - \dfrac{l(l+1)}{\rho^2}\right]$, $-2\lambda_{k,l}$ and 1. According to (C-24), the series so determined must be identically zero, that is, all its coefficients must be zero.

The lowest order term is in ρ^{s-2}. Taking its coefficient as zero, we obtain:

$$[-l(l+1)+s(s-1)]c_0 = 0 \tag{C-29}$$

If we take (C-27) into account, we see that s can take on one of two values:

$$\begin{cases} s = l+1 & \text{(C-30a)} \\ s = -l & \text{(C-30b)} \end{cases}$$

(in agreement with the general result of § A-2-c). We have seen that only (C-30a) gives a behavior at the origin that can lead to an acceptable solution [condition (C-25)]. Setting the coefficient of the general term in ρ^{q+s-2} equal to zero, we obtain (with $s = l+1$) the following recurrence relation:

$$q(q+2l+1)\ c_q = 2\left[(q+l)\lambda_{k,l} - 1\right] c_{q-1} \tag{C-31}$$

If we fix c_0, this relation enables us to calculate c_1, then c_2, and thus by recurrence all the coefficients c_q. Since c_q/c_{q-1} approaches zero when $q \longrightarrow \infty$, the corresponding series is convergent for all ρ. Thus we have determined, for any value of $\lambda_{k,l}$, the solution of (C-24) that satisfies condition (C-25).

C-3-c. Energy quantization. Radial functions

We are now going to require the preceding solution to have a physically acceptable asymptotic behavior (*cf.* § C-3-b-α). This will involve quantization of the possible values of $\lambda_{k,l}$.

If the term in brackets on the right-hand side of (C-31) does not go to zero for any integer q, expansion (C-26) is a true infinite series, for which:

$$\frac{c_q}{c_{q-1}} \underset{q\to\infty}{\sim} \frac{2\lambda_{k,l}}{q} \tag{C-32}$$

Now, the power series expansion of the function $\mathrm{e}^{2\rho\lambda_{k,l}}$ is written:

$$\begin{cases} \mathrm{e}^{2\rho\lambda_{k,l}} = \displaystyle\sum_{q=0}^{\infty} d_q\, \rho^q \\ d_q = \dfrac{(2\lambda_{k,l})^q}{q!} \end{cases} \tag{C-33}$$

which implies:

$$\frac{d_q}{d_{q-1}} = \frac{2\lambda_{k,l}}{q} \tag{C-34}$$

If we compare (C-32) and (C-34), we see[10] that, for large values of ρ, the series being considered behaves like $\mathrm{e}^{2\rho\lambda_{k,l}}$. The corresponding function $u_{k,l}$ [formula (C-23)] is then proportional to $\mathrm{e}^{+\rho\lambda_{k,l}}$, which is not physically acceptable.

Consequently, we must reject all cases in which expansion (C-26) is an infinite series. The only possible values of $\lambda_{k,l}$ are those for which (C-26) has only a finite

[10] The reader can find a fuller discussion relating to an analogous problem in Complement $\mathrm{C_V}$.

number of terms, that is, those for which $y_{k,l}$ reduces to a polynomial. The corresponding function $u_{k,l}$ is then physically acceptable, since its asymptotic behavior is dominated by $\mathrm{e}^{-\rho\lambda_{k,l}}$. Therefore, all we need is an integer k such that the term in brackets of the right-hand side of (C-31) goes to zero for $q = k$: the corresponding coefficient c_k is then zero, as are all those of higher order, since the fact that c_k is zero means that c_{k+1} is as well, and so on. For fixed l, we label the corresponding values of $\lambda_{k,l}$ by this integer k (note that k is greater than or equal to 1, since c_0 never goes to zero). We then have, according to (C-31):

$$\lambda_{k,l} = \frac{1}{k+l} \tag{C-35}$$

For a given l, the only negative energies possible are therefore [formula (C-20)]:

$$E_{k,l} = \frac{-E_I}{(k+l)^2} \quad ; \qquad k = 1, 2, 3, \ldots \tag{C-36}$$

We shall discuss this result in § C-4.

$y_{k,l}$ is therefore a polynomial, whose term of lowest order is in ρ^{l+1} and whose term of highest order is in ρ^{k+l}. Its various coefficients can be calculated in terms of c_0 by solving recurrence relation (C-31), which can be written, using (C-35):

$$c_q = -\frac{2(k-q)}{q(q+2l+1)(k+l)} c_{q-1} \tag{C-37}$$

It is easy to show that:

$$c_q = (-1)^q \left(\frac{2}{k+l}\right)^q \frac{(k-1)!}{(k-q-1)!} \frac{(2l+1)!}{q!(q+2l+1)!} c_0 \tag{C-38}$$

$u_{k,l}(\rho)$ is then given by formula (C-23), and c_0 is determined (to within a phase factor) by normalization condition (A-34) [we must first, of course, return to the variable r by using (C-19)]. Finally, we obtain the true function $R_{k,l}(r)$ by dividing $u_{k,l}(r)$ by r. The following three examples give an idea of the form of these radial functions:

$$R_{k=1,l=0}(r) = 2(a_0)^{-3/2}\, \mathrm{e}^{-r/a_0} \tag{C-39a}$$

$$R_{k=2,l=0}(r) = 2(2a_0)^{-3/2} \left(1 - \frac{r}{2a_0}\right) \mathrm{e}^{-r/2a_0} \tag{C-39b}$$

$$R_{k=1,l=1}(r) = (2a_0)^{-3/2} \frac{1}{\sqrt{3}} \frac{r}{a_0}\, \mathrm{e}^{-r/2a_0} \tag{C-39c}$$

C-4. Discussion of the results

C-4-a. Order of magnitude of atomic parameters

Formulas (C-36) and (C-39) show that, for the hydrogen atom, the ionization energy E_I, defined by (C-12a), and the Bohr radius, given by (C-12b), play an important role. These quantities give an order of magnitude of the energies and spatial extensions of the wave functions associated with the bound states of the hydrogen atom.

Relations (C-12a) and (C-12b) can be written in the form:

$$E_I = \frac{1}{2}\alpha^2 \mu c^2 \tag{C-40a}$$

$$a_0 = \frac{1}{\alpha}\lambda\!\!\!^{-}_c \tag{C-40b}$$

where α is the *fine structure constant*, a dimensionless constant which plays a very important role in physics:

$$\alpha = \frac{e^2}{\hbar c} = \frac{q^2}{4\pi\varepsilon_0 \hbar c} \simeq \frac{1}{137} \tag{C-41}$$

and where $\lambda\!\!\!^{-}_c$ is defined by:

$$\lambda\!\!\!^{-}_c = \frac{\hbar}{\mu c} \tag{C-42}$$

Since μ is almost the same as m_e, the rest mass of the electron, $\lambda\!\!\!^{-}_c$ is practically equal to the *Compton wavelength* of the electron, which is given by:

$$\frac{\hbar}{m_e c} \simeq 3.8 \times 10^{-3}\ \text{Å} \tag{C-43}$$

Relation (C-40b) therefore indicates that a_0 is on the order of one hundred times the Compton wavelength of the electron. Relation (C-40a) shows that the order of magnitude of the binding energy of the electron is between $10^{-4}\,\mu c^2$ and $10^{-5}\,\mu c^2$, where μc^2 is practically equal to the rest energy of the electron:

$$m_e c^2 \simeq 0.51 \times 10^6\ \text{eV} \tag{C-44}$$

It follows that:

$$E_I \ll m_e c^2 \tag{C-45}$$

This justifies our choice of the non-relativistic Schrödinger equation to describe the hydrogen atom. Of course, relativistic effects, although small, do exist; nevertheless, their smallness allows them to be studied by perturbation theory (*cf.* Chap. XI and XII).

C-4-b. Energy levels

α. Possible values of the quantum numbers; degeneracies

For fixed l, there exists an infinite number of possible energy values [formula (C-36)], corresponding to $k = 1, 2, 3,$ ... Each of them is at least $(2l + 1)$-fold degenerate: this is an *essential degeneracy* related to the fact that the radial equation depends only on the quantum number l and not on m (§ A-3). But, in addition, there exist *accidental degeneracies*: equation (C-36) indicates that two eigenvalues $E_{k,l}$ and $E_{k',l'}$ corresponding to different radial equations ($l' \neq l$) are equal if $k + l = k' + l'$. Figure 4, in which the first eigenvalues associated with $l = 0, 1, 2$ and 3 are shown on a common energy scale, clearly reveals several accidental degeneracies.

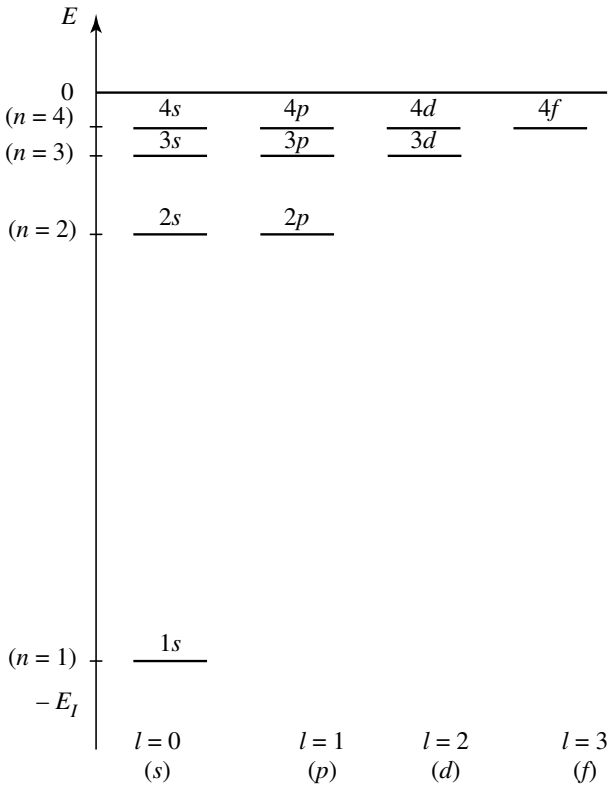


Figure 4: Energy levels of the hydrogen atom. The energy E_n of each level depends only on n. If n is fixed, several values of l are possible: $l = 0, 1, 2, \ldots n-1$. To each of these values of l correspond $(2l+1)$ possible values for m:

$$m = -l, -l+1, \ldots, l.$$

Consequently, the level E_n is n^2-fold degenerate.

In the special case of the hydrogen atom, $E_{k,l}$ does not depend on k and l separately, but only on their sum. We set:

$$n = k + l \tag{C-46}$$

The various energy states are labeled by the integer n (greater than or equal to 1), and (C-36) becomes:

$$E_n = -\frac{1}{n^2} E_I \tag{C-47}$$

According to (C-46), it is equivalent to specify k and l or n and l to determine the eigenfunctions. Following convention, from now on we shall use the quantum numbers n

and l. The energy is fixed by n, which is called the *principal quantum number*; a given value of n characterizes what is called an *electron shell*.

Since k is necessarily an integer which is greater than or equal to 1 (§ 3-c above), there is only a finite number of values of l associated with the same value of n. According to (C-46), if n is fixed, one can have:

$$l = 0, 1, 2, \ldots, n-1 \tag{C-48}$$

The shell characterized by n is said to contain n *sub-shells*[11], each one corresponding to one of the values of l given in (C-48). Finally, each sub-shell contains $(2l+1)$ distinct states, associated with the $(2l+1)$ possible values of m for fixed l.

The total degeneracy of the energy level E_n is therefore:

$$g_n = \sum_{l=0}^{n-1}(2l+1) = 2\frac{(n-1)n}{2} + n = n^2 \tag{C-49}$$

We shall see in Chapter IX that the existence of electron spin multiplies this number by 2 (if we also take into account the proton spin, which is equal to that of the electron, we obtain another factor of 2).

β. Spectroscopic notation

For historical reasons (dating from the period, before the development of quantum mechanics, in which the study of spectra resulted in an empirical classification of the numerous lines observed), letters of the alphabet are associated with the various values of l. The correspondence is as follows:

$$\begin{array}{l} l = 0 \leftrightarrow s \\ l = 1 \leftrightarrow p \\ l = 2 \leftrightarrow d \\ l = 3 \leftrightarrow f \\ l = 4 \leftrightarrow g \\ \quad\vdots \qquad\quad \vdots \end{array}$$

alphabetical order (C-50)

Therefore, *spectroscopic notation* labels a sub-shell by the corresponding number n followed by the letter that characterizes the value of l. Thus, the ground level [which is non-degenerate, according to (C-49)], sometimes called the "K shell", includes only the $1s$ sub-shell; the first excited level, or "L shell", includes the $2s$ and $2p$ sub-shells; the second excited level ("M shell") includes the $3s$, $3p$ and $3d$ sub-shells, etc. (The capital letters sometimes associated with the successive shells follow an alphabetical order, starting with the letter K.)

[11] The concept of a sub-shell can even be found in the semi-classical model of Sommerfeld. This model assigns, to each value n of Bohr's quantum number, n elliptical orbits of the same energy and different angular momenta. One of these orbits is circular; it is the one that corresponds to the maximum value of the angular momentum.

C-4-c. Wave functions

The wave functions associated with the eigenstates common to $\mathbf{L}^2$, L_z and the Hamiltonian H of the hydrogen atom are generally labeled, not by the three quantum numbers k, l, m, as we have done until now, but by n, l and m [passage from one set to the other simply involves use of relation (C-46)]. Since the operators H, $\mathbf{L}^2$ and L_z constitute a C.S.C.O. (*cf.* § A-3), specification of the three integers n, l and m, which is equivalent to that of the eigenvalues of H, $\mathbf{L}^2$ and L_z, unambiguously determines the corresponding eigenfunction $\varphi_{n,l,m}(\mathbf{r})$.

α. *Angular dependence*

As is the case for any central potential, the functions $\varphi_{n,l,m}(\mathbf{r})$ are products of a radial function and a spherical harmonic $Y_l^m(\theta, \varphi)$. To visualize their angular dependence on the axis characterized by the polar angles θ and φ, we can measure off a distance that is proportional to $|\varphi_{n,l,m}(r, \theta, \varphi)|^2$ for any *fixed* r, that is, proportional to $|Y_l^m(\theta, \varphi)|^2$. Thus, we obtain a surface of revolution about the Oz axis, since we know that $Y_l^m(\theta, \varphi)$ depends on φ only through the factor $e^{im\varphi}$ (§ D-1-b of Chapter VI); consequently, $|Y_l^m(\theta, \varphi)|^2$ is independent of φ. We can therefore represent its cross-section by a plane containing Oz. This is what is done in Figure 5, for $m = 0$ and $l = 0$, 1 and 2 [the corresponding spherical harmonics are given in Complement A_{VI}, formulas (31), (32) and (33)]: Y_0^0 is a constant, and is therefore spherically symmetric; $|Y_1^0|^2$ is proportional to $\cos^2\theta$, and $|Y_2^0|^2$ to $(3\cos^2\theta - 1)^2$.

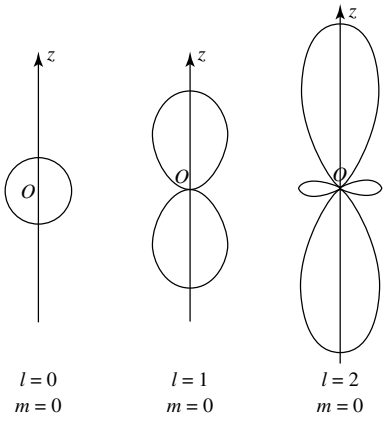


Figure 5: Angular dependence, $Y_l^m(\theta, \varphi)$, of some stationary wave functions of the hydrogen atom, corresponding to well-defined values of l and m. For each direction of polar angles θ, φ, the value of $|Y_l^m(\theta, \varphi)|^2$ is recorded; a surface of revolution about the Oz axis is thus obtained. When $l = 0$, this surface is a sphere centered at O; it becomes more complicated for higher values of l.

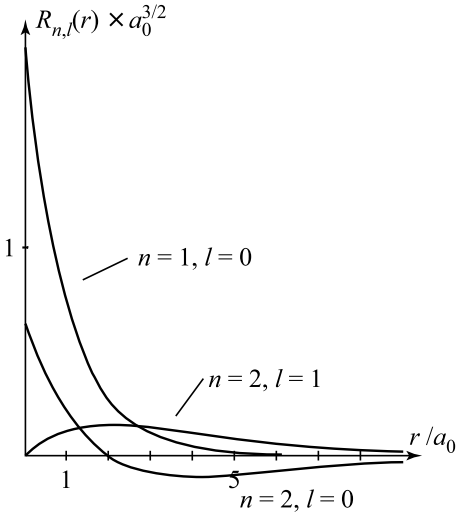


Figure 6: Radial dependence $R_{n,l}(r)$ of wave functions associated with the first few levels of the hydrogen atom. When $r \longrightarrow 0$, $R_{n,l}(r)$ behaves like r^l; only the s states (for which $l = 0$) have a non-zero position probability at the origin.

β. Radial dependence

The radial functions $R_{n,l}(r)$, each of which characterizes a sub-shell, can be calculated from the results of § C-3-c, paying attention, however, to the change in notation introduced by formula (C-46). Figure 6 represents the variation with respect to r of the three radial functions given in (C-39):

$$R_{k=1,\,l=0} \equiv R_{n=1,\,l=0} \; ; \quad R_{k=2,\,l=0} \equiv R_{n=2,\,l=0} \; ; \quad R_{k=1,\,l=1} \equiv R_{n=2,\,l=1} \tag{C-51}$$

The behavior of $R_{n,l}(r)$ in the neighborhood of $r = 0$ is that of r^l (see discussion of § A-2-c). Consequently, *only states belonging to s sub-shells* ($l = 0$) *give a non-zero probability of presence at the origin.* The greater l, the larger the region around the proton in which the position probability of the electron is negligible. This has a certain number of physical consequences, particularly in the phenomenon of electron capture by certain nuclei, and in the hyperfine structure of lines (*cf.* Chap. XII, § B-2).

Finally, we can derive formula (C-11b) for the successive Bohr radii. To do so, consider the various states for which $l = n - 1$[12]. We calculate the variation with r of the probability density for each of the preceding levels in an infinitesimal solid angle $\mathrm{d}\Omega$ about a fixed direction of polar angles θ and φ. In general, the position probability for the electron in the volume element $\mathrm{d}^3r = r^2\,\mathrm{d}r\,\mathrm{d}\Omega$ situated at the point (r, θ, φ) is given by:

$$\begin{aligned} \mathrm{d}^3\mathcal{P}_{n,l,m}(r,\theta,\varphi) &= |\varphi_{n,l,m}(r,\theta,\varphi)|^2\, r^2\,\mathrm{d}r\,\mathrm{d}\Omega \\ &= |R_{n,l}(r)|^2\, r^2\,\mathrm{d}r \times |Y_l^m(\theta,\varphi)|^2\,\mathrm{d}\Omega \end{aligned} \tag{C-52}$$

Here, we have fixed θ, φ and $\mathrm{d}\Omega$. The probability of finding the electron between r and $r + \mathrm{d}r$, inside the solid angle under consideration, is then proportional to $r^2\,|R_{n,l}(r)|^2\,\mathrm{d}r$. The corresponding density is therefore, to within a constant factor, $r^2\,|R_{n,l}(r)|^2$ (the

[12] These states correspond to the circular orbits of the Sommerfeld theory (*cf.* note 11).

factor r^2 arises from the expression for the volume element in spherical coordinates). We are interested in the case where $l = n - 1$, that is, $k = n - l = 1$; § C-3-c then indicates that the polynomial which enters into $R_{n,l}(r)$ contains only one term in $(r/a_0)^{n-1}$. The desired probability density is therefore proportional to:

$$f_n(r) = \frac{r^2}{a_0^2}\left[\left(\frac{r}{a_0}\right)^{n-1} \mathrm{e}^{-r/na_0}\right]^2 = \left(\frac{r}{a_0}\right)^{2n} \mathrm{e}^{-2r/na_0} \tag{C-53}$$

This function has a maximum for:

$$r = r_n = n^2 a_0 \tag{C-54}$$

which is the radius of the Bohr orbit corresponding to the energy E_n.

Finally, the table below gives the expressions for the wave functions of the first energy levels:

$1s$ level	$\varphi_{n=1,\,l=0,\,m=0} = \dfrac{1}{\sqrt{\pi a_0^3}}\,\mathrm{e}^{-r/a_0}$
$2s$ level	$\varphi_{n=2,\,l=0,\,m=0} = \dfrac{1}{\sqrt{8\pi a_0^3}}\left(1 - \dfrac{r}{2a_0}\right)\mathrm{e}^{-r/2a_0}$
$2p$ level	$\varphi_{n=2,\,l=1,\,m=1} = -\dfrac{1}{8\sqrt{\pi a_0^3}}\,\dfrac{r}{a_0}\,\mathrm{e}^{-r/2a_0}\sin\theta\,\mathrm{e}^{i\varphi}$
	$\varphi_{n=2,\,l=1,\,m=0} = \dfrac{1}{4\sqrt{2\pi a_0^3}}\,\dfrac{r}{a_0}\,\mathrm{e}^{-r/2a_0}\cos\theta$
	$\varphi_{n=2,\,l=1,\,m=-1} = \dfrac{1}{8\sqrt{\pi a_0^3}}\,\dfrac{r}{a_0}\,\mathrm{e}^{-r/2a_0}\sin\theta\,\mathrm{e}^{-i\varphi}$

References and suggestions for further reading:

Particle in a central potential: Messiah (1.17), Chap. IX; Schiff (1.18), § 16.

The Bohr-Sommerfeld atom and the old quantum therory: Cagnac and Pebay-Peyroula (11.2), Chaps. V, VI and XIII; Born (11.4), Chap. V, §§ 1 and 2; Pauling and Wilson (1.9), Chap. II; Tomonaga (1.8), Vol. I; Eisberg and Resnick (1.3), Chap. 4.

Hydrogen-like wave functions: Levine (12.3), § 6.5; Karplus and Porter (12.1), §§ 3.8 and 3.10; Eisberg and Resnick (1.3), §§ 7.6 and 7.7.

Degeneracy related to a $1/r$ potential (dynamical group): Borowitz (1.7), § 13.7; Schiff (1.18), § 30; Bacry (10.31), § 6.11.

Mathematical treatment of differential equations: Morse and Feshbach (10.13), Chaps. 5 and 6; Courant and Hilbert (10.11), Vol. I, § V-11.

A_{VII} : HYDROGEN-LIKE SYSTEMS	Presentation of various hydrogen-like systems to which the calculations of Chapter VII can be applied directly. The accent is placed on physical discussions and on the influence of the masses of the particles involved in the system. Simple, advised for a first reading.
B_{VII} : A SOLUBLE EXAMPLE OF A CENTRAL POTENTIAL: THE ISOTROPIC THREE-DIMENSIONAL HARMONIC OSCILLATOR	Study of another case (a three-dimensional harmonic oscillator) in which it is possible to calculate exactly the energy levels of a particle in a central potential by the method of Chapter VII (solution of the radial equation). Not theoretically difficult; may be considered as a worked example.
C_{VII} : PROBABILITY CURRENTS ASSOCIATED WITH THE STATIONARY STATES OF THE HYDROGEN ATOM	Completes the results of § C-4-c of Chapter VII concerning the properties of the stationary states of the hydrogen atom by calculating their probability currents. Short and simple, useful for Complement D_{VII}.
D_{VII} : THE HYDROGEN ATOM PLACED IN A UNIFORM MAGNETIC FIELD. PARAMAGNETISM AND DIAMAGNETISM. THE ZEEMAN EFFECT	Discussion of a certain number of physical phenomena, using the results of Chapter VII. D_{VII} : the properties of an atom in a magnetic field (diamagnetism, paramagnetism, Zeeman effect). Moderately difficult, important because of the numerous applications.
E_{VII} : STUDY OF SOME ATOMIC ORBITALS. HYBRID ORBITALS	E_{VII}: complement intended to introduce the concept of an atomic hybrid orbital, essential for understanding certain properties of the chemical bond. No theoretical difficulty. Stresses the geometrical aspect of wave functions.
F_{VII} : VIBRATIONAL-ROTATIONAL LEVELS OF DIATOMIC MOLECULES	F_{VII}: direct application of the theory of Chapter VII to the study of the vibrational-rotational spectrum of heteropolar diatomic molecules. Sequel to Complements A_V (§ 1) and C_{VI}; moderately difficult.
G_{VII} : EXERCISES	Exercise 2 studies the influence of a uniform magnetic field on the levels of a simple physical system, in an exactly soluble case. It then provides a concrete illustration of the general considerations of Complements C_{VII} and D_{VII} concerning the influence of the paramagnetic and diamagnetic terms of the Hamiltonian.

Complement A_{VII}

Hydrogen-like systems

1 Hydrogen-like systems with one electron 834
1-a Electrically neutral systems 834
1-b Hydrogen-like ions 838
2 Hydrogen-like systems without an electron 839
2-a Muonic atoms 839
2-b Hadronic atoms 840

The calculations of Chapter VII, which enabled us to determine numerous physical properties of the hydrogen atom (energy levels, spatial distribution of the wave functions, etc.) are based on the fact that the system under study is composed of two particles (an electron and a proton) whose mutual attraction energy is inversely proportional to the distance between them. There exist numerous other systems in physics that fulfil this condition: deuterium or tritium, muonium, positronium, muonic atoms, etc. The results obtained in Chapter VII are therefore directly applicable to these examples. All we need to do is change the constants introduced in the calculations (masses and charges of the two particles). This is what we shall do in this complement, in which we shall study, in particular, how the Bohr radius and ionization energy E_I are modified in each of the systems to be considered. The wave functions associated with their stationary states and the corresponding energies will then be obtained by replacing, in formulas (C-39) and (C-47) of Chapter VII, a_0 and E_I by their new values, which give the order of magnitude of the spatial extension of the wave functions and the binding energies of these systems.

We recall the expressions for a_0 and E_I:

$$a_0 = \lambda\!\!\!^{-}_c \frac{1}{\alpha} = \frac{\hbar^2}{\mu e^2} \tag{1}$$

$$E_I = \frac{1}{2}\mu c^2\, \alpha^2 = \frac{\mu e^4}{2\hbar^2} \tag{2}$$

where μ is the reduced mass of the electron-proton system:

$$\mu = \mu(\mathrm{H}) = \frac{m_e m_p}{m_e + m_p} \simeq m_e \left(1 - \frac{m_e}{m_p}\right) \tag{3}$$

and e^2 characterizes the intensity of the attractive potential $V(r)$:

$$V(r) = -\frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{4}$$

In the case of hydrogen, we have seen that:

$$a_0(\mathrm{H}) \simeq 0.52 \text{ Å} \tag{5a}$$

$$E_I \simeq 13.6 \text{ eV} \simeq 2.2 \times 10^{-18} \text{ J} \tag{5b}$$

How can we obtain the corresponding values for a system of two arbitrary particles, of masses m_1 and m_2, whose attraction energy is:

$$V'(r) = -\frac{Ze^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{6}$$

(where Z is a dimensionless parameter)? All we must do is calculate the reduced mass μ of the system by replacing m_e and m_p by m_1 and m_2 in (3):

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{7a}$$

and substitute the result obtained into (1) and (2), being careful to perform the substitution:

$$e^2 \Longrightarrow Z\, e^2 \tag{7b}$$

This is what we shall do in a certain number of physical examples.

1. Hydrogen-like systems with one electron

1-a. Electrically neutral systems

α. Heavy isotopes of hydrogen

The physical systems closest to the hydrogen atom are its two isotopes, deuterium and tritium. In these atoms, the proton is replaced by a nucleus having the same charge but possessing either one or two neutrons in addition to the proton. The mass of the deuterium nucleus is approximately $2m_p$, and that of the tritium nucleus, $3m_p$. The reduced masses therefore become, in these two cases:

$$\mu(\text{deuterium}) \simeq m_e \left(1 - \frac{m_e}{2m_p}\right) \tag{8a}$$

$$\mu(\text{tritium}) \simeq m_e \left(1 - \frac{m_e}{3m_p}\right) \tag{8b}$$

Since:

$$\frac{m_e}{m_p} \simeq \frac{1}{1\,836} \ll 1 \tag{9}$$

it is clear that the reduced masses of deuterium and tritium are very close to that of hydrogen, and that they can be replaced by m_e without great inaccuracy.

If we substitute either (3), or (8a), or (8b) into formulas (1) and (2), we see that the Bohr radii and energies of the hydrogen, deuterium and tritium atoms are practically the same. Nevertheless, there are slight differences, of the order of a thousandth in relative value. These differences can be detected experimentally. For example, with an optical spectrograph of sufficient resolution, it can be observed that the wavelengths of the lines emitted by hydrogen atoms are slightly greater than those emitted by deuterium, which are in turn greater than those emitted by tritium. This slight shift in the emitted wavelengths is related to the fact that the nucleus is not infinitely heavy, and does not remain fixed while the electron moves; this is called the "nuclear finite mass effect". Experiments have verified that formulas (7a), (1) and (2) account very precisely for this effect.

β. Muonium

The muon is a particle whose fundamental properties are the same as those of the electron, except for a difference in mass (the mass m_μ of the muon is approximately equal to 207 m_e). In particular, the muon is not sensitive to nuclear forces (strong interactions). There are two types of muons, the μ^- and the μ^+, whose charges are respectively equal to those of the electron e^- and the positron[1] e^+. Like all other charged particles, the muon is sensitive to electromagnetic interactions.

We can therefore consider a physical system formed by a μ^+ muon and an electron e^-, in which the electrostatic attraction is the same as for a proton and an electron. Bound states therefore exist. This is, so to speak, a light isotope of hydrogen, in which the μ^+ muon replaces the proton. This "isotope" is called muonium (its atomic mass is on the order of $m_\mu/m_p \simeq 0.1$).

It is not difficult to use the results in Chapter VII to find the ionization energy and Bohr radius associated with muonium; (1), (2) and (7) yield:

$$a_0(\text{muonium}) = a_0(\text{H})\frac{1+m_e/m_\mu}{1+m_e/m_p} \simeq a_0(\text{H})\left(1+\frac{1}{200}\right) \tag{10a}$$

$$E_I(\text{muonium}) = E_I(\text{H})\frac{1+m_e/m_p}{1+m_e/m_\mu} \simeq E_I(\text{H})\left(1-\frac{1}{200}\right) \tag{10b}$$

Since the muon is approximately ten times lighter than the proton, the nuclear finite mass effect is about ten times greater for muonium than for hydrogen; however, since the electron is distinctly lighter than the muon, this effect remains small (on the order of 0.5%). For example, the wavelengths of the optical lines emitted by muonium are be close to those of the corresponding lines for hydrogen. The frequency of the resonance line has been measured with a very high accuracy in 1999 [*cf.* ref. (11.26)], and has shown that the charges of the electron and of the muon are the same within 10^{-9}.

Experimentally, the existence of muonium was revealed by its instability: the μ^+ muon decays, emitting a positron and two neutrinos, and the lifetime of muonium is 2.2×10^{-6} s. The positron resulting from this decay can be detected. It is emitted preferentially in the direction of the μ^+ muon spin[2] (non-conservation of parity in weak interactions). Detection of the positrons then leads to the experimental determination of this direction. Since, in addition, the spin of the μ^+ muon of a muonium atom is coupled to that of the electron (hyperfine structure coupling; *cf.* Chap. XII and complements), its precession frequency in a magnetic field is different from that of a free muon. Measurement of this frequency thus reveals the existence of muonium atoms.

The study of muonium is of very great interest, theoretically as well as experimentally. The two particles which constitute this system are not subjected to strong interactions, so that its energy levels (in particular, the hyperfine structure of the ground state $1s$) can be calculated with great precision, without bringing in any "nuclear correction" (for the hydrogen atom, on the other hand, one must take into account the internal structure and polarizability of the proton, which are due to strong interactions). Comparison between theoretical predictions and experimental results provides a very severe test of quantum electrodynamics. For instance, a measurement of the hyperfine

[1] In addition, like e^- and e^+, μ^- and μ^+ are antiparticles of each other.

[2] Like the electron, the muon has a spin of 1/2 with which is associated a magnetic moment $\mathbf{M}_\mu = \frac{q_\mu}{m_\mu}\mathbf{S}$.

structure of muonium can give an excellent determinations of the fine structure constant $\alpha = e^2/\hbar c$.

γ. Positronium

Positronium is a bound system composed of an electron e^- and a positron e^+. Like muonium, it can be said by extension to be an isotope of hydrogen, with the proton being replaced by a positron. However, it must be noted that the situation is not quite the same: in the hydrogen atom, the proton (which is much heavier than the electron) remains almost motionless, while in positronium, the positron, the antiparticle of the electron, has the same mass and consequently the same velocity as the electron when the center of mass of positronium is fixed (*cf.* Fig. 1-b).

According to (7a), the reduced mass associated with positronium is:

$$\mu(\text{positronium}) = \frac{m_e}{2} \tag{11}$$

Therefore:

$$a_0(\text{positronium}) \simeq 2\, a_0(\text{H}) \tag{12a}$$

$$E_I(\text{positronium}) \simeq \frac{1}{2}\, E_I(\text{H}) \tag{12b}$$

Thus, for a given state of positronium, the average electron-positron distance is twice the electron-proton distance for the corresponding state of the hydrogen atom (*cf.* Fig. 1). The differences between the energies of the stationary states, however, are twice as small, and the optical line spectrum emitted by positronium is obtained by doubling all the wavelengths of that of hydrogen.

Comment:

One should not conclude from formula (12a) that the radius of positronium is twice that of the hydrogen atom. The Bohr radius gives an idea of the extension

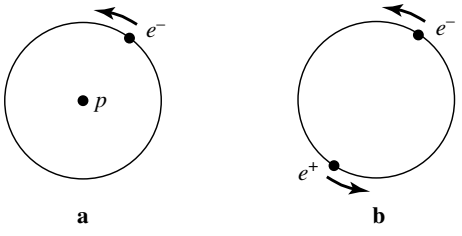


Figure 1: Schematic representation of the hydrogen (electron + proton system) and positronium (electron + positron system) atoms. Since the proton is much heavier than the electron, it is located practically at the center of mass of the hydrogen atom; the electron "revolves" about the proton at a distance $a_0(H)$. On the other hand, the positron is equal in mass to the electron; these particles therefore both revolve about their center of mass, the distance between them being a_0 (positronium) $= 2a_0(H)$.

of the wave functions associated with the "relative particle" (*cf.* Chap. VII, § B), whose position $\mathbf{r}_1 - \mathbf{r}_2$ is related to the distance between the two particles and not to the distance between them and the center of mass G. Figure 1 clearly shows, moreover, that the hydrogen and positronium atoms are of equal size. In general, all hydrogen-like systems for which the attractive potential is given by (6) with $Z = 1$ have exactly the same radius, since formula (B-5) of Chapter VII shows that:

$$\mathbf{r}_1 - \mathbf{r}_G = \frac{m_2}{m_1 + m_2}\mathbf{r} = \frac{\mu}{m_1}\mathbf{r} \tag{13}$$

Using (1), which gives the order of magnitude of the spatial extension of the wave function $\varphi_{100}(\mathbf{r})$ of the ground state, we see that the "radius" ρ of the atom can be defined by:

$$\rho = \frac{\hbar^2}{m_1 Z e^2} \tag{14}$$

where m_1 is the mass of the lighter particle (the heavier particle is found closer to the center of mass). In all the systems considered until now, $Z = 1$ and $m_1 = m_e$; their radii are therefore the same. We shall see cases later in which the radius ρ is smaller, either because $m_1 \neq m_e$ or because $Z \neq 1$.

The optical spectrum of positronium has been observed in 1975 [*cf.* ref. (11-23)]. As for the (hyperfine) structure of the ground state (due to the interaction between the magnetic moments of the electron and the positron), it has been accurately determined (*cf.* Complement C_{XII}).

The fact that positronium, like muonium, is a purely electrodynamic system (neither the electron nor the positron is sensitive to strong interactions) explains the importance attached to its theoretical and experimental study.

Let us also point out that positronium is an unstable system. Since the ground state is a $1s$ state, the electron and positron come into contact and annihilate, yielding two or three photons, depending on the hyperfine structure level they started from. The study of the corresponding decay rate is also of great interest in quantum electrodynamics.

δ. Hydrogen-like systems in solid state physics

Atomic physics is not the only domain of application of the theory presented in Chapter VII. For example, the "donor atoms" localized in semiconductors constitute approximately hydrogen-like systems in solid state physics.

Consider a silicon crystal; in the silicon lattice, each atom uses its four valence electrons to form four tetrahedral bonds with its neighbors. If a pentavalent atom like phosphorus (a donor atom) is introduced into the lattice in place of a silicon atom, it must lose a valence electron, and its overall charge becomes positive. It then behaves like a center which can retain the electron and form a hydrogen-like system with it. Actually, the force acting on the electron cannot be calculated directly from Coulomb's law in a vacuum since silicon has a large dielectric constant $\varepsilon \simeq 12$, so that (4) must be replaced by:

$$V(r) = -\frac{e^2}{\varepsilon|\mathbf{r}_1 - \mathbf{r}_2|} \tag{15}$$

To be completely rigorous, we should have to replace the electron mass by the "effective mass" m^* of the electron in silicon, which is different from the free electron mass because of interactions with the charges of the nuclei in the crystal. Nevertheless, we shall confine ourselves to a qualitative discussion, noting that the effect of the high value of ε is to decrease e^2 in (15), that is, according to (1), to increase the Bohr radius by a factor on the order of 10. The donor atom impurity is therefore similar to a very large hydrogen atom, whose wave functions are spread over distances much greater than the unit cell length of the silicon lattice.

Let us briefly describe another hydrogen-like system in solid state physics: the exciton. Consider a semi-conductor crystal. In the absence of external perturbations, the outside electrons of the atoms forming the crystal are all in states belonging to the "valence band" (the temperature is assumed to be sufficiently low; *cf.* Complement C_{XIV}). By suitably illuminating the crystal, we can, through the absorption of a photon, cause an electron to go into the "conduction band" (which contains a set of energy levels which are higher than those of the valence band). There is then one electron missing from the valence band. We can treat this band as if it contained a particle of charge opposite to that of an electron, called a "hole". The hole can attract an electron of the valence band and form a bound system with it: the exciton. The exciton, like hydrogen, has a series of energy levels between which it can undergo transitions. Its existence can be demonstrated by a measurement of light absorption by the crystal.

1-b. Hydrogen-like ions

The neutral helium atom is composed of two electrons and a positively charged nucleus of charge $-2q_e$. Such a system, which consists of three particles, cannot be studied with the theory of Chapter VII. However, if an electron is somehow removed from the helium atom, the He^+ ion is similar to a hydrogen atom; the only differences are in the nuclear charge, which is twice that of the proton (the total charge of the ion is positive and equal to $-q_e$) and its mass (which, for ^{4}He, is approximately four times that of the proton). Of course, there are other hydrogen-like ions: the Li^{++} ion (the lithium atom, when it is not ionized, has $Z = 3$ electrons), the Be^{+++} ion ($Z = 4$), etc...

Let us then consider a system formed by a nucleus, of mass M and positive charge $-Zq_e$, and an electron. If we substitute (7b) in (1) and (2), we obtain:

$$a_0(Z) \simeq \frac{a_0(H)}{Z} \tag{16}$$

$$E_I(Z) \simeq Z^2\, E_I(H) \tag{17}$$

(since $M \gg m_e$, we have neglected the difference between the reduced mass of hydrogen and that of the hydrogen-like ion under study; the consequences of the nuclear finite mass effect on a_0 and E_I are, in effect, negligible compared to those due to the charge variation). Hydrogen-like ions are therefore all smaller than the hydrogen atom, as one would expect since the nucleus and the electron are more strongly bound. Moreover, their energy increases rapidly with Z (quadratically): for example, the energy which must be supplied to a Li^{++} ion to remove its last electron is greater than 100 eV. This is why the electromagnetic frequencies that can be emitted or absorbed by a hydrogen-like ion fall into the ultraviolet, and even, when Z is large enough, into the domain of X-rays.

2. Hydrogen-like systems without an electron

Thus far, the systems we have considered all include an electron. Nevertheless, there exist numerous other particles having the same charge q_e, able to form a hydrogen-like system with a nucleus of charge $-Zq_e$. We shall give a few examples. The "atoms" we are going to describe here are of course less common than the "usual" atoms which appear in Mendeleev's classification. They are unstable, and, in order to study them, it is necessary to use high-energy particle accelerators to produce the particles needed for their formation. This is why they are sometimes called "exotic atoms".

2-a. Muonic atoms

We have already mentioned some essential features of the muon and pointed out the existence of the μ^- muon. When this particle is attracted by a positively charged atomic nucleus, it can form a bound system called a "muonic atom"[3].

Consider, for example, the simplest muonic atom, which is composed of a μ^- muon and a proton; this is a neutral system whose Bohr radius is:

$$a_0(\mu^-, p^+) \simeq \frac{\hbar^2}{m_\mu e^2} \simeq \frac{a_0(H)}{200} \tag{18}$$

and whose ionization energy is:

$$E_I(\mu^-, p^+) \simeq \frac{m_\mu e^4}{2\hbar^2} \simeq 200\, E_I(H) \tag{19}$$

The size of this muonic atom is therefore on the order of several thousandths of an Angström. Its spectrum is obtained from that of hydrogen by dividing the wavelengths by 200; it therefore falls into the domain of soft X-rays.

What happens if, instead of revolving about a proton, the μ^- is captured by a nucleus N whose charge is Z times greater, like lead[4] for example, for which $Z = 82$? Formulas (1) and (2) then yield:

$$a_0(\mu^-, N) \simeq \frac{a_0(H)}{200\, Z} \tag{20}$$

$$E_I(\mu^-, N) \simeq 200\, Z^2\, E_I(H) \tag{21}$$

Setting $Z = 82$ in these formulas, we find for the transitions of the lead muonic atom energies equal to several MeV (1 MeV = 10^6 electron volts). However, it must be noted that formulas (1) and (2) are no longer valid in this case, since equation (20) would yield:

$$a_0(\mu^-, \mathrm{Pb}) \simeq 3 \times 10^{-5}\ \text{Å} = 3\ \text{Fermi} \tag{22}$$

[3] We could also imagine a bound system composed of a μ^+ muon and a μ^- muon. However, given the low intensity of muon beams that can be produced, such an atom is very difficult to create.

[4] Such a system can be formed by directing a μ^- beam onto a target of lead atoms. When a μ^- is captured by a lead nucleus, it revolves about it at a distance about 200 times smaller than the distance to the electrons of the innermost shell of the atom. The nuclear charge is therefore practically the only one to affect the muon. Thus, in studying the states of the muonic atom, we can simply ignore the electrons.

that is, a radius slightly smaller than the radius of the lead nucleus. The calculations of Chapter VII are therefore no longer valid. This is because they are based on form (6) of the potential $V(r)$, which is correct[5] only when the particles under study, separated by distances much greater than their dimensions, can be considered to be point particles. This hypothesis, very well satisfied for hydrogen, is not valid in the case studied here.

However, (20) and (21) give the correct order of magnitude of the energies and radius of the lead muonic atom. The physical consequences of the existence of a non-zero spatial extension of the nucleus ("volume effect") will be studied in greater detail in Complement D_{XI}. We take this occasion to note, however, that one of the reasons for interest in muonic atoms is precisely related to this type of effect: the μ^- muon "explores", as it were, the internal structure of the nucleus[6], and the energy levels of muonic atoms depend on the electrical charge distribution and magnetism inside the nucleus (recall that the muon is not sensitive to nuclear forces). Thus, the study of these states can furnish information very useful in nuclear physics.

2-b. Hadronic atoms

"Hadrons" are those particles sensitive to strong interactions, as opposed to "leptons", which are not. Electrons and muons, whose bound states in a Coulomb potential we have studied thus far, are leptons. Protons, neutrons and mesons such as the π meson, etc... are hadrons. Of the latter, those that are negatively charged can form a hydrogen-like bound system with an atomic nucleus, called a "hadronic atom". For example, the nucleus-π meson system yields a "pionic atom"; the nucleus-Σ^- particle system, a "sigmaonic atom"[7]; the nucleus-K^- meson system, a "kaonic atom"; the nucleus-antiproton system, an "antiprotonic atom", etc... All of the systems just cited have actually been observed and studied. They are all unstable, but their lifetimes are long enough for us to observe some of their spectral lines. Hydrogen atom theory, which takes into account only the electrostatic interaction between the two particles under consideration, does not, of course, apply to such systems, in which strong interactions play an important role. However, since they are of very short range, strong interactions can be neglected in the study of excited states of the hadronic atom (other than the s states), in which the two particles are far apart. The theory of Chapter VII is then applicable, as are formulas (1) and (2), which, in all these cases, lead to much smaller Bohr radii and much greater energies than for hydrogen. Thus, measurement of the spectral frequencies emitted by pionic atoms gives a very precise determination of the mass of the π^- meson.

References and suggestions for further reading:

Exotic atoms: see the subsection "Exotic atoms" of section 11 of the bibliography; see also Cagnac and Pebay-Peyroula (11.2), Chap. XIX, § 7; Weissenberg (16.19), Chap. 4, § 2 and Chap. 6.

Excitons: Kittel (13.2), Chap. 17, p. 538; Ziman (13.3), § 6.7.

[5] Inside the nucleus, the potential is approximately parabolic (*cf.* Complements A_V, § 4 and D_{XI}).

[6] The concept of the impenetrability of two solid bodies is macroscopic. In quantum mechanics, nothing prevents the wave functions of two particles of different nature from overlapping.

[7] The term "mesic atom" is sometimes used to denote a system involving a meson. Similarly, since the Σ^- is a hyperon (a particle heavier than the proton), the sigmaonic atom is sometimes called a "hyperonic atom".

Complement B_{VII}

A soluble example of a central potential: the isotropic three-dimensional harmonic oscillator

1 Solving the radial equation 842
2 Energy levels and stationary wave functions 845

In this complement we shall examine a special case of a central potential for which the radial equation is exactly soluble: the isotropic three-dimensional harmonic oscillator. We have already treated this problem (Complement E_V) by considering the state space $\mathcal{E}_\mathbf{r}$ as the tensor product $\mathcal{E}_x \otimes \mathcal{E}_y \otimes \mathcal{E}_z$; this amounts, in the $\{|r\rangle\}$ representation, to separating the variables in Cartesian coordinates. We thus obtained three differential equations, one in the x-variable, one in y, and the third in z. Here we intend to seek the stationary states that are also eigenstates of $\mathbf{L}^2$ and L_z, by separating the variables in polar coordinates. We shall then indicate how the two bases of $\mathcal{E}_\mathbf{r}$ obtained by these two different methods are related to each other.

We shall also study, in Complement A_{VIII}, the stationary states of well-defined angular momentum of a free particle. This can be considered to be another special case of a central potential $[V(r) \equiv 0]$ which leads to an exactly soluble radial equation.

A three-dimensional harmonic oscillator is composed of a (spinless) particle of mass μ subjected to the potential:

$$V(x,y,z) = \frac{1}{2}\mu\left[\omega_x^2 x^2 + \omega_y^2 y^2 + \omega_z^2 z^2\right] \tag{1}$$

where ω_x, ω_y and ω_z are real positive constants. The oscillator is said to be isotropic if:

$$\omega_x = \omega_y = \omega_z = \omega \tag{2}$$

Since the potential (1) is the sum of a function of x alone, a function of y alone and a function of z alone, we can solve the eigenvalue equation of the Hamiltonian:

$$H = \frac{\mathbf{P}^2}{2\mu} + V(\mathbf{R}) \tag{3}$$

by separating the variables x, y and z in the $\{|r\rangle\}$ representation. This is what was done in Complement E_V. The energy levels, for an isotropic oscillator, are then found to be of the form:

$$E_n = \left(n + \frac{3}{2}\right)\hbar\omega \tag{4}$$

where n is any positive integer or zero. The degree of degeneracy g_n of the level E_n is equal to:

$$g_n = \frac{1}{2}(n+1)(n+2) \tag{5}$$

and the associated eigenfunctions are:

$$\varphi_{n_x,n_y,n_z}(x,y,z) = \left(\frac{\beta^2}{\pi}\right)^{3/4} \frac{1}{\sqrt{2^{n_x+n_y+n_z} n_x! n_y! n_z!}} \, \mathrm{e}^{-\frac{\beta^2}{2}(x^2+y^2+z^2)} \times H_{n_x}(\beta x)\, H_{n_y}(\beta y)\, H_{n_z}(\beta z) \quad (6)$$

with:

$$\beta = \sqrt{\frac{\mu\omega}{\hbar}} \quad (7)$$

[$H_p(u)$ denotes the Hermite polynomial of degree p; *cf.* Complement B_V]. φ_{n_x,n_y,n_z} is an eigenfunction of the Hamiltonian H with the eigenvalue E_n such that:

$$n = n_x + n_y + n_z \quad (8)$$

If the oscillator under consideration is isotropic[1], the potential (1) is a function only of the distance r between the particle and the origin:

$$V(r) = \frac{1}{2}\mu\omega^2 r^2 \quad (9)$$

Consequently, the three components of the orbital angular momentum $\mathbf{L}$ are constants of the motion. We want to find the common eigenstates of H, $\mathbf{L}^2$ and L_z. To do so, we could proceed, as in Complement D_{VI}, by introducing operators related to right and left circular quanta and to "longitudinal" quanta corresponding to the third degree of freedom along Oz (an outline of this method is given at the end of this complement). However, we prefer to use this example to illustrate the method elaborated in Chapter VII (§ A) and solve the radial equation by the polynomial method.

1. Solving the radial equation

For a fixed value of the quantum number l, the radial functions $R_{k,l}(r)$ and energies $E_{k,l}$ are given by the equation: [*cf.* relation (A-21) of Chapter VII]:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r + \frac{1}{2}\mu\omega^2 r^2 + \frac{l(l+1)\hbar^2}{2\mu r^2}\right] R_{k,l}(r) = E_{k,l}\, R_{k,l}(r) \quad (10)$$

We set:

$$R_{k,l}(r) = \frac{1}{r}u_{k,l}(r) \quad (11a)$$

$$\varepsilon_{k,l} = \frac{2\mu E_{k,l}}{\hbar^2} \quad (11b)$$

Equation (10) then becomes:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \beta^4 r^2 - \frac{l(l+1)}{r^2} + \varepsilon_{k,l}\right] u_{k,l}(r) = 0 \quad (12)$$

[1] Separation of the polar variables r, θ, φ is possible only for an isotropic oscillator.

[where β is the constant defined in (7)]. We must add the condition at the origin:

$$u_{k,l}(0) = 0 \tag{13}$$

For large r, (12) virtually reduces to:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \beta^4 r^2\right] u_{k,l}(r) \underset{r\to\infty}{\simeq} 0 \tag{14}$$

The asymptotic behavior of the solutions of equation (12) is therefore dominated by $\mathrm{e}^{\beta^2 r^2/2}$ or $\mathrm{e}^{-\beta^2 r^2/2}$. Only the second possibility is physically acceptable. This leads us to the change of functions:

$$u_{k,l}(r) = \mathrm{e}^{-\beta^2 r^2/2} y_{k,l}(r) \tag{15}$$

It is easy to find that $y_{k,l}(r)$ must satisfy:

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2} y_{k,l} - 2\beta^2 r \frac{\mathrm{d}}{\mathrm{d}r} y_{k,l} + \left[\varepsilon_{k,l} - \beta^2 - \frac{l(l+1)}{r^2}\right] y_{k,l} = 0 \tag{16a}$$

$$y_{k,l}(0) = 0 \tag{16b}$$

Now we shall seek $y_{k,l}(r)$ in the form of a power series in r:

$$y_{k,l}(r) = r^s \sum_{q=0}^{\infty} a_q \, r^q \tag{17}$$

where, by definition, a_0 is the coefficient of the first non-zero term:

$$a_0 \neq 0 \tag{18}$$

When we substitute expansion (17) into equation (16a), the term of lowest order is in r^{s-2}; its coefficient is zero if:

$$[s(s-1) - l(l+1)]a_0 = 0 \tag{19}$$

With conditions (18) and (16b) taken into account, the only way to satisfy relation (19) is to choose:

$$s = l + 1 \tag{20}$$

(this result could have been predicted; *cf.* § A-2-c of Chapter VII). The next term in the expansion of equation (16a) is in r^{s-1}, and its coefficient is equal to:

$$[s(s+1) - l(l+1)]a_1 \tag{21}$$

Since s is already fixed by (20), this coefficient can go to zero only if:

$$a_1 = 0 \tag{22}$$

Finally, let us set the coefficient of the general term in r^{q+s} equal to zero:

$$[(q+s+2)(q+s+1)-l(l+1)]a_{q+2}+[\varepsilon_{k,l}-\beta^2-2\beta^2(q+s)]a_q=0 \tag{23}$$

that is, using (20):

$$(q+2)(q+2l+3)a_{q+2}=[(2q+2l+3)\beta^2-\varepsilon_{k,l}]a_q \tag{24}$$

We therefore obtain a recurrence relation for the coefficients a_q of expansion (17).

Note, first of all, that this recurrence relation, combined with result (22), implies that *all coefficients a_q of odd rank q are zero.* As for the coefficients of even rank, they must all be proportional to a_0. If the value of $\varepsilon_{k,l}$ is such that no integer q makes the term in brackets on the right-hand side of (24) go to zero, we find the solution $y_{k,l}$ of (16) in the form of an infinite power series, for which:

$$\frac{a_{q+2}}{a_q} \underset{q\to\infty}{\sim} \frac{2\beta^2}{q} \tag{25}$$

This behavior is the same as that of the coefficients appearing in the expansion of the function $e^{\beta^2 r^2}$, since:

$$e^{\beta^2 r^2}=\sum_{p=0}^{\infty} c_{2p}\, r^{2p} \tag{26}$$

with:

$$c_{2p}=\frac{\beta^{2p}}{p!} \tag{27}$$

and, consequently:

$$\frac{c_{2p+2}}{c_{2p}} \underset{p\to\infty}{\sim} \frac{\beta^2}{p} \tag{28}$$

Since it is $2p$ that corresponds to the even integer q of the expansion of $y_{k,l}$, (28) is indeed identical to (25). From this, we can see that if (17) really contains an infinite number of terms, the asymptotic behavior of $y_{k,l}$ is dominated by $e^{\beta^2 r^2}$, which renders this function physically unacceptable [*cf.* relation (15)].

The only cases which are interesting from a physical point of view are therefore those in which there exists an even integer k, positive or zero, such that:

$$\varepsilon_{k,l}=(2k+2l+3)\beta^2 \tag{29}$$

Recurrence relation (24) indicates that the coefficients of even rank greater than k are then zero. Since all the coefficients of odd rank are also zero, expansion (17) reduces to a polynomial, and the radial function $u_{k,l}(r)$ given by (15) decreases exponentially as r goes to infinity.

2. Energy levels and stationary wave functions

Using definitions (7) and (11b), relation (29) gives the energies $E_{k,l}$ associated with a given value of l:

$$E_{k,l} = \hbar\omega \left(k + l + \frac{3}{2} \right) \tag{30}$$

where k is any even positive integer or zero. Since $E_{k,l}$ actually depends only on the sum:

$$n = k + l \tag{31}$$

accidental degeneracies appear: the energy levels of the isotropic three-dimensional harmonic oscillator are of the form:

$$E_n = \left(n + \frac{3}{2} \right) \hbar\omega \tag{32}$$

l is any positive integer or zero, and k is any even positive integer or zero; n can therefore take on all integral values, positive or zero. This is in agreement with result (4).

We shall fix an energy E_n, that is, an integer n, positive or zero. The values of k and l which can be associated with it according to (31) are the following:

$$(k,l) = (0,n), \quad (2,n-2), \ldots (n-2,2), (n,0) \quad \text{if } n \text{ is even} \tag{33a}$$

$$(k,l) = (0,n), \quad (2,n-2), \ldots (n-3,3), (n-1,1) \quad \text{if } n \text{ is odd} \tag{33b}$$

From this, we can immediately get the values of l associated with the first values of n:

$$\begin{aligned} &n = 0 : l = 0 \\ &n = 1 : l = 1 \\ &n = 2 : l = 0, 2 \\ &n = 3 : l = 1, 3 \\ &n = 4 : l = 0, 2, 4 \end{aligned} \tag{34}$$

Figure 1 represents, with the same conventions as for the hydrogen atom (*cf.* Figure 4 of Chapter VII), the lowest energy levels of an isotropic three-dimensional harmonic oscillator.

For each pair (k,l), there exists one and only one radial function $u_{k,l}(r)$, which corresponds to $(2l+1)$ common eigenfunctions of H, $\mathbf{L}^2$ and L_z:

$$\varphi_{k,l,m}(\mathbf{r}) = \frac{1}{r} u_{k,l}(r)\ Y_l^m(\theta,\varphi) \tag{35}$$

Consequently, the degree of degeneracy of the energy E_n under consideration is equal to:

$$g_n = \sum_{l=0,2,\ldots,n} (2l+1) \qquad \text{if } n \text{ is even} \tag{36a}$$

$$g_n = \sum_{l=1,3,\ldots,n} (2l+1) \qquad \text{if } n \text{ is odd} \tag{36b}$$

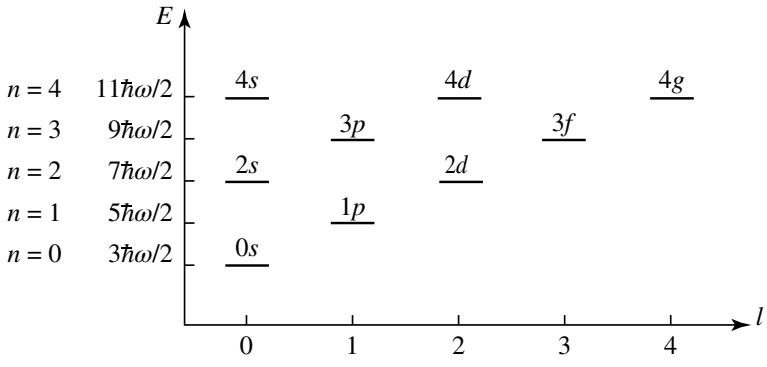


Figure 1: Lowest energy levels of the three-dimensional harmonic oscillator. When n is even, l can take on $\frac{n}{2}+1$ values: $l = n,\ n-2,\ ...,\ 0$. When n is odd, l can take on $\frac{n+1}{2}$ values: $l = n, n-2,\ ...,\ 1$. With the possible values of m $(-l \leq m \leq l)$ taken into account, the degree of degeneracy of the level E_n is $\frac{(n+1)(n+2)}{2}$.

These sums are simple to calculate, and we again obtain result (5):

$$\text{for even } n: \qquad g_n = \sum_{p=0}^{n/2}(4p+1) = \frac{1}{2}(n+1)(n+2) \tag{37a}$$

$$\text{for odd } n: \qquad g_n = \sum_{p=0}^{(n-1)/2}(4p+3) = \frac{1}{2}(n+1)(n+2) \tag{37b}$$

For each of the pairs (k, l) given in (33), the results of § 1 enable us to determine the corresponding radial function $u_{k,l}(r)$ (to within the factor a_0) and, therefore, the $(2l+1)$ common eigenfunctions of H and $\mathbf{L}^2$, of eigenvalues E_n and $l(l+1)\hbar^2$. We shall calculate, for example, the wave functions associated in this way with the three lowest energy levels.

For the ground state $E_0 = \frac{3}{2}\hbar\omega$, we must have:

$$k = l = 0 \tag{38}$$

$y_{0,0}(r)$ then reduces to $a_0 r$. If we choose a_0 to be real and positive, the normalized function $\varphi_{k=l=m=0}$ can be written:

$$\varphi_{0,0,0}(\mathbf{r}) = \left(\frac{\beta^2}{\pi}\right)^{3/4} \mathrm{e}^{-\beta^2 r^2/2} \tag{39}$$

Since the ground state is not degenerate $(g_0 = 1)$, $\varphi_{0,0,0}$ is the same as the function $\varphi_{n_x=n_y=n_z=0}$ which is found by separating the Cartesian variables x, y, and z [*cf.* formula (6)].

With the first excited state $E_1 = \frac{5}{2}\hbar\omega$, which is three-fold degenerate, is again associated a single pair (k, l):

$$\begin{cases} k = 0 \\ l = 1 \end{cases} \tag{40}$$

and $y_{0,1} = a_0 r^2$. The three functions of the basis defined by $\mathbf{L}^2$ and L_z are therefore:

$$\varphi_{0,1,m}(\mathbf{r}) = \sqrt{\frac{8}{3}} \frac{\beta^{3/2}}{\pi^{1/4}} \, \beta r \, \mathrm{e}^{-\beta^2 r^2/2} \, Y_1^m(\theta, \varphi) \qquad m = 1, 0, -1 \tag{41}$$

We know [*cf.* Complement $\mathrm{A_{VI}}$, formulas (32)] that the spherical harmonics Y_1^m are such that:

$$\begin{aligned} r\, Y_1^0(\theta, \varphi) &= \sqrt{\frac{3}{4\pi}} \, z \\ \frac{r}{\sqrt{2}} \left[Y_1^{-1} - Y_1^1 \right] &= \sqrt{\frac{3}{4\pi}} \, x \\ \frac{r}{\sqrt{2}} \left[Y_1^{-1} + Y_1^1 \right] &= -i \sqrt{\frac{3}{4\pi}} \, y \end{aligned} \tag{42}$$

and that the Hermite polynomial of order 1 is [*cf.* Complement $\mathrm{B_{VI}}$, formulas (18)]:

$$H_1(u) = 2u \tag{43}$$

Consequently, it is clear that the three functions $\varphi_{0,1,m}$ are related to the functions φ_{n_x,n_y,n_z} of basis (6) by the equations:

$$\begin{aligned} \varphi_{n_x=0,n_y=0,n_z=1} &= \varphi_{k=0,l=1,m=0} \\ \varphi_{n_x=1,n_y=0,n_z=0} &= \frac{1}{\sqrt{2}} \left[\varphi_{k=0,l=1,m=-1} - \varphi_{k=0,l=1,m=1} \right] \\ \varphi_{n_x=0,n_y=1,n_z=0} &= \frac{i}{\sqrt{2}} \left[\varphi_{k=0,l=1,m=-1} + \varphi_{k=0,l=1,m=1} \right] \end{aligned} \tag{44}$$

Finally, consider the second excited state, of energy $E_2 = \frac{7}{2}\hbar\omega$. It is six-fold degenerate, and the quantum numbers k and l can take on the values:

$$k = 0, \quad l = 2 \tag{45a}$$

$$k = 2, \quad l = 0 \tag{45b}$$

The function $y_{0,2}(r)$ which corresponds to the values (45a) is simply $a_0 r^3$. For the values (45b), $y_{2,0}$ contains two terms; using (24) and (29), we easily find:

$$y_{2,0}(r) = a_0 r \left[1 - \frac{2}{3} \beta^2 r^2 \right] \tag{46}$$

The six basis functions in the eigensubspace associated with E_2 are thus of the form:

$$\varphi_{0,2,m}(\mathbf{r}) = \sqrt{\frac{16}{15}} \frac{\beta^{3/2}}{\pi^{1/4}} \, \beta^2 r^2 \, \mathrm{e}^{-\beta^2 r^2/2} \, Y_2^m(\theta, \varphi) \qquad m = 2, 1, 0, -1, -2 \tag{47a}$$

$$\varphi_{2,0,0}(\mathbf{r}) = \sqrt{\frac{3}{2}} \frac{\beta^{3/2}}{\pi^{3/4}} \left(1 - \frac{2}{3} \beta^2 r^2 \right) \mathrm{e}^{-\beta^2 r^2/2} \tag{47b}$$

Knowing the explicit expressions for the spherical harmonics [formulas (33) of Complement A_{VI}] and the Hermite polynomials [formulas (18) of Complement B_V], we can easily prove the following relations:

$$\varphi_{k=2,l=0,m=0} = -\sqrt{\frac{1}{3}}\left[\varphi_{n_x=2,n_y=0,n_z=0} + \varphi_{n_x=0,n_y=2,n_z=0} + \varphi_{n_x=0,n_y=0,n_z=2}\right]$$

$$\frac{1}{\sqrt{2}}\left[\varphi_{k=0,l=2,m=2} + \varphi_{k=0,l=2,m=-2}\right] = \frac{1}{\sqrt{2}}\left[\varphi_{n_x=2,n_y=0,n_z=0} - \varphi_{n_x=0,n_y=2,n_z=0}\right]$$

$$\frac{1}{\sqrt{2}}\left[\varphi_{k=0,l=2,m=2} - \varphi_{k=0,l=2,m=-2}\right] = i\varphi_{n_x=1,n_y=1,n_z=0}$$

$$\frac{1}{\sqrt{2}}\left[\varphi_{k=0,l=2,m=1} - \varphi_{k=0,l=2,m=-1}\right] = -\varphi_{n_x=1,n_y=0,n_z=1}$$

$$\frac{1}{\sqrt{2}}\left[\varphi_{k=0,l=2,m=1} + \varphi_{k=0,l=2,m=-1}\right] = -i\varphi_{n_x=0,n_y=1,n_z=1}$$

$$\varphi_{k=0,l=2,m=0} = \sqrt{\frac{2}{3}}\left[\varphi_{n_x=0,n_y=0,n_z=2} - \frac{1}{2}\varphi_{n_x=2,n_y=0,n_z=0} - \frac{1}{2}\varphi_{n_x=0,n_y=2,n_z=0}\right] \quad (48)$$

Comment:

As we pointed out in the beginning of this complement, we can apply a method analogous to the one presented in Complement D_{VI} to the isotropic three-dimensional harmonic oscillator. If a_x, a_y and a_z are the annihilation operators which act in the state spaces $\mathcal{E}_x$, $\mathcal{E}_y$ and $\mathcal{E}_z$ respectively, we define:

$$a_r = \frac{1}{\sqrt{2}}(a_x - ia_y) \quad (49a)$$

$$a_l = \frac{1}{\sqrt{2}}(a_x + ia_y) \quad (49b)$$

It can be shown that a_r and a_l behave like independent annihilation operators (Complement D_{VI}, § 3-b). The Hamiltonian H and the angular momentum operators can then be expressed in terms of a_r, a_l, a_z and their adjoints:

$$H = \hbar\omega\left(N_r + N_l + N_z + \frac{3}{2}\right) \quad (50a)$$

$$L_z = \hbar\left(N_r - N_l\right) \quad (50b)$$

$$L_+ = \hbar\sqrt{2}\left(a_z^\dagger a_l - a_r^\dagger a_z\right) \quad (50c)$$

$$L_- = \hbar\sqrt{2}\left(a_l^\dagger a_z - a_z^\dagger a_r\right) \quad (50d)$$

The common eigenvectors $|\chi_{n_r,n_l,n_z}\rangle$ of the observables N_r, N_l and N_z can be obtained through the action of the creation operators $a_r^\dagger$, $a_l^\dagger$ and $a_z^\dagger$ on the ground

state $|0,0,0\rangle$ of the Hamiltonian H [this state is unique to within a constant factor; *cf.* formulas (6) and (39)]:

$$|\chi_{n_r,n_l,n_z}\rangle = \frac{1}{\sqrt{n_r!n_l!n_z!}}(a_r^\dagger)^{n_r}(a_l^\dagger)^{n_l}(a_z^\dagger)^{n_z}|0,0,0\rangle \tag{51}$$

According to (50a) and (50b), $|\chi_{n_r,n_l,n_z}\rangle$ is an eigenvector of H and L_z with the eigenvalues $(n_r+n_l+n_z+3/2)\hbar\omega$ and $(n_r-n_l)\hbar$. The eigensubspace $\mathcal{E}_n$, associated with a given energy E_n can therefore be spanned by the set of vectors $|\chi_{n_r,n_l,n_z}\rangle$ such that:

$$n_r + n_l + n_z = n \tag{52}$$

Of these, the eigenvector of L_z with the largest eigenvalue compatible with E_n is $|\chi_{n,0,0}\rangle$, whose eigenvalue is $n\hbar$. This ket, according to (50c), satisfies:

$$L_+|\chi_{n,0,0}\rangle = 0 \tag{53}$$

Consequently[2], it is an eigenvector of $\mathbf{L}^2$ with the eigenvalue $n(n+1)\hbar^2$, and it can be identified with the ket of the $\{|\varphi_{k,l,m}\rangle\}$ basis such that:

$$\begin{aligned} &k + l = n \\ &l = m = n \end{aligned} \tag{54}$$

Therefore:

$$|\varphi_{k=0,l=n,m=n}\rangle = |\chi_{n_r=n,n_l=0,n_z=0}\rangle \tag{55}$$

Application of the operator L_- [formula (50d)] to both sides of relation (55) yields:

$$|\varphi_{0,n,n-1}\rangle = -|\chi_{n-1,0,1}\rangle \tag{56}$$

The eigenvalue $(n-2)\hbar$ of L_z, unlike the two preceding ones, is two-fold degenerate in $\mathcal{E}_n$: two orthogonal vectors, $|\chi_{n-2,0,2}\rangle$ and $|\chi_{n-1,1,0}\rangle$ correspond to it. Using (50d) again in order to apply L_- to (56), we find that:

$$|\varphi_{0,n,n-2}\rangle = \sqrt{\frac{2(n-1)}{2n-1}}|\chi_{n-2,0,2}\rangle - \frac{1}{\sqrt{2n-1}}|\chi_{n-1,1,0}\rangle \tag{57}$$

It can be shown that the action of L_+ on the linear combination orthogonal to (57) yields the null vector. This linear combination must therefore be an eigenvector of $\mathbf{L}^2$ with the eigenvalue $(n-2)(n-1)\hbar^2$. This gives, to within a phase factor:

$$|\varphi_{2,n-2,n-2}\rangle = \frac{1}{\sqrt{2n-1}}|\chi_{n-2,0,2}\rangle + \sqrt{\frac{2(n-1)}{2n-1}}|\chi_{n-1,1,0}\rangle \tag{58}$$

[2]This result follows directly from relation (C-7b) of Chapter VI, which, applied to $|\chi_{n,0,0}\rangle$, yields:

$$\mathbf{L}^2|\chi_{n,0,0}\rangle = \hbar^2(n^2+n)|\chi_{n,0,0}\rangle$$

We can thus relate, by iteration[3], the two bases, $\{|\chi_{n_r,n_l,n_z}\rangle\}$ and $\{|\varphi_{k,l,m}\rangle\}$. Of course, replacing $a_r^\dagger$ and $a_l^\dagger$ by functions of $a_x^\dagger$ and $a_y^\dagger$ in (51), we can express $|\chi_{n_r,n_l,n_z}\rangle$ as a linear combination of the vectors $|\varphi_{n_x,n_y,n_z}\rangle$ whose wave functions are given by (6).

References and suggestions for further reading:

Other soluble examples (spherical square well, etc.): Messiah (1.17), Chap. IX, § 10; Schiff (1.18), § 15; see also Flügge (1.24), §§ 58 to 79.

[3]An argument analogous to the one just outlined will be used in Chapter X to add two angular momenta.

Complement C_{VII}

Probability currents associated with the stationary states of the hydrogen atom

1 General expression for the probability current 851

2 Application to the stationary states of the hydrogen atom 852

2-a Structure of the probability current 852

2-b Effect of a magnetic field 853

The normalized wave functions $\varphi_{n,l,m}(\mathbf{r})$ associated with the stationary states of the hydrogen atom were determined in Chapter VII. $\varphi_{n,l,m}(\mathbf{r})$ is the product of the spherical harmonic $Y_l^m(\theta,\varphi)$ and the function $R_{n,l}(r)$ calculated in § C-3 of that chapter:

$$\varphi_{n,l,m}(\mathbf{r}) = R_{n,l}(r)Y_l^m(\theta,\varphi) \tag{1}$$

The spatial variation of the probability density:

$$\rho_{n,l,m}(\mathbf{r}) = |\varphi_{n,l,m}(\mathbf{r})|^2 \tag{2}$$

was then studied, at least for the lowest energy states.

It is important, however, to understand that a stationary state is not characterized only by the value of the probability density $\rho_{n,l,m}(\mathbf{r})$ at all points of space. We must also associate with it a probability current, which can be expressed as:

$$\mathbf{J}_{n,l,m}(\mathbf{r}) = \frac{\hbar}{2\mu i}\varphi^*_{n,l,m}(\mathbf{r})\,\boldsymbol{\nabla}\varphi_{n,l,m}(\mathbf{r}) + c.c. \tag{3}$$

[we assume here that the vector potential $\mathbf{A}(\mathbf{r},t)$ is zero; μ denotes the mass of the particle].

Thus, we can associate with the quantum state of a particle, a "fluid" (called the "probability fluid") whose density at each point of space is $\rho(\mathbf{r})$. This fluid is not motionless; it is in a state of flow characterized by the current density $\mathbf{J}$. In a stationary state, ρ and $\mathbf{J}$ are not time-dependent: the fluid is in a steady state of flow.

To complete the results of Chapter VII concerning the physical properties of the stationary states of the hydrogen atom, we shall now study the probability currents $\mathbf{J}_{n,l,m}(\mathbf{r})$.

1. General expression for the probability current

Consider an arbitrary normalized wave function $\psi(\mathbf{r})$. We introduce the real quantities $\alpha(\mathbf{r})$ [the modulus of $\psi(\mathbf{r})$] and $\xi(\mathbf{r})$ [the argument of $\psi(\mathbf{r})$] by setting:

$$\psi(\mathbf{r}) = \alpha(\mathbf{r})\,\mathrm{e}^{i\xi(\mathbf{r})} \tag{4}$$

with:

$$\alpha(\mathbf{r}) \geq 0 \quad ; \qquad 0 \leq \xi(\mathbf{r}) < 2\pi \tag{5}$$

If we substitute (4) into the expressions for the probability density $\rho(\mathbf{r})$ and the current $\mathbf{J}(\mathbf{r})$, we obtain [still assuming the vector potential $\mathbf{A}(\mathbf{r})$ to be zero]:

$$\rho(\mathbf{r}) = \alpha^2(\mathbf{r}) \tag{6}$$

$$\mathbf{J}(\mathbf{r}) = \frac{\hbar}{\mu}\alpha^2(\mathbf{r})\,\boldsymbol{\nabla}\xi(\mathbf{r}) \tag{7}$$

Therefore, $\rho(\mathbf{r})$ depends only on the modulus of the wave function, while $\mathbf{J}(\mathbf{r})$ brings in its phase [for example, $\mathbf{J}(\mathbf{r})$ is zero if this phase is constant throughout all space].

Comment:

If the wave function $\psi(\mathbf{r})$ is given, it is obvious that $\rho(\mathbf{r})$ and $\mathbf{J}(\mathbf{r})$ are perfectly well-defined. Conversely, is there always one and only one function $\psi(\mathbf{r})$ corresponding to given values of $\rho(\mathbf{r})$ and $\mathbf{J}(\mathbf{r})$?

According to (6), the modulus $\alpha(\mathbf{r})$ of the wave function can be obtained directly[1] from $\rho(\mathbf{r})$; the argument $\xi(\mathbf{r})$ must satisfy the equation:

$$\boldsymbol{\nabla}\xi(\mathbf{r}) = \frac{\mu}{\hbar}\frac{\mathbf{J}(\mathbf{r})}{\rho(\mathbf{r})} \tag{8}$$

We know that such an equation has a solution only if:

$$\boldsymbol{\nabla} \times \frac{\mathbf{J}(\mathbf{r})}{\rho(\mathbf{r})} = \mathbf{0} \tag{9}$$

It then has an infinite number of solutions, which differ from each other by a constant. Since this constant corresponds to a global phase factor, it follows, if condition (9) is satisfied, that the wave function of the particle is perfectly well-defined by the specification of $\rho(\mathbf{r})$ and $\mathbf{J}(\mathbf{r})$. If condition (9) is not fulfilled, no wave function exists that corresponds to the values of $\rho(\mathbf{r})$ and $\mathbf{J}(\mathbf{r})$ under consideration.

2. Application to the stationary states of the hydrogen atom

2-a. Structure of the probability current

When the wave function has the form of (1), where $R_{n,l}(r)$ is a real function and $Y_l^m(\theta,\varphi)$ is the product of $\mathrm{e}^{im\varphi}$ and a real function, we have:

$$\begin{aligned} \alpha_{n,l,m}(\mathbf{r}) &= |R_{n,l}(r)|\,|Y_l^m(\theta,\varphi)| \\ \xi_{n,l,m}(\mathbf{r}) &= m\varphi \end{aligned} \tag{10}$$

Applying (7) and using the expression for the gradient in polar coordinates, we get:

$$\mathbf{J}_{n,l,m}(\mathbf{r}) = \frac{\hbar}{\mu}\,m\,\frac{\rho_{n,l,m}(\mathbf{r})}{r\sin\theta}\,\mathbf{e}_\varphi(\mathbf{r}) \tag{11}$$

[1] Of course, in order to be a probability density, $\rho(\mathbf{r})$ must be positive everywhere.

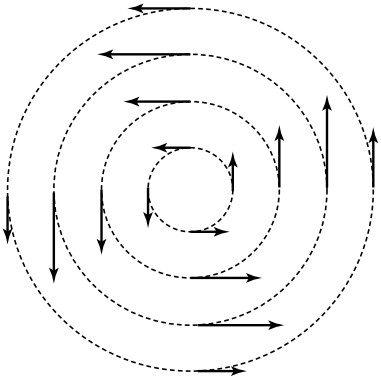


Figure 1: Structure of the probability current associated with a stationary state $|\varphi_{n,l,m}\rangle$ of the hydrogen atom (shown in a plane perpendicular to Oz). The index m refers to the eigenvalue $m\hbar$ of L_z. If $m > 0$, the probability fluid rotates counterclockwise about Oz; if $m < 0$, clockwise. If $m = 0$, the probability current is zero at all points of space.

where $\mathbf{e}_\varphi(\mathbf{r})$ is the unit vector that forms with Oz and $\mathbf{r}$ a right-handed Cartesian coordinate frame.

Probability current variations in a plane perpendicular to Oz are shown in Figure 1. According to (11), the current at each point M is perpendicular to the plane defined by M and the Oz axis: the probability fluid rotates about Oz. Since $|\mathbf{J}|$ is not proportional to $r \sin\theta\, \rho(\mathbf{r})$, the system does not rotate as a whole. The eigenvalue $m\hbar$ of the observable L_z can be interpreted as the classical angular momentum associated with this rotational motion of the probability fluid. The contribution of the volume element d^3r, situated at the point $\mathbf{r}$, to the angular momentum with respect to the origin can be written:

$$\mathrm{d}\boldsymbol{\mathcal{L}} = \mu \mathbf{r} \times \mathbf{J}_{n,l,m}(\mathbf{r})\ \mathrm{d}^3r \tag{12}$$

By symmetry, the resultant of all these elementary momenta is directed along Oz; it is equal to:

$$\mathcal{L}_z = \mu \int \mathrm{d}^3r\, \mathbf{e}_z \cdot [\mathbf{r} \times \mathbf{J}_{n,l,m}(\mathbf{r})] \tag{13}$$

Using expression (11) for $\mathbf{J}_{n,l,m}(\mathbf{r})$, we easily obtain:

$$\begin{aligned} \mathcal{L}_z &= \mu \int \mathrm{d}^3r\ \ r|\mathbf{J}_{n,l,m}(\mathbf{r})| \sin\theta \\ &= m\hbar \int \mathrm{d}^3r\, \rho_{n,l,m}(\mathbf{r}) \\ &= m\hbar \end{aligned} \tag{14}$$

2-b. Effect of a magnetic field

The results obtained thus far are valid only if the vector potential $\mathbf{A}(\mathbf{r})$ is zero. Let us examine what happens when this is not the case. Imagine, for example, that we place the hydrogen atom in a uniform magnetic field $\mathbf{B}$. This field can be described by the vector potential:

$$\mathbf{A}(\mathbf{r}) = -\frac{1}{2}\mathbf{r} \times \mathbf{B} \tag{15}$$

What is the probability current associated with the ground state?

For simplicity, we shall assume that the magnetic field $\mathbf{B}$ does not modify the wave function of the ground state[2]. The probability current can then be calculated from the general expression for $\mathbf{J}$ [*cf.* relation (D-20) of Chapter III]. This yields:

$$\begin{aligned}\mathbf{J}_{n,l,m}(\mathbf{r}) &= \frac{1}{2\mu}\left\{\varphi^*_{n,l,m}(\mathbf{r})\left[\frac{\hbar}{i}\boldsymbol{\nabla} - q\mathbf{A}(\mathbf{r})\right]\varphi_{n,l,m}(\mathbf{r}) + c.c.\right\} \\ &= \frac{1}{\mu}\rho_{n,l,m}(\mathbf{r})\left[\hbar\boldsymbol{\nabla}\xi_{n,l,m}(\mathbf{r}) - q\mathbf{A}(\mathbf{r})\right] \end{aligned} \tag{16}$$

For the ground state and a field $\mathbf{B}$ directed along Oz, we obtain, using (15):

$$\mathbf{J}_{1,0,0}(\mathbf{r}) = \frac{\omega_c}{2}\rho_{1,0,0}(\mathbf{r})\,\mathbf{e}_z \times \mathbf{r} \tag{17}$$

where the cyclotron frequency ω_c is defined by:

$$\omega_c = -\frac{qB}{\mu} \tag{18}$$

Therefore, the probability current of the ground state is not zero in the presence of a magnetic field as it is when $\mathbf{B} = \mathbf{0}$. Expression (17) indicates that the probability fluid rotates as a whole about $\mathbf{B}$ with the angular velocity $\omega_c/2$. Physically, this result arises from the fact that when the magnetic field $\mathbf{B}$ is turned on, a transient electric field $\mathbf{E}(t)$ must exist. Under its influence, the electron, while remaining in the ground state, goes into rotation about the proton, with an angular velocity depending only on the value of $\mathbf{B}$ (and not on the precise way in which the field was turned on during the transition period).

Comment:

The particular choice of gauge made in (15) allowed us, with only a negligible error, to retain the same wave functions we used in the absence of a field (*cf.* footnote 2). With another gauge, the wave functions would have been different (*cf.* Complement H_{III}) and, in (16), the term containing $\mathbf{A}(\mathbf{r})$ explicitly would not have been the only one to contribute to the value of $\mathbf{J}(\mathbf{r})$, to first order in $\mathbf{B}$. Nevertheless, we would have found (17) at the end of the calculation, since the physical result must not be gauge-dependent.

[2]Since the Hamiltonian H depends on $\mathbf{B}$, this is obviously not rigorously correct. Nevertheless, by considering the expression for H [*cf.* formulas (6) and (7) of Complement D_{VII}], we can show that, for the gauge chosen in (15) and $\mathbf{B}$ directed along Oz, the functions $\varphi_{n,l,m}(\mathbf{r})$ are eigenfunctions of H to within a term of second order in $\mathbf{B}$. By using the perturbation theory of Chapter XI, one can show that, for magnetic fields normally produced in the laboratory, this second-order term is negligible.

Complement D_{VII}

The hydrogen atom placed in a uniform magnetic field. Paramagnetism and diamagnetism. The Zeeman effect

1 The Hamiltonian of the problem. The paramagnetic term and the diamagnetic term 856

1-a Expression for the Hamiltonian 856

1-b Order of magnitude of the various terms 857

1-c Interpretation of the paramagnetic term 859

1-d Interpretation of the diamagnetic term 860

2 The Zeeman effect 862

2-a Energy levels of the atom in the presence of a magnetic field 862

2-b Electric dipole oscillations 863

2-c Frequency and polarization of emitted radiation 865

In Chapter VII, we studied the quantum mechanical properties of a free hydrogen atom, that is, of the system formed by an electron and a proton exerting an electrostatic attraction on each other but not interacting with any external field. This complement is devoted to the study of the new effects that appear when this atom is placed in a static magnetic field. We shall confine ourselves to the case in which this field is uniform, as it always is, moreover, in practice, since the magnetic fields that can be produced in the laboratory vary very little in relative value over distances comparable to atomic dimensions.

We have already studied the behavior of an electron subjected either to an electric field alone (*cf.* for example, Chapter VII) or to a magnetic field alone (*cf.* Complement E_{VI}). Here, we shall generalize these discussions by calculating the energy levels of an electron subjected both to the influence of the internal electric field of the atom and to an external magnetic field. Under these conditions, the exact solution of the Schrödinger equation may seem to be a very complicated problem. However, we shall see that this problem can be simplified considerably by means of certain approximations. In the first place, we shall totally neglect the nuclear finite mass effect[1]. Then we shall use the fact that, in practice, the effect of the external magnetic field is much smaller than that of the internal electric field of the atom: the atomic level shifts due to the magnetic field are much smaller than the energy separations in a zero field.

The discussion presented in this complement will enable us to introduce and explain certain effects which are important in atomic physics. We shall see, in particular, how atomic paramagnetism and diamagnetism appear in the quantum mechanical formalism.

[1] For the hydrogen atom, such an approximation is justified by the fact that the proton is considerably heavier than the electron. For muonium (*cf.* Complement A_{VII}), the approximation is not as good and it becomes totally inapplicable for the case of positronium. Moreover, in the presence of a magnetic field, it is not possible to separate the motion of the center of mass rigorously. If one wished to take the nuclear finite mass effect into account in this complement, it would not suffice to replace the mass m_e of the electron by the reduced mass μ of the electron-proton system.

In addition, we shall be able to predict the modifications occurring in the optical spectrum emitted by hydrogen atoms when they are placed in a static magnetic field (the Zeeman effect).

1. The Hamiltonian of the problem. The paramagnetic term and the diamagnetic term

1-a. Expression for the Hamiltonian

Consider a spinless particle, of mass m_e and charge q, subjected simultaneously to a scalar central potential $V(r)$ and a vector potential $\mathbf{A}(\mathbf{r})$. Its Hamiltonian is:

$$H = \frac{1}{2m_e}\left[\mathbf{P} - q\mathbf{A}(\mathbf{R})\right]^2 + V(\mathbf{R}) \tag{1}$$

When the magnetic field $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r})$ is uniform, the vector potential $\mathbf{A}$ can be put into the form:

$$\mathbf{A}(\mathbf{r}) = -\frac{1}{2}\mathbf{r} \times \mathbf{B} \tag{2}$$

To substitute this expression into (1), we shall calculate the quantity:

$$[\mathbf{P} - q\mathbf{A}(\mathbf{R})]^2 = \mathbf{P}^2 + \frac{q}{2}[\mathbf{P} \cdot (\mathbf{R} \times \mathbf{B}) + (\mathbf{R} \times \mathbf{B}) \cdot \mathbf{P}] + \frac{q^2}{4}(\mathbf{R} \times \mathbf{B})^2 \tag{3}$$

Now, $\mathbf{B}$ is actually a constant, and not an operator. All observables therefore commute with $\mathbf{B}$, so we can write, using the rules of vector calculus:

$$[\mathbf{P} - q\mathbf{A}(\mathbf{R})]^2 = \mathbf{P}^2 + \frac{q}{2}[\mathbf{B} \cdot (\mathbf{P} \times \mathbf{R}) - (\mathbf{R} \times \mathbf{P}) \cdot \mathbf{B}] + \frac{q^2}{4}[\mathbf{R}^2\mathbf{B}^2 - (\mathbf{R} \cdot \mathbf{B})^2] \tag{4}$$

On the right-hand side of this expression, the angular momentum $\mathbf{L}$ of the particle appears:

$$\mathbf{L} = \mathbf{R} \times \mathbf{P} = -\mathbf{P} \times \mathbf{R} \tag{5}$$

We can therefore write H in the form:

$$H = H_0 + H_1 + H_2 \tag{6}$$

where H_0, H_1 and H_2 are defined by:

$$H_0 = \frac{\mathbf{P}^2}{2m_e} + V(\mathbf{R}) \tag{7a}$$

$$H_1 = -\frac{\mu_B}{\hbar}\mathbf{L} \cdot \mathbf{B} \tag{7b}$$

$$H_2 = \frac{q^2\mathbf{B}^2}{8m_e}\mathbf{R}_\perp^2 \tag{7c}$$

In these relations, μ_B denotes the Bohr magneton (whose dimensions are those of a magnetic moment):

$$\mu_B = \frac{q\hbar}{2m_e} \tag{8}$$

and the operator $\mathbf{R}_\perp$ is the projection of $\mathbf{R}$ onto a plane perpendicular to $\mathbf{B}$:

$$\mathbf{R}_\perp^2 = \mathbf{R}^2 - \frac{(\mathbf{R}\cdot\mathbf{B})^2}{\mathbf{B^2}} \tag{9}$$

If we choose a system of orthonormal axes $Oxyz$ such that $\mathbf{B}$ is parallel to Oz, we have:

$$\mathbf{R}_\perp^2 = X^2 + Y^2 \tag{10}$$

Comment:

When the field $\mathbf{B}$ is zero, H becomes equal to H_0, which is the sum of the kinetic energy $\mathbf{P}^2/2m_e$ and the potential energy $V(\mathbf{R})$. Nevertheless, we must not conclude from this that when $\mathbf{B}$ is not zero, $\mathbf{P}^2/2m_e$ still represents the kinetic energy of the electron. We have seen (*cf.* Complement $\mathrm{H_{III}}$) that the physical meaning of operators acting in the state space changes when the vector potential is not zero. For example, the momentum $\mathbf{P}$ no longer represents the mechanical momentum $\mathbf{\Pi} = m_e\mathbf{V}$, and the kinetic energy is then equal to:

$$\frac{\mathbf{\Pi}^2}{2m_e} = \frac{1}{2m_e}[\mathbf{P} - q\mathbf{A}(\mathbf{R})]^2 \tag{11}$$

The meaning of the term $\mathbf{P}^2/2m_e$, taken alone, depends on the gauge chosen. With the one defined by (2), it can easily be shown to correspond to the "relative" kinetic energy $\mathbf{\Pi}_{\boldsymbol{R}}^2/2m_e$, where $\mathbf{\Pi}_{\boldsymbol{R}}$ is the mechanical momentum of the particle with respect to the "Larmor frame" rotating about $\mathbf{B}$ with angular velocity $\omega_L = -qB/2m_e$. The term H_2 then describes the kinetic energy $\mathbf{\Pi}_{\boldsymbol{E}}^2/2m_e$ related to the velocity of the frame. As for H_1, it corresponds to the cross term $\mathbf{\Pi}_{\boldsymbol{E}} \cdot \mathbf{\Pi}_{\boldsymbol{R}}/m_e$.

1-b. Order of magnitude of the various terms

In the presence of the magnetic field $\mathbf{B}$, two new terms, H_1 and H_2, therefore appear in H. Before examining their physical meaning in greater detail, we shall calculate the order of magnitude of the energy differences ΔE (or the frequency differences $\Delta E/h$) associated with them.

As far as H_0 is concerned, we already know the corresponding energy differences ΔE_0 (*cf.* Chap. VII). The associated frequencies are of the order of:

$$\frac{\Delta E_0}{h} \simeq 10^{14} \text{ to } 10^{15} \text{ Hz} \tag{12}$$

Also, by using (7b), we see that ΔE_1 is approximately given by:

$$\frac{\Delta E_1}{h} \simeq \frac{1}{h}\left(\frac{\mu_B}{\hbar}\hbar B\right) = \frac{\omega_L}{2\pi} \tag{13}$$

where ω_L is the Larmor angular velocity[2]:

$$\omega_L = -\frac{qB}{2m_e} \tag{14}$$

[2]Note that the Larmor frequency $\frac{\omega_L}{2\pi}$ is half the cyclotron frequency.

A simple numerical calculation shows that, for an electron, the Larmor frequency is such that:

$$\frac{\nu_L}{B} = \frac{\omega_L}{2\pi B} \simeq 1.40 \times 10^{10} \text{ Hz/tesla } = 1.40 \text{ MHz/gauss} \tag{15}$$

Now, with the fields usually produced in the laboratory (which rarely exceed 100,000 gauss), we have:

$$\frac{\omega_L}{2\pi} \lesssim 10^{11} \text{ Hz} \tag{16}$$

Comparing (12) and (16), we see that:

$$\Delta E_1 \ll \Delta E_0 \tag{17}$$

Let us show, similarly, that:

$$\Delta E_2 \ll \Delta E_1 \tag{18}$$

To do so, we shall evaluate the order of magnitude ΔE_2 of the energies associated with H_2. The matrix elements of the operator $\mathbf{R}_\perp^2 = X^2 + Y^2$ are of the same order of magnitude as a_0^2, where $a_0 = \hbar^2/m_e e^2$ characterizes atomic dimensions. Thus we obtain:

$$\Delta E_2 \simeq \frac{q^2 B^2}{m_e} a_0^2 \tag{19}$$

We then find the ratio:

$$\frac{\Delta E_2}{\Delta E_1} \simeq \frac{q^2 B^2}{m_e} a_0^2 \frac{1}{\hbar \omega_L} = 2\hbar \frac{qB}{m_e} \frac{m_e a_0^2}{\hbar^2} \tag{20}$$

Now, according to formulas (C-12a) and (C-12b) of Chapter VII:

$$\Delta E_0 \simeq \frac{\hbar^2}{m_e a_0^2} \tag{21}$$

Relation (20) therefore yields, with (13) taken into account:

$$\frac{\Delta E_2}{\Delta E_1} \simeq \frac{\Delta E_1}{\Delta E_0} \tag{22}$$

which, according to (17), proves (18).

Therefore, the effects of the magnetic field always remain, in practice, much smaller than those due to the internal field of the atom. Moreover, it is generally sufficient, when we study them, to retain only the term H_1, compared to which H_2 is negligible (H_2 will be taken into account only in the special cases in which the contribution of H_1 is zero)[3].

[3]The Zeeman effect of a three-dimensional harmonic oscillator can be calculated without approximations (*cf.* problem 2 of Complement G_{VII}). This is true because $V(\mathbf{R})$ and H_2 then have analogous forms. This example is interesting since it enables us to analyze the contributions of H_1 and H_2 in a soluble case.

1-c. Interpretation of the paramagnetic term

Consider, first of all, the term H_1 given by (7b). We shall see that it can be interpreted to be the coupling energy $-\mathbf{M}_1 \cdot \mathbf{B}$ of the field $\mathbf{B}$ and the magnetic moment $\mathbf{M}_1$ related to the revolution of the electron in its orbit.

For this purpose, we shall begin by calculating the magnetic moment $\boldsymbol{\mathcal{M}}$ classically associated with a charge q in a circular orbit of radius r (Fig. 1). If the speed of the particle is v, its motion is equivalent to a current:

$$i = q\frac{v}{2\pi r} \tag{23}$$

Since the surface S defined by this current is:

$$S = \pi r^2 \tag{24}$$

the magnetic moment $\boldsymbol{\mathcal{M}}$ is given by:

$$|\boldsymbol{\mathcal{M}}| = i \times S = \frac{q}{2} rv \tag{25}$$

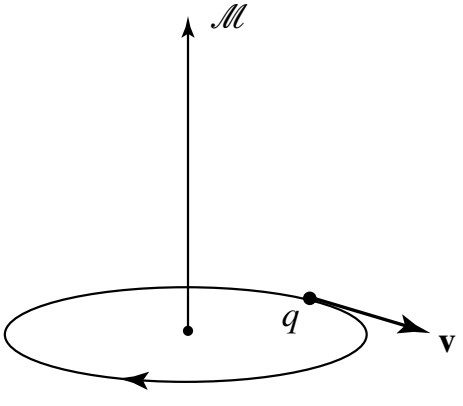


Figure 1: Classically, the motion of an electron in its orbit can be regarded as a current loop of magnetic moment $\boldsymbol{\mathcal{M}}$.

Introducing the angular momentum $\boldsymbol{\mathcal{L}}$, which, since the velocity is tangential, has a modulus of:

$$|\boldsymbol{\mathcal{L}}| = m_e\, rv \tag{26}$$

we can write (25) in the form:

$$\boldsymbol{\mathcal{M}} = \frac{q}{2m_e}\boldsymbol{\mathcal{L}} \tag{27}$$

(this is a vector relation since $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{M}}$ are parallel, as both are perpendicular to the plane of the classical orbit).

The quantum mechanical analogue of (27) is the operator relation:

$$\mathbf{M}_1 = \frac{q}{2m_e}\,\boldsymbol{L} \tag{28}$$

We can therefore write H_1 in the form:

$$H_1 = -\mathbf{M}_1 \cdot \mathbf{B} \tag{29}$$

This confirms the interpretation given above: H_1 corresponds to the coupling between the magnetic field $\mathbf{B}$ and the permanent atomic magnetic moment ($\mathbf{M}_1$ is independent of $\mathbf{B}$). H_1 is called the *paramagnetic coupling term.*

Comments:

(*i*) According to (28), the eigenvalues of any component of the magnetic moment $\mathbf{M}_1$ are of the form:

$$\left(\frac{q}{2m_e}\right) \times (m\hbar) = m\mu_B \tag{30}$$

where m is an integer. μ_B therefore gives the order of magnitude of the magnetic moment associated with the orbital moment of the electron. This is why definition (8) is useful. In the MKSA system:

$$\mu_B \simeq -9.27 \times 10^{-24} \text{ Joule/tesla} \tag{31}$$

(*ii*) As we shall see in Chapter IX, the electron possesses, in addition to the orbital angular momentum $\mathbf{L}$, a spin angular momentum $\mathbf{S}$. With this observable is associated a magnetic moment, $\mathbf{M}_S$, proportional to $\mathbf{S}$:

$$\mathbf{M}_S = 2\frac{\mu_B}{\hbar}\mathbf{S} \tag{32}$$

Although the magnetic effects due to the spin are important, we shall ignore them for now (we shall return to them in Complement D_{XII}).

(*iii*) The classical argument presented above is not completely correct. We have confused the angular momentum:

$$\boldsymbol{\mathcal{L}} = \mathbf{r} \times \mathbf{p} \tag{33}$$

with the moment of the mechanical momentum:

$$\boldsymbol{\lambda} = \mathbf{r} \times m_e\mathbf{v} = \boldsymbol{\mathcal{L}} - q\mathbf{r} \times \mathbf{A}(\mathbf{r}) \tag{34}$$

In fact, the error is small. As we shall see in the next section, it simply amounts to neglecting H_2 relative to H_1.

1-d. Interpretation of the diamagnetic term

Consider a zero angular momentum state of the hydrogen atom (for example, the ground state). The correction supplied by H_1 to the energy of this state is also zero. Thus, to determine the effect of the field $\mathbf{B}$, we must now take into account the presence of H_2. How should the corresponding energy be interpreted?

We have seen (*cf.* Complement C_{VII}, § 2-b) that, in the presence of a uniform magnetic field, the probability current associated with the electron is modified. This current is cylindrically symmetric with respect to $\mathbf{B}$. It corresponds to a uniform rotation of the probability fluid, clockwise if q is positive and counterclockwise if q is negative. With the corresponding electric current is then associated a magnetic moment $\langle \mathbf{M}_2 \rangle$ antiparallel to $\mathbf{B}$, and, therefore, a positive coupling energy, which explains the physical origin of the term H_2.

To see this more precisely, we shall return to the classical argument of the preceding section, taking into account the fact [*cf.* comment (*iii*) of § 1-c] that the magnetic moment $\boldsymbol{\mathcal{M}}$ is in fact proportional to $\boldsymbol{\lambda} = \mathbf{r} \times m_e \mathbf{v}$ (and not to $\boldsymbol{\mathcal{L}} = \mathbf{r} \times \mathbf{p}$):

$$\boldsymbol{\mathcal{M}} = \frac{q}{2m_e} \boldsymbol{\lambda} = \frac{q}{2m_e} [\boldsymbol{\mathcal{L}} - q\mathbf{r} \times \mathbf{A}(\mathbf{r})] \tag{35}$$

When $\boldsymbol{\mathcal{L}}$ is zero, $\boldsymbol{\mathcal{M}}$ reduces, in gauge (2), to:

$$\boldsymbol{\mathcal{M}}_2 = \frac{q^2}{4m_e} \, \mathbf{r} \times (\mathbf{r} \times \mathbf{B}) = \frac{q^2}{4m_e} \Big[(\mathbf{r} \cdot \mathbf{B}) \mathbf{r} - \mathbf{r}^2 \mathbf{B} \Big] \tag{36}$$

$\boldsymbol{\mathcal{M}}_2$ is proportional to the value of the magnetic field[4]. It therefore represents the *moment induced* by $\mathbf{B}$ in the atom. Its coupling energy with $\mathbf{B}$ is:

$$\begin{aligned} W_2 &= -\int_{\mathbf{0}}^{\mathbf{B}} \boldsymbol{\mathcal{M}}_2(\mathbf{B}') \cdot \mathrm{d}\mathbf{B}' = -\frac{1}{2} \boldsymbol{\mathcal{M}}_2(\mathbf{B}) \cdot \mathbf{B} \\ &= \frac{q^2}{8m_e} \left[\mathbf{r}^2 \mathbf{B}^2 - (\mathbf{r}.\mathbf{B})^2 \right] \\ &= \frac{q^2}{8m_e} \mathbf{r}_\perp^2 \mathbf{B}^2 \end{aligned} \tag{37}$$

as we found in (7c). Therefore, the interpretation given above has been confirmed: H_2 describes the coupling between the field $\mathbf{B}$ and the magnetic moment $\mathbf{M}_2$ induced in the atom. Since the induced moment, according to Lenz' law, opposes the applied field, the coupling energy is positive. H_2 is called the *diamagnetic term* of the Hamiltonian.

Comment:

As we have already pointed out [*cf.* (18)], atomic diamagnetism is a weak phenomenon which is concealed by paramagnetism when both are present. As is shown by (37) (and the calculations of § 1-b), this result is related to the small size of the atomic radius: for magnetic fields of the type usually produced, the magnetic flux intercepted by an atom is very small. It must not be concluded that we can always neglect H_2 relative to H_1, whatever the physical problem. For example, in the case of a free electron (for which the radius of the classical orbit would be infinite in a zero magnetic field), we saw in Complement E_{VI} that the contribution of the diamagnetic term is as important as that of the paramagnetic term.

[4] $\boldsymbol{\mathcal{M}}_2$ is not collinear with $\mathbf{B}$. However, it can be shown that, in the ground state of the hydrogen atom, the mean value $\langle \mathbf{M}_2 \rangle$ of the operator associated with $\boldsymbol{\mathcal{M}}_2$ is antiparallel to $\mathbf{B}$. This is in agreement with the result obtained above from the structure of the probability current.

2. The Zeeman effect

Now that we have explained the physical significance of the various terms appearing in the Hamiltonian, we shall look more closely at their effects on the spectrum of the hydrogen atom. More precisely, we shall examine the way in which the emission of the optical line called the "resonance line" ($\lambda \simeq 1\ 200$ Å) is modified when the hydrogen atom is placed in a static magnetic field. We shall see that this changes not only the frequency, but also the polarization, of the atomic lines: this is what is usually called the "Zeeman effect".

Important comment: In reality, because of the existence of electron and proton spins, the resonance line of hydrogen includes several neighboring components (fine and hyperfine structure; *cf.* Chap. XII). Moreover, the spin degrees of freedom profoundly modify the effect of a magnetic field on the various components of the resonance line (the Zeeman effect of the hydrogen atom is sometimes called "anomalous"). Since we are ignoring the effects of spin here, the following calculations do not truly correspond to the real physical situation. However, they can easily be generalized to take spins into account (*cf.* Complement D_{XII}). Moreover, the results we shall obtain (the appearance of several Zeeman components of different frequencies and polarizations) remain qualitatively valid.

2-a. Energy levels of the atom in the presence of a magnetic field

The resonance line of hydrogen corresponds to an atomic transition between the ground state $1s$ ($n = 1; l = m = 0$) and the excited state $2p$ ($n = 2; l = 1; m = +1, 0, -1$). While the angular momentum is zero in the ground state, it is not so in the excited state; in calculating optical line modifications in the presence of the magnetic field $\mathbf{B}$, we therefore make a small error by neglecting the effects of the diamagnetic term H_2, which amounts to taking $H_0 + H_1$ for the Hamiltonian.

We denote by $|\varphi_{n,l,m}\rangle$ the common eigenstates of H_0 (eigenvalue $E_n = -E_I/n^2$), $\mathbf{L}^2$ [eigenvalue $l(l+1)\hbar^2$] and L_z (eigenvalue $m\hbar$). The wave functions of these states have been calculated in Chapter VII:

$$\varphi_{n,l,m}(r, \theta, \varphi) = R_{n,l}(r)\, Y_l^m(\theta, \varphi) \tag{38}$$

We choose the Oz axis parallel to $\mathbf{B}$; it is not difficult to see that the states $|\varphi_{n,l,m}\rangle$ are then also eigenvectors of $H_0 + H_1$:

$$\begin{aligned}(H_0 + H_1)|\varphi_{n,l,m}\rangle &= \left(H_0 - \frac{\mu_B}{\hbar} B\, L_z\right) |\varphi_{n,l,m}\rangle \\ &= (E_n - m\mu_B B)|\varphi_{n,l,m}\rangle \end{aligned} \tag{39}$$

If we neglect the diamagnetic term, the stationary states of the atom placed in the field $\mathbf{B}$ are therefore still the $|\varphi_{n,l,m}\rangle$; only the corresponding energies are modified.

In particular, for the states involved in the resonance line, we see that:

$$(H_0 + H_1)|\varphi_{1,0,0}\rangle = -E_I|\varphi_{1,0,0}\rangle \tag{40a}$$

$$(H_0 + H_1)|\varphi_{2,1,m}\rangle = [-E_I + \hbar(\Omega + m\omega_L)]\, |\varphi_{2,1,m}\rangle \tag{40b}$$

where:

$$\Omega = \frac{E_2 - E_1}{\hbar} = \frac{3E_I}{4\hbar} \tag{41}$$

is the angular frequency of the resonance line in a zero field.

2-b. Electric dipole oscillations

α. Matrix elements of the operator associated with the dipole

Let:

$$\mathbf{D} = q\,\mathbf{R} \tag{42}$$

be the electric dipole operator of the atom. To calculate the mean value $\langle \mathbf{D} \rangle$ of this dipole, we begin by evaluating the matrix elements of $\mathbf{D}$.

Under reflection through the origin, $\mathbf{D}$ is changed into $-\mathbf{D}$: the electric dipole is therefore an odd operator (*cf.* Complement $\mathrm{F_{II}}$). Now, the states $|\varphi_{n,l,m}\rangle$ also have a well-defined parity: since their angular dependence is given by $Y_l^m(\theta,\varphi)$, their parity is $+1$ if l is even and -1 if l is odd (*cf.* Complement $\mathrm{A_{VI}}$). It follows, in particular, that:

$$\begin{cases} \langle \varphi_{1,0,0}|\mathbf{D}|\varphi_{1,0,0}\rangle = \mathbf{0} \\ \langle \varphi_{2,1,m'}|\mathbf{D}|\varphi_{2,1,m}\rangle = \mathbf{0} \end{cases} \tag{43}$$

for all m and m'.

The non-zero matrix elements of $\mathbf{D}$ are therefore necessarily non-diagonal elements. To calculate the matrix elements $\langle \varphi_{2,1,m}|\mathbf{D}|\varphi_{1,0,0}\rangle$, it is convenient to note that x, y and z can easily be expressed in terms of the spherical harmonics:

$$\begin{cases} x = \sqrt{\dfrac{2\pi}{3}}\, r\, \left[Y_1^{-1}(\theta,\varphi) - Y_1^1(\theta,\varphi) \right] \\ y = i\sqrt{\dfrac{2\pi}{3}}\, r\, \left[Y_1^{-1}(\theta,\varphi) + Y_1^1(\theta,\varphi) \right] \\ z = \sqrt{\dfrac{4\pi}{3}}\, r\, Y_1^0(\theta,\varphi) \end{cases} \tag{44}$$

In the expressions for the desired matrix elements, we therefore have:

– on the one hand, a radial integral, which we shall set equal to χ:

$$\chi = \int_0^\infty R_{2,1}(r)\, R_{1,0}(r)\, r^3\, \mathrm{d}r \tag{45}$$

– and on the other hand, an angular integral that, thanks to relations (44), reduces to a scalar product of spherical harmonics, which can be calculated directly from their orthogonality relations. We obtain, finally:

$$\begin{cases} \langle \varphi_{2,1,1}|D_x|\varphi_{1,0,0}\rangle = -\langle \varphi_{2,1,-1}|D_x|\varphi_{1,0,0}\rangle = -\dfrac{q\chi}{\sqrt{6}} \\ \langle \varphi_{2,1,0}|D_x|\varphi_{1,0,0}\rangle = 0 \end{cases} \tag{46a}$$

$$\begin{cases} \langle \varphi_{2,1,1}|D_y|\varphi_{1,0,0}\rangle = \langle \varphi_{2,1,-1}|D_y|\varphi_{1,0,0}\rangle = \dfrac{iq\chi}{\sqrt{6}} \\ \langle \varphi_{2,1,0}|D_y|\varphi_{1,0,0}\rangle = 0 \end{cases} \tag{46b}$$

$$\begin{cases} \langle \varphi_{2,1,1}|D_z|\varphi_{1,0,0}\rangle = \langle \varphi_{2,1,-1}|D_z|\varphi_{1,0,0}\rangle = 0 \\ \langle \varphi_{2,1,0}|D_z|\varphi_{1,0,0}\rangle = \dfrac{q\chi}{\sqrt{3}} \end{cases} \tag{46c}$$

β. Calculation of the mean value of the dipole

The results of § 2-b-α indicate that, if the system is in a stationary state, the mean value of the operator **D** is zero. Let us assume, rather, that the state vector of the system is initially a linear superposition of the ground state $1s$ and one of the $2p$ states:

$$|\psi_m(0)\rangle = \cos\alpha\ |\varphi_{1,0,0}\rangle + \sin\alpha\ |\varphi_{2,1,m}\rangle \tag{47}$$

with $m = +1$, 0 or -1 (α is a real parameter). We then immediately obtain the state vector at time t:

$$|\psi_m(t)\rangle = \cos\alpha\ |\varphi_{1,0,0}\rangle + \sin\alpha\ \mathrm{e}^{-i(\Omega+m\omega_L)t}\ |\varphi_{2,1,m}\rangle \tag{48}$$

(we have omitted the global phase factor $\mathrm{e}^{iE_I t/\hbar}$, which is of no physical consequence).

To calculate the mean value of the electric dipole:

$$\langle \mathbf{D}\rangle_m(t) = \langle\psi_m(t)|\mathbf{D}|\psi_m(t)\rangle \tag{49}$$

we shall use results (46) and (48), and cite three cases:

(*i*) if $m = 1$, we obtain:

$$\begin{cases} \langle D_x\rangle_1 = -\dfrac{q\chi}{\sqrt{6}}\ \sin 2\alpha\ \cos\left[(\Omega+\omega_L)t\right] \\ \langle D_y\rangle_1 = -\dfrac{q\chi}{\sqrt{6}}\ \sin 2\alpha\ \sin\left[(\Omega+\omega_L)t\right] \\ \langle D_z\rangle_1 = 0 \end{cases} \tag{50}$$

The vector $\langle \mathbf{D}\rangle_1(t)$ therefore rotates in the xOy plane about the Oz axis, in the counterclockwise direction and with the angular velocity $\Omega + \omega_L$.

(*ii*) if $m = 0$:

$$\begin{cases} \langle D_x\rangle_0 = \langle D_y\rangle_0 = 0 \\ \langle D_z\rangle_0 = \dfrac{q\chi}{\sqrt{3}} \sin 2\alpha\ \cos\Omega t \end{cases} \tag{51}$$

The motion of $\langle \mathbf{D}\rangle_0(t)$ is now a linear oscillation along the Oz axis, of angular frequency Ω.

(*iii*) if $m = -1$:

$$\begin{cases} \langle D_x\rangle_{-1} = \dfrac{q\chi}{\sqrt{6}}\ \sin 2\alpha\ \cos\left[(\Omega-\omega_L)t\right] \\ \langle D_y\rangle_{-1} = -\dfrac{q\chi}{\sqrt{6}}\ \sin 2\alpha\ \sin\left[(\Omega-\omega_L)t\right] \\ \langle D_z\rangle_{-1} = 0 \end{cases} \tag{52}$$

The vector $\langle \mathbf{D}\rangle_{-1}(t)$ again rotates in the xOy plane about Oz, but this time in the clockwise direction and with the angular velocity $\Omega - \omega_L$.

2-c. Frequency and polarization of emitted radiation

In the three cases ($m = +1$, 0 and -1), the mean value of the electric dipole is an oscillating function of time. It is clear that such a dipole radiates electromagnetic energy.

Since the atomic dimensions are negligible compared to the optical wavelength, the atom's radiation at great distances can be treated like that of a dipole. We shall assume that the characteristics of the light emitted (or absorbed) by the atom during transition between a state $|\varphi_{2,1,m}\rangle$ and the ground state are correctly given by the classical calculation of radiation of a dipole[5] equal to the mean quantum mechanical value $\langle \mathbf{D} \rangle_m(t)$.

To state the problem precisely, imagine we want to study the radiation emitted by a sample containing a great number of hydrogen atoms, which have somehow been excited into a $2p$ state. In most experiments actually performed, the excitation of the atoms is isotropic, and the three states $|\varphi_{2,1,1}\rangle$, $|\varphi_{2,1,0}\rangle$ and $|\varphi_{2,1,-1}\rangle$ occur with equal probability. Therefore, we shall begin by calculating the radiation diagram for each of the cases cited in the preceding sections. Then we shall obtain the radiation actually emitted by the atomic system by taking, for each spatial direction, the sum of the light intensities emitted in each case.

(*i*) If $m = 1$, the angular frequency of the emitted radiation is $(\Omega+\omega_L)$. The optical line frequency is therefore slightly shifted by the magnetic field. In accordance with the laws of classical electromagnetism applied to a rotating dipole such as $\langle \mathbf{D} \rangle_1(t)$, the radiation emitted in the Oz direction is circularly polarized (the corresponding polarization is called σ_+). However, the radiation emitted in a direction of the xOy plane is linearly polarized (parallel to this plane). In other directions, the polarization is elliptical.

(*ii*) If $m = 0$, we must consider a dipole oscillating linearly along Oz, with angular frequency Ω, that is, the same as in a zero field. The wavelength of the radiation is therefore not changed by the field $\mathbf{B}$. Its polarization is always linear, whatever the propagation direction being considered. For example, for a propagation direction situated in the xOy plane, this polarization is parallel to Oz (π polarization). No radiation is emitted in the Oz direction (an oscillating linear dipole does not radiate along its axis).

(*iii*) If $m = -1$, the results are analogous to those for $m = 1$. The only difference is that the angular frequency of the radiation is $(\Omega - \omega_L)$ instead of $(\Omega + \omega_L)$, and the dipole rotates in the opposite direction; this changes, for example, the direction of the circular polarization (σ_- polarization).

.

If we now assume that there are equal numbers of excited atoms in the three states $m = +1$, 0 and -1, we see that:

- in an arbitrary spatial direction, three optical frequencies are emitted: $\Omega/2\pi$, $(\Omega \pm \omega_L)/2\pi$. The polarization associated with the first one is linear, and that associated with the others is, in general, elliptical;

- in a direction perpendicular to the field $\mathbf{B}$, the three polarizations are linear (*cf.* Fig. 2). The first one is parallel to $\mathbf{B}$, and the other two are perpendicular. The intensity of the central line is twice that of each of the shifted lines [*cf.* formulas

[5] If we wanted to treat the problem entirely quantum mechanically, we should have to use the quantum mechanical theory of radiation. In particular, the return of the atom to the ground state by spontaneous emission of a photon could only be understood in the framework of this theory. However, the results we shall obtain here semi-classically would remain essentially valid as far as radiation is concerned.

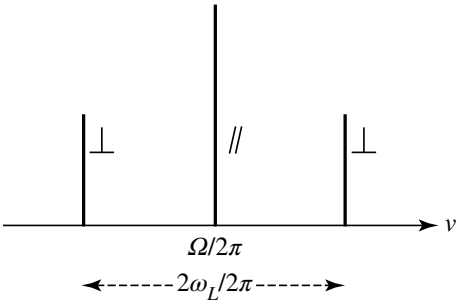


Figure 2: The Zeeman components of the resonance line of hydrogen observed in a direction perpendicular to the magnetic field **B** *(ignoring electron spin). We obtain a component of unshifted frequency* $\Omega/2\pi$*, polarized parallel to* **B***, and two components shifted by* $\pm\omega_L/2\pi$*, polarized perpendicularly to* **B***.*

(50), (51) and (52)]. In a direction parallel to **B**, only the two shifted frequencies $(\Omega \pm \omega_L)/2\pi$ are emitted, and the associated light polarizations are both circular but opposite in direction (*cf.* Fig. 3).

Comment:

The atom therefore emits σ_+ polarized radiation in going from the state $|\varphi_{2,1,1}\rangle$ to the state $|\varphi_{1,0,0}\rangle$, σ_- in going from $|\varphi_{2,1,-1}\rangle$ to $|\varphi_{1,0,0}\rangle$, and π in going from $|\varphi_{2,1,0}\rangle$ to $|\varphi_{1,0,0}\rangle$. Formulas (46) furnish a simple rule for finding these polarizations. Consider the operators $D_x + iD_y$, $D_x - iD_y$ and D_z; their only non-zero matrix elements between the $2p$ and $1s$ states taken in this order are:

$$\langle\varphi_{2,1,1}|D_x + iD_y|\varphi_{1,0,0}\rangle, \quad \langle\varphi_{2,1,-1}|D_x - iD_y|\varphi_{1,0,0}\rangle$$
$$\text{and} \quad \langle\varphi_{2,1,0}|D_z|\varphi_{1,0,0}\rangle$$

To the σ_+, σ_- and π polarizations, therefore, correspond the operators $D_x + iD_y$, $D_x - iD_y$ and D_z, respectively. This is a general rule: there is emission of electric dipole radiation when the operator **D** has a non-zero matrix element between the atom's initial state and its final state. The polarization of this radiation is σ_+, σ_- or π depending on whether the non-zero matrix element[6] is that of $D_x + iD_y$, $D_x - iD_y$ or D_z.

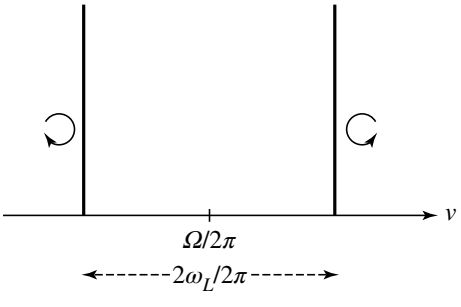


Figure 3: When the observation is carried out along the direction of the field **B***, only two Zeeman components are obtained, circularly polarized in opposite directions and shifted by* $\pm\omega_L/2\pi$*.*

[6]The order of the states in the matrix element must be respected in order not to confuse σ_+ with σ_-.

References and suggestions for further reading:

Paramagnetism and diamagnetism: Feynman II (7.2), Chaps. 34 and 35; Cagnac and Pebay-Peyroula (11.2), Chaps. VIII and IX; Kittel (13.2), Chap. 14; Slater (1.6), Chap. 14; Flügge (1.24), §§ 128 and 160.

Dipole radiation: Cagnac and Pebay-Peyroula (11.2), Annex III; Panofsky and Phillips (7.6), § 14-7; Jackson (7.5), § 9-2.

Angular momentum of radiation and selection rules: Cagnac and Pebay-Peyroula (11.2), Chap. XI.

Complement E_{VII}

Some atomic orbitals. Hybrid orbitals

1 Introduction 869
2 Atomic orbitals associated with real wave functions 870
2-a s orbitals ($l = 0$) 870
2-b p orbitals ($l = 1$) 870
2-c Other values of l 874
3 sp hybridization 876
3-a Introduction of sp hybrid orbitals 876
3-b Properties of sp hybrid orbitals 876
3-c Example: the structure of acetylene 878
4 sp^2 hybridization 878
4-a Introduction of sp^2 hybrid orbitals 878
4-b Properties of sp^2 hybrid orbitals 880
4-c Example: the structure of ethylene 881
5 sp^3 hybridization 882
5-a Introduction of sp^3 hybrid orbitals 882
5-b Properties of sp^3 hybrid orbitals 882
5-c Example: The structure of methane 883

1. Introduction

In § C of Chapter VII, we determined an orthonormal basis of stationary states for the electron of the hydrogen atom. The corresponding wave functions are:

$$\varphi_{n,l,m}(\mathbf{r}) = R_{n,l}(r)\ Y_l^m(\theta, \varphi) \tag{1}$$

and the quantum numbers n, l, m refer, respectively, to the energy $E_n = -E_I/n^2$, the square of the angular momentum $l(l+1)\hbar^2$, and the Oz-component of the angular momentum $m\hbar$.

By linearly superposing stationary states of the same energy, that is, of the same quantum number n, we can construct new stationary states which no longer necessarily correspond to well-defined values of l and m. In this complement, we intend to study the properties of some of these new stationary states – in particular, the angular dependence of the associated wave functions.

The wave functions (1) are often called *atomic orbitals*. A linear superposition of orbitals of the same n but different l and m is called a *hybrid orbital*. We shall see that a hybrid orbital can extend further in certain spatial directions than the (pure) orbitals from which it is constructed. It is this property, important in the formation of chemical bonds, which justifies the introduction of hybrid orbitals.

Although the calculations presented in this complement are rigorously valid only for the hydrogen atom, we shall also indicate qualitatively how they explain the geometrical structure of the various bonds formed by an atom with several valence electrons.

2. Atomic orbitals associated with real wave functions

In expression (1), the radial function $R_{n,l}(r)$ is real. However, $Y_l^m(\theta,\varphi)$, except for $m=0$, is a complex function of φ, since:

$$Y_l^m(\theta,\varphi) = F_l^m(\theta)\ \mathrm{e}^{im\varphi} \tag{2}$$

where $F_l^m(\theta)$ is a real function of θ.

Atomic orbitals are therefore generally complex functions. By superposing the $\varphi_{n,l,m}(\mathbf{r})$ and $\varphi_{n,l,-m}(\mathbf{r})$ orbitals, one can, however, construct real orbitals whose advantage is their simple angular dependence, which can be represented graphically without having to take the square of the modulus of the wave function (as we did in § C-4-c-α of Chapter VII).

2-a. s orbitals ($l=0$)

When $l=m=0$, the wave function $\varphi_{n,0,0}(\mathbf{r})$ is real, and we say we are dealing with an "s orbital". We shall denote the corresponding stationary state by $|\varphi_{ns}\rangle$. To represent the angular dependence of the ns orbital, we fix r and measure off in each direction of polar angles θ and φ a line segment of length $\varphi_{ns}(r,\theta,\varphi)$. The surface obtained by varying θ and φ is a sphere centered at O (Fig. 1).

2-b. p orbitals ($l=1$)

α. p_z, p_x, p_y orbitals

If we use the expression for the three spherical harmonics $Y_l^m(\theta,\varphi)$ given in Complement A_{VI} [formulas (32)], we obtain, for the three atomic orbitals $\varphi_{n,1,m}(\mathbf{r})$ corresponding to $l=1$:

$$\begin{cases} \varphi_{n,1,1}(\mathbf{r}) = -\sqrt{\dfrac{3}{8\pi}}\, R_{n,1}(r)\sin\theta\ \mathrm{e}^{i\varphi} & \text{(3a)} \\ \varphi_{n,1,0}(\mathbf{r}) = \sqrt{\dfrac{3}{4\pi}}\, R_{n,1}(r)\cos\theta & \text{(3b)} \\ \varphi_{n,1,-1}(\mathbf{r}) = \sqrt{\dfrac{3}{8\pi}}\, R_{n,1}(r)\sin\theta\ \mathrm{e}^{-i\varphi} & \text{(3c)} \end{cases}$$

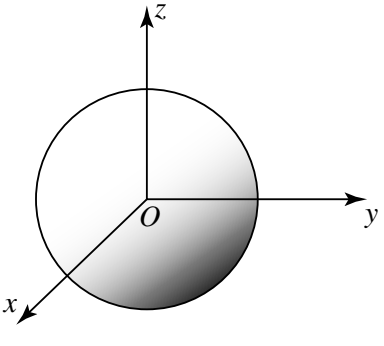


Figure 1: An s orbital is spherically symmetric: the wave function depends on neither θ nor φ.

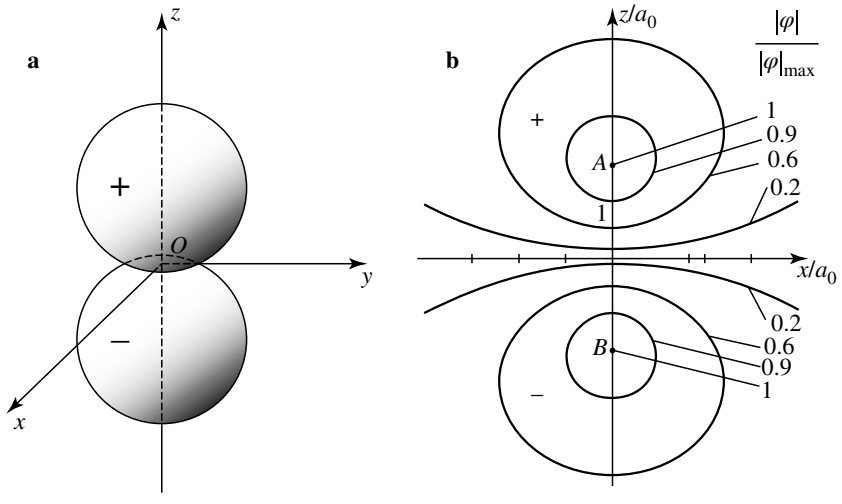


Figure 2: Two possible representations of a p_z orbital ($l = 1, m = 0$)
fig. a: angular dependence of this orbital. With r fixed, we display $|\varphi_{n,l=1,m=0}(r, \theta, \varphi)|$ for each direction θ, φ. Thus we obtain two spheres, tangential at O to the xOy plane. The sign indicated on each of them is that of the wave function (which is real).
fig. b: cross sections in the xOz plane of a family of surfaces, each one corresponding to a given value for $|\varphi_{n,l=1,m=0}(r, \theta, \varphi)|$ [we have chosen values equal to 0.2, 0.6, 0.9 times the maximum value of $|\varphi|$ at points A and B]. These are surfaces of revolution about Oz. The sign indicated is that of the wave function (which is real). Unlike the one in figure a, the representation in figure b depends on the radial part of the wave function (the one chosen here corresponds to the $n = 2$ state of the hydrogen atom).

We now form the three linear superpositions:

$$\begin{cases} \varphi_{n,1,0}(\mathbf{r}) & \text{(4a)} \\ -\dfrac{1}{\sqrt{2}}[\varphi_{n,1,1}(\mathbf{r}) - \varphi_{n,1,-1}(\mathbf{r})] & \text{(4b)} \\ \dfrac{i}{\sqrt{2}}[\varphi_{n,1,1}(\mathbf{r}) + \varphi_{n,1,-1}(\mathbf{r})] & \text{(4c)} \end{cases}$$

It is easy to see that the three preceding wave functions can also be written:

$$\begin{cases} \sqrt{\dfrac{3}{4\pi}}\, R_{n,1}(r)\, \dfrac{z}{r} & \text{(5a)} \\ \sqrt{\dfrac{3}{4\pi}}\, R_{n,1}(r)\, \dfrac{x}{r} & \text{(5b)} \\ \sqrt{\dfrac{3}{4\pi}}\, R_{n,1}(r)\, \dfrac{y}{r} & \text{(5c)} \end{cases}$$

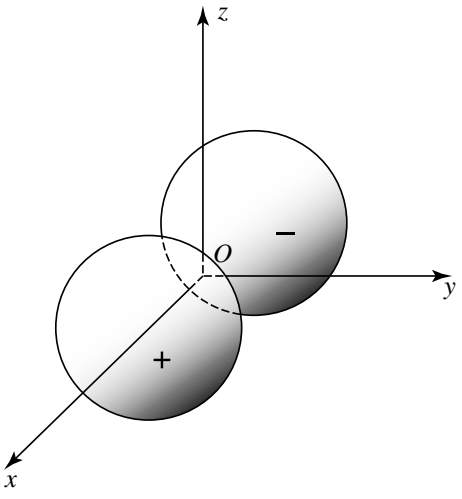


Figure 3: Angular dependence of a p_x orbital (the representation adopted is that of Figure 2-a).

These are real functions of r, θ, φ which, like the $\varphi_{n,1,m}(\mathbf{r})$, are orthonormal and form a basis in the subspace $\mathcal{E}_{n,l=1}$. The wave functions (5) will be denoted $\varphi_{np_z}(\mathbf{r})$, $\varphi_{np_x}(\mathbf{r})$ and $\varphi_{np_y}(\mathbf{r})$ respectively, and will be called "p_z, p_x, p_y orbitals".

Two distinct geometrical representations enable us to visualize the form of an orbital $\psi(r, \theta, \varphi)$. First of all, if we are interested in the angular dependence of the orbital, we fix r and measure off a line segment of length $|\psi(r, \theta, \varphi)|$ along each direction of polar angles θ and φ. Thus, the angular dependence of the $2p_z$ orbital is that of $z/r = \cos\theta$. As φ varies between 0 and 2π, and θ varies between 0 and π, the end of the line segment of length $|\cos\theta|$ drawn in the direction of polar angles θ and φ describes two spheres centered on the Oz axis, tangential at O to the xOy plane and mirror images with respect to the xOy plane (Fig. 2-a). The sign indicated in the figure is that of the wave function, which is real. Another possible representation of the orbital $\psi(r, \theta, \varphi)$ is obtained by tracing a family of surfaces, each one corresponding to a given value of $|\psi(r, \theta, \varphi)|$ (surfaces of equal probability density). This is what is done for the $2p_z$ orbital in Figure 2-b (here again, the sign indicated is that of the wave function, which is real). In the rest of this complement, we shall use one or the other of these two representations.

The p_x and p_y orbitals can be obtained respectively from the p_z orbitals by rotations through angles of $+\pi/2$ and $-\pi/2$ about Oy and Ox (*cf.* Figures 3 and 4, which use a geometric representation identical to the one in Figure 2-a).

Unlike an s orbital, which is spherically symmetric, the p_z, p_x, p_y orbitals therefore point along the Oz, Ox, Oy axes, respectively.

β. p_u orbitals

The choice of the Ox, Oy, Oz axes is obviously arbitrary. By linearly superposing the p_x, p_y and p_z orbitals, we should therefore be able to construct a p_u orbital having the same form but directed along an arbitrary Ou axis.

Let Ou be such an axis, forming angles α, β, γ with Ox, Oy, Oz. We have,

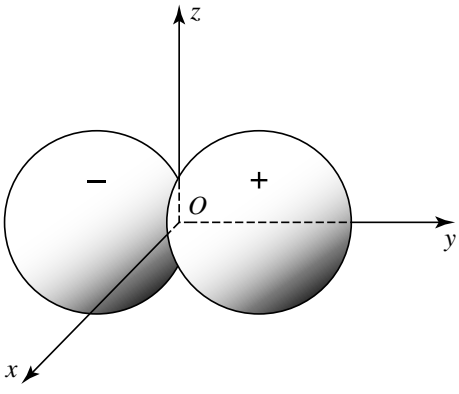


Figure 4: Angular dependence of a p_y orbital.

obviously:

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1 \tag{6}$$

Consider the state:

$$\cos\alpha\,|np_x\rangle + \cos\beta\,|np_y\rangle + \cos\gamma\,|np_z\rangle \tag{7}$$

which, according to (6), is normalized. We can, using formulas (5), put the corresponding wave function in the form:

$$\sqrt{\frac{3}{4\pi}}\,R_{n,1}(r)\frac{x\cos\alpha + y\cos\beta + z\cos\gamma}{r} = \sqrt{\frac{3}{4\pi}}\,R_{n,1}(r)\,\frac{u}{r} \tag{8}$$

where:

$$u = x\cos\alpha + y\cos\beta + z\cos\gamma \tag{9}$$

is the projection of $\mathbf{r}$ onto the Ou axis. Comparison with (5) indicates that the orbital so constructed is indeed a p_u orbital.

Therefore, any real and normalized linear superposition of p_x, p_y and p_z orbitals:

$$\lambda\;\varphi_{np_x}(\mathbf{r}) + \mu\;\varphi_{np_y}(\mathbf{r}) + \nu\;\varphi_{np_z}(\mathbf{r}) \tag{10}$$

can be considered to be a p_u orbital directed along the Ou direction, defined by:

$$\begin{cases} \cos\alpha = \lambda \\ \cos\beta = \mu \\ \cos\gamma = \nu \end{cases} \tag{11}$$

γ. *Example: structure of the H_2O and H_3N molecules*

To a first approximation (*cf.* Complement A_{XIV}), in a many-electron atom, each electron can be considered to move independently of the others in a central potential $V_c(r)$ which is the sum of the electrostatic attractive potential of the nucleus and a "mean potential" due to the repulsion of the other electrons. Each electron can therefore be found in a state characterized by

three quantum numbers, n, l, m. However, since the potential $V_c(r)$ no longer varies exactly like $1/r$, the energy no longer depends only on n, but also on l. We shall see in Complement A_{XIV} that the energy of the $2s$ state is slightly lower than that of the $2p$ state; the $3s$ state is also lower than the $3p$ state, which is, in turn, lower than the $3d$ state, etc.

The existence of spin and Pauli's principle (which we shall study in Chapters IX and XIV) imply that the $1s$, $2s$, ... sub-shells can contain only two electrons; the $2p$, $3p$, ... sub-shells, six electrons, ...; the nl sub-shells, $2(2l+1)$ electrons (the factor $2l+1$ arises from the degeneracy related to L_z, and the factor 2, from the electron spin).

Thus, for the oxygen atom, which has eight electrons, the $1s$ and $2s$ sub-shells are filled and contain four electrons in all. The four remaining electrons are in the $2p$ sub-shell: two of them (with opposite spins) can fill one of the three $2p$ orbitals, for example, $2p_z$; the other two are then distributed in the remaining $2p_x$ and $2p_y$ orbitals. These last two electrons are the valence electrons: they are "unpaired", which means that the orbitals in which they can be found can accept another electron. The $2p_x$ and $2p_y$ wave functions of the valence electrons of oxygen are therefore directed along two perpendicular axes. Now, it can be shown that, the greater the overlapping of the wave functions of the two electrons participating in a chemical bond, the greater the stability of this bond. The two hydrogen atoms that will bind with the oxygen atom to form a water molecule must therefore be centered respectively on the Ox and Oy axes. Then the spherical $1s$ orbital of the valence electron of each hydrogen will maximally overlap one of the $2p_x$ and $2p_y$ orbitals of the valence electrons of oxygen. Figure 5 represents the shape of the probability clouds associated with the valence electrons of the oxygen and hydrogen atoms in the water molecule. The graphical representation used is analogous to the one in Figure 2-b. We have drawn, for each electron, a surface defined as follows: the probability density has the same value at all points of this surface; this value is chosen in such a way that the total probability contained inside the surface has a fixed value close to 1 (0.9, for example).

The preceding argument enables us to understand the form of the H_2O molecule. The angle between the two OH bonds should be close to 90°. Actually, the angle found experimentally is 104°. The deviation from the value 90° arises, in part, from the electrostatic repulsion between the two protons of the hydrogen atoms, which tends to open up the angle between the two OH bonds[1].

An analogous argument explains the pyramidal form of the NH_3 molecule. The three valence electrons of nitrogen occupy the $2p_x$, $2p_y$, $2p_z$ orbitals, directed at right angles to each other. Here again, the electrostatic repulsion between the protons of the three hydrogen atoms causes the bond angle to go from 90° to 108° (through a slight hybridization of the $2s$ and $2p$ orbitals).

2-c. Other values of l

We have confined ourselves thus far to the s and p orbitals. Actually, an orthonormal basis of real orbitals can be constructed for each value of l. If we note that [*cf.* relation (D-29) of Chapter VI]:

$$[Y_l^m(\theta,\varphi)]^* = (-1)^m Y_l^{-m}(\theta,\varphi) \tag{12}$$

[1] The opening up of the angle between the two OH bonds can be described as the result of a slight sp^3 hybridization of the $2p$ and $2s$ orbitals (*cf.* § 5).

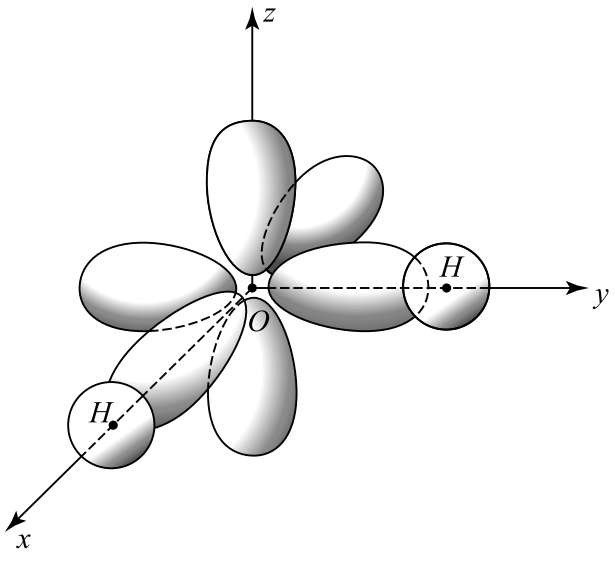


Figure 5: Schematic structure of the water molecule H_2O. The $2p_x$ and $2p_y$ orbitals yield bonds making an angle of approximately 90° (the real angle is 104° because of the electrostatic repulsion between the two protons).

---

we immediately see that (for $m \neq 0$) the two complex functions $\varphi_{n,l,m}(\mathbf{r})$ and $\varphi_{n,l,-m}(\mathbf{r})$ can be replaced by the two functions:

$$\frac{1}{\sqrt{2}}\left[\varphi_{n,l,m}(\mathbf{r}) + (-1)^m \varphi_{n,l,-m}(\mathbf{r})\right] \tag{13a}$$

$$\frac{i}{\sqrt{2}}\left[\varphi_{n,l,m}(\mathbf{r}) - (-1)^m \varphi_{n,l,-m}(\mathbf{r})\right] \tag{13b}$$

which are real and orthonormal.

Thus, for $l = 2$ ("d orbitals"), we can construct five real orbitals, for which the angular dependence is given by:

$$\sqrt{\frac{1}{2}}\left(3\cos^2\theta - 1\right),\ \sqrt{6}\sin\theta\cos\theta\cos\varphi,\ \sqrt{6}\sin\theta\cos\theta\sin\varphi,$$
$$\sqrt{\frac{3}{2}}\sin^2\theta\cos 2\varphi,\ \sqrt{\frac{3}{2}}\sin^2\theta\sin 2\varphi$$

($d_{3z^2-r^2}$, d_{zx}, d_{zy}, $d_{x^2-y^2}$, d_{xy} orbitals).

The form of these orbitals is a little more complicated than that of the s and p orbitals to which we shall confine ourselves here. However, it is possible to apply to them arguments of the same type as those below.

3. *sp* hybridization

3-a. Introduction of *sp* hybrid orbitals

Returning to the hydrogen atom, we shall consider the subspace $\mathcal{E}_{ns} \oplus \mathcal{E}_{np}$, subtended by the four real orbitals $\varphi_{ns}(\mathbf{r})$, $\varphi_{np_x}(\mathbf{r})$, $\varphi_{np_y}(\mathbf{r})$ and $\varphi_{np_z}(\mathbf{r})$ (which correspond to the same energy). We shall show that, by linearly superposing ns and np orbitals, we can construct other real orbitals, which form an orthonormal basis in $\mathcal{E}_{ns} \oplus \mathcal{E}_{np}$ and possess some interesting properties.

We begin by linearly superposing the $\varphi_{ns}(\mathbf{r})$ and $\varphi_{np_z}(\mathbf{r})$ orbitals alone, without using $\varphi_{np_x}(\mathbf{r})$ and $\varphi_{np_y}(\mathbf{r})$. We therefore replace the two functions $\varphi_{ns}(\mathbf{r})$ and $\varphi_{np_z}(\mathbf{r})$ by the two real orthonormal linear combination:

$$\begin{cases} \cos\alpha\,\varphi_{ns}(\mathbf{r}) + \sin\alpha\,\varphi_{np_z}(\mathbf{r}) & \text{(14a)} \\ \sin\alpha\,\varphi_{ns}(\mathbf{r}) - \cos\alpha\,\varphi_{np_z}(\mathbf{r}) & \text{(14b)} \end{cases}$$

In addition, we require the two orbitals (14a) and (14b) to have the same geometrical form. Since this form depends only on the relative amounts of the s and p orbitals in the linear superposition, we see immediately that we must have $\sin\alpha = \cos\alpha$, that is, $\alpha = \pi/4$. The two new orbitals we are introducing are therefore of the form:

$$\begin{cases} \varphi_{n,s,p_z}(\mathbf{r}) = \dfrac{1}{\sqrt{2}}\left[\varphi_{ns}(\mathbf{r}) + \varphi_{np_z}(\mathbf{r})\right] & \text{(15a)} \\ \varphi'_{n,s,p_z}(\mathbf{r}) = \dfrac{1}{\sqrt{2}}\left[\varphi_{ns}(\mathbf{r}) - \varphi_{np_z}(\mathbf{r})\right] & \text{(15b)} \end{cases}$$

and correspond to what is called "*sp* hybridization". Thus we have constructed a new orthonormal basis of $\mathcal{E}_{ns} \oplus \mathcal{E}_{np}$, composed of $\varphi_{n,s,p_z}(\mathbf{r})$, $\varphi'_{n,s,p_z}(\mathbf{r})$, $\varphi_{np_x}(\mathbf{r})$ and $\varphi_{np_y}(\mathbf{r})$.

3-b. Properties of *sp* hybrid orbitals

To study the angular dependence of the $\varphi_{n,s,p_z}(\mathbf{r})$ and $\varphi'_{n,s,p_z}(\mathbf{r})$ hybrid orbitals, we now choose a given value r_0 of r and set:

$$\begin{aligned} \lambda &= \sqrt{\frac{1}{4\pi}} R_{n,0}(r_0) \\ \mu &= \sqrt{\frac{3}{4\pi}} R_{n,1}(r_0) \end{aligned} \qquad (16)$$

Thus we obtain, using (5) and (15), the angular functions:

$$\begin{cases} \dfrac{1}{\sqrt{2}}(\lambda + \mu\cos\theta) \\ \dfrac{1}{\sqrt{2}}(\lambda - \mu\cos\theta) \end{cases} \qquad (17)$$

which we shall represent, using the same method as in § 2 (*cf.* Fig. 2-a), by measuring off, along each direction of polar angles θ and φ, a line segment of length $\dfrac{1}{\sqrt{2}}[\lambda + \mu\cos\theta]$

or $\frac{1}{\sqrt{2}}[\lambda - \mu\cos\theta]$ and indicating by a plus or minus sign whether the wave function is positive or negative. Figure 6 represents the cross sections in the xOz plane of the surfaces so obtained, which have cylindrical symmetry with respect to Oz (we have assumed $\mu > \lambda > 0$). The $\varphi_{n,s,p_z}(\mathbf{r})$ orbital can be transformed into the $\varphi'_{n,s,p_z}(\mathbf{r})$ by a reflection through the point O. It can be seen that the $\varphi_{n,s,p_z}(\mathbf{r})$ orbital has no simple symmetry with respect to the point O. This asymmetry is due to the fact that the $\varphi_{np_z}(\mathbf{r})$ and $\varphi_{ns}(\mathbf{r})$ orbitals of which it is formed (and which are shown in Figure 6-c) have opposite parities. In the region where $z > 0$, $\varphi_{ns}(\mathbf{r})$ and $\varphi_{np_z}(\mathbf{r})$ have the same sign and add, while in the region where $z < 0$, $\varphi_{ns}(\mathbf{r})$ and $\varphi_{np_z}(\mathbf{r})$ have opposite signs and subtract. The conclusions are reversed for $\varphi'_{n,s,p_z}(\mathbf{r})$.

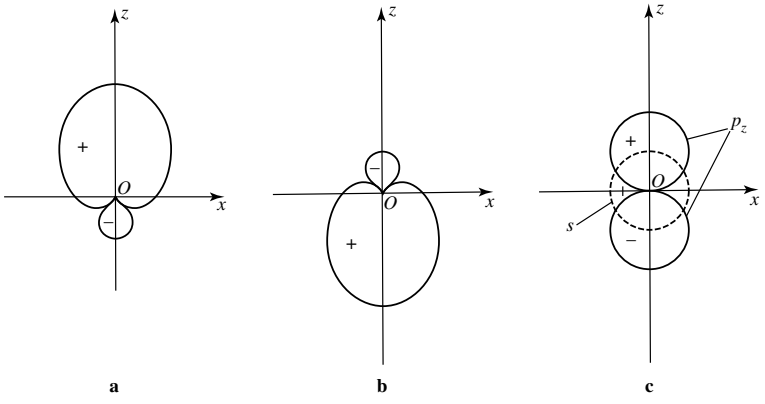


Figure 6: Angular dependence of the $\varphi_{n,s,p_z}(\mathbf{r})$ (fig. a) and $\varphi'_{n,s,p_z}(\mathbf{r})$ (fig. b), hybrid orbitals obtained from the $\varphi_{n,s}(\mathbf{r})$ and $\varphi_{n,p_z}(\mathbf{r})$ orbitals, which have opposite parities (fig. c). A hybrid orbital can extend further in certain directions than the pure orbitals from which it is obtained.

The $\varphi_{n,s,p_z}(\mathbf{r})$ orbital therefore extends further in the positive direction of the Oz axis than in the negative direction since, for fixed r, the values it takes on are greater (in absolute value) for $\theta = 0$ than for $\theta = \pi$. In general, for large values of r, λ and μ are such that the values of the $\varphi_{n,s,p_z}(\mathbf{r})$ orbital in the positive direction of the Oz axis are larger than those taken on separately by the $\varphi_{ns}(\mathbf{r})$ and $\varphi_{np_z}(\mathbf{r})$ orbitals [the same conclusions are valid for the $\varphi'_{n,s,p_z}(\mathbf{r})$ orbital and the negative direction of the Oz axis].

This property plays an important role in the study of the chemical bond. To understand this qualitatively, assume that, in a particular atom A, one of the valence electrons can be either in the ns orbital or in one of the np orbitals. Then suppose that another atom B is in the neighborhood of the first one, and call Oz the axis joining A and B. The $\varphi_{n,s,p_z}(\mathbf{r})$ orbital of A will overlap more with the orbitals of the valence electrons of B than the $\varphi_{ns}(\mathbf{r})$ or $\varphi_{np_z}(\mathbf{r})$ orbitals. Thus we see that hybridization of the orbitals of A can lead to a greater stability of the chemical bond, since this stability increases, as we have already pointed out, with the overlapping of the electronic orbitals of A and B involved in the bond.

3-c. Example: the structure of acetylene

The carbon atom has six electrons. When this atom is free, two of these electrons are in the $1s$ sub-shell, two in the $2s$ sub-shell, and two in the $2p$ sub-shell. Only these last two are unpaired, and we should therefore expect carbon to be bivalent. This is indeed what is observed in some of its compounds. However, carbon is usually present in a quadrivalent form. This arises from the fact that when a carbon atom is bound to other atoms, one of its $2s$ electrons can leave this sub-shell and place itself in the third $2p$ orbital, which is unoccupied in the free carbon atom. There are then four unpaired electrons, whose wave functions are the result of a hybridization of the four orbitals, $2s$, $2p_x$, $2p_y$ and $2p_z$.

Thus, in the acetylene molecule C_2H_2, the four valence electrons of each carbon atom are distributed as follows: two electrons are in the $\varphi_{2,s,p_z}(\mathbf{r})$ and $\varphi'_{2,s,p_z}(\mathbf{r})$ hybrid orbitals we have just introduced, and the other two are in the $\varphi_{2p_x}(\mathbf{r})$ and $\varphi_{2p_y}(\mathbf{r})$ orbitals studied in § 2-b. According to Figures 6-a and 6-b, the two electrons of each carbon atom that occupy the $\varphi_{2,s,p_z}(\mathbf{r})$ and $\varphi'_{2,s,p_z}(\mathbf{r})$ hybrid orbitals participate in bonds separated by an angle of 180°: the first one with the other carbon atom, and the second one with one of the two hydrogen atoms (whose valence electrons occupy $1s$ orbitals). Thus we understand why the C_2H_2 molecule is linear (*cf.* Fig. 7, where we have used the same type of graphical representation as in Figure 5).

As for the $2p_x$ orbitals centered on each of the carbon atoms, they present a partial lateral overlapping, as do the two $2p_y$ orbitals, as is shown by the solid lines in Figure 7. They contribute to the reinforcement of the chemical stability of the molecule. The two carbon atoms thus form a *triple bond* between them. One bond is produced by the $\varphi_{2,s,p_z}(\mathbf{r})$ and $\varphi'_{2,s,p_z}(\mathbf{r})$ hybrid orbitals, each centered on one of the two atoms and cylindrically symmetric with respect to the Oz axis (σ bond). Two bonds are produced by the $\varphi_{2p_x}(\mathbf{r})$ and $\varphi_{2p_y}(\mathbf{r})$ orbitals, which are symmetric with respect to the xOz and yOz planes (π bonds).

Comment:

As we have already pointed out, the $2p$ sub-shell, in a many-electron atom, has an energy greater than that of the $2s$ sub-shell. The passage of an electron from the $2s$ sub-shell to the $2p$ sub-shell is therefore not energetically favorable. However, the energy needed for this excitation is fully compensated by the increase in stability due to the hybrid orbitals involved in the C – H and C – C bonds.

4. sp^2 hybridization

4-a. Introduction of sp^2 hybrid orbitals

We shall now return to the four orbitals $\varphi_{ns}(\mathbf{r})$, $\varphi_{np_x}(\mathbf{r})$, $\varphi_{np_y}(\mathbf{r})$ and $\varphi_{np_z}(\mathbf{r})$, and replace the first three by the three following real combinations:

$$\begin{cases} \varphi_{n,s,p_x,p_y}(\mathbf{r}) = a\,\varphi_{ns}(\mathbf{r}) + b\,\varphi_{np_x}(\mathbf{r}) + c\,\varphi_{np_y}(\mathbf{r}) & \text{(18a)} \\ \varphi'_{n,s,p_x,p_y}(\mathbf{r}) = a'\,\varphi_{ns}(\mathbf{r}) + b'\,\varphi_{np_x}(\mathbf{r}) + c'\,\varphi_{np_y}(\mathbf{r}) & \text{(18b)} \\ \varphi''_{n,s,p_x,p_y}(\mathbf{r}) = a''\,\varphi_{ns}(\mathbf{r}) + b''\,\varphi_{np_x}(\mathbf{r}) + c''\,\varphi_{np_y}(\mathbf{r}) & \text{(18c)} \end{cases}$$

We require the three wave functions (18) to be equivalent, that is, to be transformable into each other under rotation about Oz. Consequently, the proportion of the $\varphi_{ns}(\mathbf{r})$ orbital must be the same in each of them:

$$a = a' = a'' \tag{19}$$

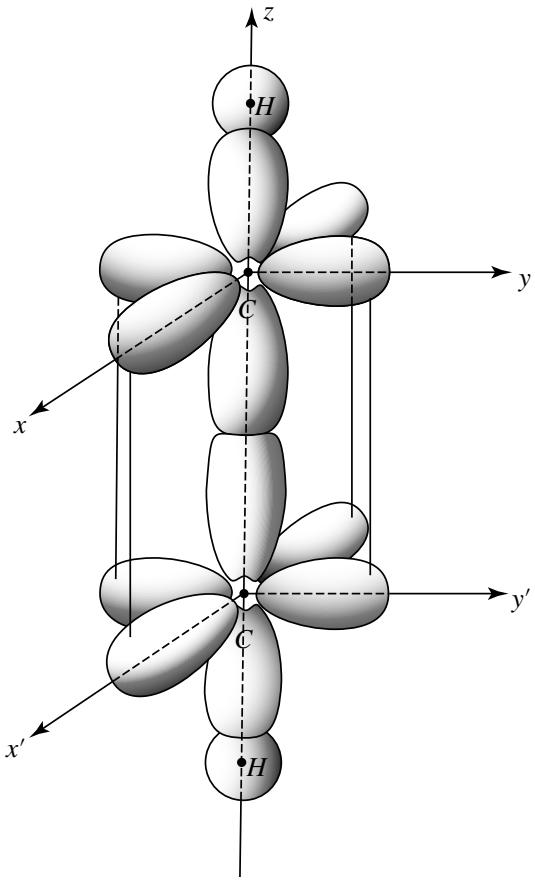


Figure 7: Schematic structure of the acetylene molecule C_2H_2. For each carbon atom, two electrons are in the sp_z hybrid orbitals (cf. Fig. 6) and contribute to the $C-H$ and $C-C$ bonds (σ bonds). In addition, two electrons are in the p_x and p_y orbitals and form additional bonds between the two carbon atoms (π bonds, weaker than σ bonds), shown by the vertical lines in the figure. The $C-C$ bond is therefore a "triple bond".

It is always possible to choose the axes so as to make the first orbital (18a) symmetric about the xOz plane. We can therefore choose:

$$c = 0 \tag{20}$$

By taking the three orbitals (18) to be normalized and orthogonal, we obtain six relations which enable us to determine[2] the six coefficients a, b, b', b'', c', c''. A simple calculation

[2] Actually, the signs of a, b and c' can be chosen arbitrarily.

yields:

$$\begin{cases} \varphi_{n,s,p_x,p_y}(\mathbf{r}) = \dfrac{1}{\sqrt{3}}\varphi_{ns}(\mathbf{r}) + \sqrt{\dfrac{2}{3}}\varphi_{np_x}(\mathbf{r}) & \text{(21a)} \\ \varphi'_{n,s,p_x,p_y}(\mathbf{r}) = \dfrac{1}{\sqrt{3}}\varphi_{ns}(\mathbf{r}) - \sqrt{\dfrac{1}{6}}\varphi_{np_x}(\mathbf{r}) + \dfrac{1}{\sqrt{2}}\varphi_{np_y}(\mathbf{r}) & \text{(21b)} \\ \varphi''_{n,s,p_x,p_y}(\mathbf{r}) = \dfrac{1}{\sqrt{3}}\varphi_{ns}(\mathbf{r}) - \sqrt{\dfrac{1}{6}}\varphi_{np_x}(\mathbf{r}) - \dfrac{1}{\sqrt{2}}\varphi_{np_y}(\mathbf{r}) & \text{(21c)} \end{cases}$$

We have thus produced what is called "sp^2 hybridization". The three hybrid orbitals (21) and the $\varphi_{np_z}(\mathbf{r})$ orbital form a new orthonormal basis in the space $\mathcal{E}_{ns} \oplus \mathcal{E}_{np}$.

4-b. Properties of sp^2 hybrid orbitals

We shall use the same graphical representation as in Figure 6. The $\varphi_{n,s,p_x,p_y}(\mathbf{r})$ orbital has cylindrical symmetry with respect to Ox. Figure 8-a represents the cross section in the xOy plane of the surface which describes its angular dependence for fixed r. The form of the curve obtained is completely analogous to that of Figure 6-a: the orbital points along the positive direction of the Ox axis.

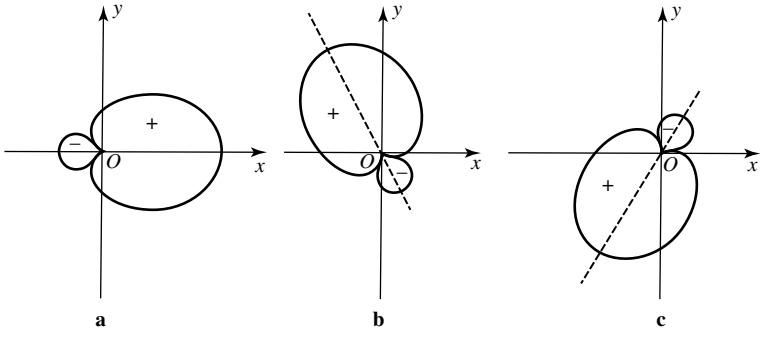


Figure 8: Angular dependence of the three orthogonal sp^2 orbitals. The φ_{n,s,p_x,p_y}, φ'_{n,s,p_x,p_y} and φ''_{n,s,p_x,p_y} orbitals can be transformed into each other by rotations through 120° about Oz.

By using expression (4b) for $\varphi_{np_x}(\mathbf{r})$, we can easily obtain the action on $|\varphi_{np_x}\rangle$ of the operator which performs a rotation through an angle α about Oz, $\mathrm{e}^{-i\alpha L_z/\hbar}$:

$$\mathrm{e}^{-i\alpha L_z/\hbar}|\varphi_{np_x}\rangle = \cos\alpha\,|\varphi_{np_x}\rangle + \sin\alpha\,|\varphi_{np_y}\rangle \tag{22}$$

Also, we obviously have:

$$\mathrm{e}^{-i\alpha L_z/\hbar}|\varphi_{ns}\rangle = |\varphi_{ns}\rangle \tag{23}$$

Formulas (21) then indicate that:

$$|\varphi'_{n,s,p_x,p_y}\rangle = \mathrm{e}^{-2i\frac{\pi}{3}L_z/\hbar}\,|\varphi_{n,s,p_x,p_y}\rangle \tag{24a}$$

$$|\varphi''_{n,s,p_x,p_y}\rangle = \mathrm{e}^{2i\frac{\pi}{3}L_z/\hbar}\,|\varphi_{n,s,p_x,p_y}\rangle \tag{24b}$$

The two orbitals (21b) and (21c) can therefore be obtained from the orbital (21a) by rotations through angles $2\pi/3$ and $-2\pi/3$ about Oz. Figures (8-b) and (8-c) give the cross sections in the xOy plane of the surfaces describing their angular dependence.

4-c. Example: the structure of ethylene

As in the acetylene molecule, each of the two carbon atoms of the ethylene molecule C_2H_4 has four valence electrons (one electron in the $2s$ sub-shell and three electrons in the $2p$ sub-shell).

Three of these four electrons occupy sp^2 hybrid orbitals of the type we just considered. They are the ones that, for each carbon atom, form the bonds with the neighboring carbon atom and the two hydrogen atoms of the CH_2 group. Thus we see why the three bonds, C – C, C – H, C – H, originating from one carbon atom are coplanar and form angles of 120° with each other (*cf.* Fig. 9, in which we have used the same graphical representation as in Figures 5 and 7). The remaining electron of each carbon atom occupies the $2p_z$ orbital. The $2p_z$ orbitals of the two carbons present a partial lateral overlapping, shown by the solid lines in Figure 9.

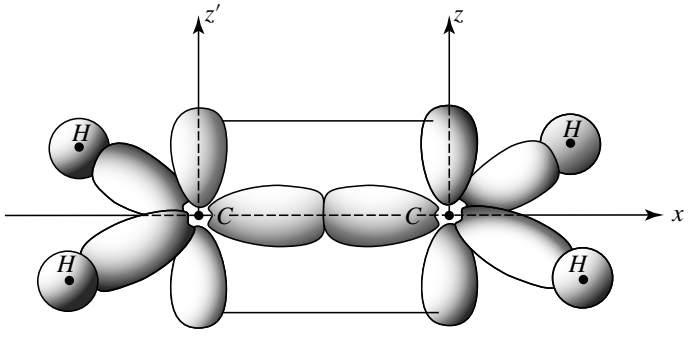


Figure 9: Schematic structure of the ethylene molecule C_2H_4. The two carbon atoms form a double bond with each other: one σ bond due to sp^2 orbitals of the type of those shown in Figure 8 (the other two sp^2 hybrid orbitals at 120° with this one form the C – H bonds), and one π bond, due to the overlapping of the p_z orbitals.

The two carbon atoms of the ethylene molecule are therefore connected by a *double bond*: one bond formed by two hybrid orbitals of the sp^2 type, cylindrically symmetric with respect to the Ox axis joining the two carbon atoms (σ bond), and one bond formed by two $2p_z$ orbitals, symmetric about the xOz plane (π bond). It is the latter bond which blocks the rotation of one CH_2 group with respect to the other one. If one of the CH_2 groups were to rotate with respect to the other one about the axis joining the two carbons, the axes of the two orbitals $2p_z$ and $2p_{z'}$ (Fig. 9) would no longer be parallel. This would diminish their lateral overlapping

and, consequently, the stability of the system. Thus we see why the six atoms of the ethylene molecule are in the same plane.

5. sp^3 hybridization

5-a. Introduction of sp^3 hybrid orbitals

We shall now superpose the four orbitals, $\varphi_{ns}(\mathbf{r})$, $\varphi_{np_x}(\mathbf{r})$, $\varphi_{np_y}(\mathbf{r})$, $\varphi_{np_z}(\mathbf{r})$, to form the four hybrid orbitals:

$$\begin{cases} \varphi_{n,s,p_x,p_y,p_z}(\mathbf{r}) = a\,\varphi_{ns}(\mathbf{r}) + b\,\varphi_{np_x}(\mathbf{r}) + c\,\varphi_{np_y}(\mathbf{r}) + d\,\varphi_{np_z}(\mathbf{r}) & \text{(25a)} \\ \varphi'_{n,s,p_x,p_y,p_z}(\mathbf{r}) = a'\,\varphi_{ns}(\mathbf{r}) + b'\,\varphi_{np_x}(\mathbf{r}) + c'\,\varphi_{np_y}(\mathbf{r}) + d'\,\varphi_{np_z}(\mathbf{r}) & \text{(25b)} \\ \varphi''_{n,s,p_x,p_y,p_z}(\mathbf{r}) = a''\,\varphi_{ns}(\mathbf{r}) + b''\,\varphi_{np_x}(\mathbf{r}) + c''\,\varphi_{np_y}(\mathbf{r}) + d''\,\varphi_{np_z}(\mathbf{r}) & \text{(25c)} \\ \varphi'''_{n,s,p_x,p_y,p_z}(\mathbf{r}) = a'''\,\varphi_{ns}(\mathbf{r}) + b'''\,\varphi_{np_x}(\mathbf{r}) + c'''\,\varphi_{np_y}(\mathbf{r}) + d'''\,\varphi_{np_z}(\mathbf{r}) & \text{(25d)} \end{cases}$$

We again require the four orbitals to have the same geometrical form, which means that:

$$a = a' = a'' = a''' \tag{26}$$

We can arbitrarily choose the symmetry axis of one of the orbitals, then the plane containing this axis and that of a second orbital. This reduces the number of free parameters to 10; we can find them by taking the four orbitals (25) to be orthonormal.

We shall content ourselves here with giving a possible set of such hybrid orbitals, defined by:

$$\begin{cases} a = b = c = d = \dfrac{1}{2} \\ a' = -b' = -c' = d' = \dfrac{1}{2} \\ a'' = -b'' = c'' = -d'' = \dfrac{1}{2} \\ a''' = b''' = -c''' = -d''' = \dfrac{1}{2} \end{cases} \tag{27}$$

and which can easily be shown to be orthonormal and of the same geometrical form. All the other possible sets can be obtained from this one by rotation.

We have thus produced what is called a "sp^3 hybridization". The four orbitals (25) corresponding to the coefficients (27) form a new orthonormal basis in the space $\mathcal{E}_{ns} \oplus \mathcal{E}_{np}$.

5-b. Properties of sp^3 hybrid orbitals

The four orbitals constructed in § 5-a are analogous in form to those studied in §§ 3 and 4. They point respectively in the directions of the vectors whose components are:

$$\begin{cases} (1,1,1) \\ (-1,-1,1) \\ (-1,1,-1) \\ (1,-1,-1) \end{cases} \tag{28}$$

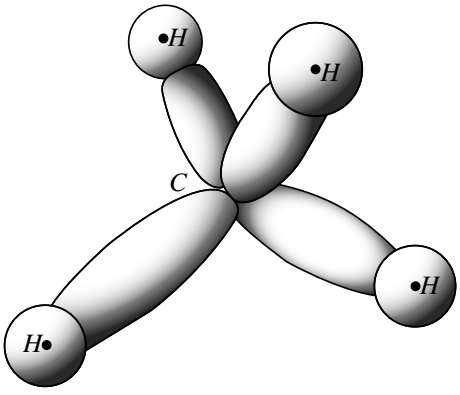


Figure 10: Schematic structure of the methane molecule. The sp^3 orbitals produce bonds arranged like the straight lines joining the center of a tetrahedron to its four corners (angles of 109° 28′).

The axes of the four sp^3 orbitals are therefore arranged like the straight lines joining the center of a regular tetrahedron to the four corners of this tetrahedron. The angle between any two of these straight lines is equal to 109° 28′.

5-c. Example: The structure of methane

In the methane molecule CH_4, the four valence electrons of the carbon atom each occupy one of the four sp^3 hybrid orbitals studied above. This immediately explains why the four hydrogen atoms form the corners of a regular tetrahedron centered on the carbon atom (Fig. 10).

In the ethane molecule C_2H_6, one of the hydrogen atoms of methane is replaced by a CH_3 group. The two carbon atoms are then connected by a single bond, formed by two sp^3 hybrid orbitals, cylindrically symmetric with respect to the straight line joining the two carbon atoms. The absence of a double bond permits the practically free rotation of one CH_3 group with respect to the other one.

References and suggestions for further reading:

Various geometrical representations of orbitals: Levine (12.3), § 6.6; Karplus and Porter (12.1), § 3.10.

Hybrid orbitals: Karplus and Porter (12.1), § 6.3; Alonso and Finn III (1.4), § 5-5; Eyring et al. (12.5), Chap. XII, § 12 b; Coulson (12.6), Chap. VIII; Pauling (12.2), Chap. III, §§ 13 and 14.

Complement F_{VII}

Vibrational-rotational levels of diatomic molecules

1 Introduction 885

2 Approximate solution of the radial equation 886

2-a The zero angular momentum states ($l = 0$) 886

2-b General case (l any positive integer) 888

2-c The vibrational-rotational spectrum 889

3 Evaluation of some corrections 892

3-a More precise study of the form of the effective potential $V_{\text{eff}}(r)$ 893

3-b Energy levels and wave functions of the stationary states . . 894

3-c Interpretation of the various corrections 895

1. Introduction

In this complement, we shall use the results of Chapter VII to study quantum mechanically the stationary states of the system formed by the two nuclei of a diatomic molecule. We shall simultaneously take into account all the degrees of freedom of the system: vibration of the two nuclei about their equilibrium position and rotation of the system about the center of mass. We shall show that the results obtained in Complements A_V and C_{VI}, in which one degree of freedom at a time was considered, are valid to a first approximation. In addition, a certain number of corrections due to the "centrifugal distortion" of the molecule and to the vibration-rotation coupling will be calculated and interpreted.

We saw in § 1-a of Complement A_V (the Born-Oppenheimer approximation) that the potential energy $V(r)$ of interaction between the two nuclei depends only on the distance r between them and has the form shown in Figure 1: $V(r)$ is attractive at large distances, repulsive at short distances, and has a minimum at $r = r_e$ of depth V_0. Let m_1 and m_2 be the masses of the two nuclei. Since $V(r)$ depends only on r, we can, according to § B of Chapter VII, study separately the motion of the center of mass (a free particle of mass $M = m_1 + m_2$) and the relative motion in the center of mass reference frame, which is equivalent to that of a fictitious particle of mass:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{1}$$

placed in the potential $V(r)$ of Figure 1.

If we are interested only in the relative motion, the stationary states of the system, according to the results of § A of Chapter VII, are described by the wave functions:

$$\varphi_{v,l,m}(r, \theta, \varphi) = \frac{1}{r}\, u_{v,l}(r)\, Y_l^m(\theta, \varphi) \tag{2}$$

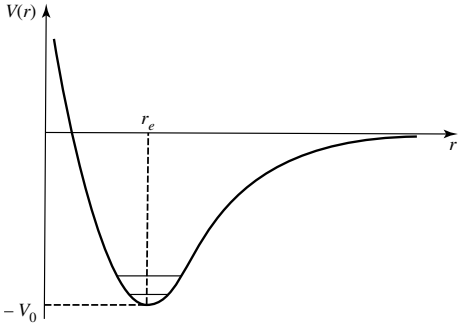


Figure 1: Variation of the interaction potential energy $V(r)$ between the nuclei of a diatomic molecule as a function of the distance r between them. $V(r)$ takes on its minimum value $-V_0$ at $r = r_e$. The first vibrational states are represented by the horizontal lines in the potential well.

where the corresponding energies $E_{v,l}$ and the radial functions $u_{v,l}(r)$ are given by the equation:

$$\left[-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2}\right] u_{v,l}(r) = E_{v,l}\ u_{v,l}(r) \tag{3}$$

Comment:

Rigorously speaking, we implicitly assume in all of this complement (as in A_V and C_{VI}) that the projection of the total orbital angular momentum of the electrons onto the internuclear axis is zero, as is their total spin. The total angular momentum of the molecule then arises only from the rotation of the two nuclei. Such a situation is found in virtually all diatomic molecules in their ground states. In the general case, terms will also appear in the nuclear interaction energy which do not depend exclusively on the distance r.

2. Approximate solution of the radial equation

The radial equation has the same form as the eigenvalue equation of the Hamiltonian of a one-dimensional problem in which a particle of mass μ is placed in the effective potential:

$$V_{\text{eff}}(r) = V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2} \tag{4}$$

2-a. The zero angular momentum states ($l = 0$)

For $l = 0$, the "centrifugal potential" $l(l+l)\hbar^2/2\mu r^2$ is zero, and $V_{\text{eff}}(r)$ is then the same as $V(r)$. In the neighborhood of the minimum at $r = r_e$, $V(r)$ can be expanded in powers of $r - r_e$:

$$V(r) = -V_0 + f(r - r_e)^2 - g(r - r_e)^3 + ... \tag{5}$$

The coefficients f and g are positive, since $r = r_e$ is a minimum and since the potential increases faster for $r < r_e$ than for $r > r_e$.

We begin by neglecting the term in $(r - r_e)^3$ and terms of higher order. The potential is then purely parabolic, and we know the eigenstates and eigenvalues of the Hamiltonian. If we set:

$$\omega = \sqrt{\frac{2f}{\mu}} \tag{6}$$

we obtain levels whose energy is:

$$E_{v,0} = -V_0 + \left(v + \frac{1}{2}\right)\hbar\omega \tag{7}$$

$(v = 0, 1, 2, \ldots)$

with the associated wave functions (*cf.* Chap. V and Complement $\mathrm{B_V}$):

$$u_v(r) = \left(\frac{\beta^2}{\pi}\right)^{1/4} \frac{1}{\sqrt{2^v v!}} \mathrm{e}^{-\beta^2(r-r_e)^2/2} H_v[\beta(r - r_e)] \tag{8}$$

with:

$$\beta = \sqrt{\frac{\mu\omega}{\hbar}} \tag{9}$$

(H_v is a Hermite polynomial). In Figure 1, we have represented the first two energy levels by horizontal lines. The length of the lines gives an idea of the extension $(\Delta r)_v$ of the wave functions corresponding to these states. Recall [Chap. V, formula (D-5a)] that:

$$(\Delta r)_v \simeq \sqrt{\left(v + \frac{1}{2}\right)\frac{\hbar}{\mu\omega}} \tag{10}$$

For the preceding calculation to be valid, it is necessary, in a region of width $(\Delta r)_v$ about $r = r_e$, for the term in $(r - r_e)^3$ of (5) to be always negligible compared to the term in $(r - r_e)^2$. We must therefore have:

$$f \gg g(\Delta r)_v = g(\Delta r)_0 \sqrt{v + \frac{1}{2}} \tag{11}$$

where $(\Delta r)_0$ is the extension of the ground state:

$$(\Delta r)_0 = \sqrt{\frac{\hbar}{\mu\omega}} \tag{12}$$

This implies, in particular that:

$$f \gg g(\Delta r)_0 \tag{13}$$

Condition (13) is always satisfied in practice. We shall confine ourselves in what follows to quantum numbers v which are small enough for (11) to be satisfied as well.

Comment:

Expansion (5) is obviously not valid at $r = 0$, where $V(r)$ is infinite. The preceding argument implicitly assumes that:

$$(\Delta r)_v \ll r_e \tag{14}$$

In this case, the wave functions (8) are practically zero at the origin, and almost identical to the exact solutions of the radial equation (3), which must be rigorously zero at $r = 0$ (*cf.* § A-2-c of Chapter VII).

2-b. General case (l any positive integer)

α. Evaluation of the effect of the centrifugal potential

At $r = r_e$, the centrifugal potential is equal to:

$$\frac{l(l+1)\hbar^2}{2\mu r_e^2} = Bh\, l(l+1) \tag{15}$$

where:

$$B = \frac{\hbar}{4\pi\mu r_e^2} \tag{16}$$

is the rotational constant introduced in Complement C_{VI}. We have already pointed out in that complement (§ 4-a-β) that the energy $2Bh$ (the distance between two adjacent lines of the pure rotational spectrum) is always very much smaller than $\hbar\omega$ (the vibrational quantum):

$$2Bh \ll \hbar\omega \tag{17}$$

We shall confine ourselves here to rotational quantum numbers l sufficiently small that:

$$Bh\, l(l+1) \ll \hbar\omega \tag{18}$$

In a domain of small width Δr about $r = r_e$, the variation of the centrifugal potential is of the order of:

$$\frac{l(l+1)\hbar^2}{\mu r_e^3}\Delta r = 2Bh\, l(l+1)\frac{\Delta r}{r_e} \tag{19}$$

That of the potential $V(r)$ is approximately:

$$f(\Delta r)^2 = \frac{1}{2}\mu\omega^2(\Delta r)^2 = \frac{1}{2}\hbar\omega\frac{(\Delta r)^2}{(\Delta r)_0^2} \tag{20}$$

where we have used (12). We know from § 2-a that the extension Δr of the wave functions we shall be considering is negligible with respect to r_e, but certainly at least of the order of $(\Delta r)_0$. Consequently, in the region of space in which the wave functions have significant amplitudes, the variation (19) of the centrifugal potential is, according to (18), much smaller than that of $V(r)$ found in (20). We can then, to a first approximation, replace the centrifugal potential, in equation (4), by its value (15) at $r = r_e$. This gives for the effective potential:

$$V_{\text{eff}}(r) \simeq V(r) + Bh\, l(l+1) \tag{21}$$

β. Energy levels and stationary wave functions

By using (21) and neglecting terms of order greater than two in expansion (5), we can put the radial equation (3) in the form:

$$\left[-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{1}{2}\mu\omega^2(r-r_e)^2\right]u_{v,l}(r)=\left[E_{v,l}+V_0-Bh\,l(l+1)\right]u_{v,l}(r) \tag{22}$$

which is completely analogous to the eigenvalue equation of a one-dimensional harmonic oscillator.

It thus follows that the term in brackets on the right-hand side must be equal to $(v+1/2)\hbar\omega$, where $v=0,1,2\ldots$; this yields the possible energies $E_{v,l}$ of the molecule:

$$E_{v,l}=-V_0+\left(v+\frac{1}{2}\right)\hbar\omega+Bh\,l(l+1) \tag{23}$$

with:

$$\begin{cases} v=0,1,2... \\ l=0,1,2... \end{cases}$$

As for the radial functions, they do not depend on l, since the differential appearing on the left-hand side of (22) does not depend on l. Consequently, we have:

$$u_{v,l}(r)=u_v(r) \tag{24}$$

where $u_v(r)$ was given in (8). Expression (2) for the wave functions of the stationary states can then be written, in this approximation:

$$\varphi_{v,l,m}(r,\theta,\varphi)=\frac{1}{r}u_v(r)\,Y_l^m(\theta,\varphi) \tag{25}$$

Thus we see that the energies of the stationary states are the sums of the energies calculated in Complements $\mathrm{A_V}$ and $\mathrm{C_{VI}}$, in which only one degree of freedom at a time (vibration or rotation) was taken into account. In addition, the wave functions are the products of the wave functions found in these two complements, to within a factor of $1/r$.

Figure 2 shows the first two vibrational levels $v=0$ and $v=1$, with their rotational structure due to the term $Bh\,l(l+1)$.

2-c. The vibrational-rotational spectrum

We shall confine ourselves to the study of the infrared absorption or emission spectrum, thereby assuming the molecule to be heteropolar (calculations analogous to those presented in § l-c-β of $\mathrm{A_V}$ and § 4-b of $\mathrm{C_{VI}}$ could be performed when dealing with homopolar molecules and the Raman effect).

α. Selection rules

Recall that the dipole moment $D(r)$ of the molecule is directed along the straight line joining the two nuclei and can be expanded in powers of $r-r_e$ about r_e:

$$D(r)=d_0+d_1(r-r_e)+... \tag{26}$$

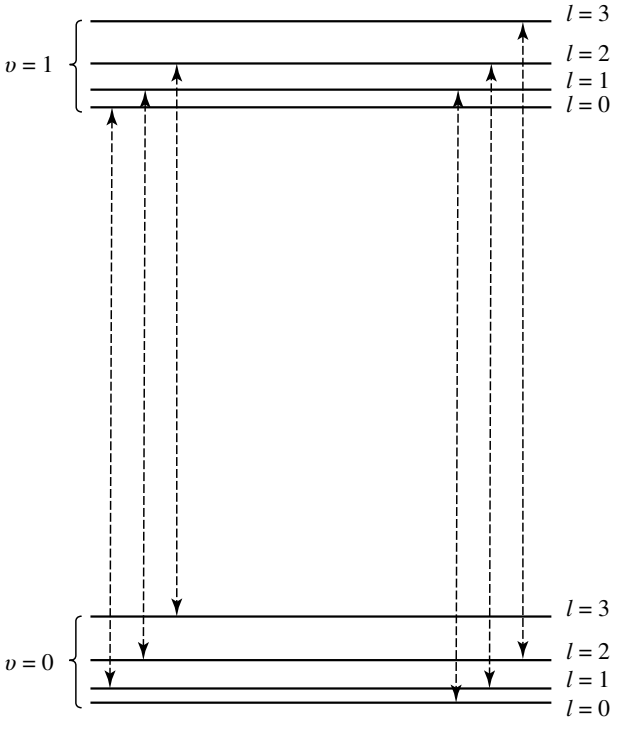


Figure 2: Diagram showing the first two vibrational levels ($v = 0$ and $v = 1$) of a diatomic molecule and their rotational structure ($l = 0, 1, 2, ...$). Within the limits of the approximations, this rotational structure is the same for the various vibrational levels. For a heteropolar molecule, the transitions represented by the vertical arrows in the figure yield the lines of the vibrational-rotational spectrum of the molecule. These lines fall in the infrared. These transitions obey the selection rule $\Delta l = l' - l = \pm 1$.

The projection of this dipole moment onto Oz is equal to $D(r)\cos\theta$ (where θ is the angle between the axis of the molecule and Oz).

We want to determine the frequency spectrum of electromagnetic waves polarized along Oz that the molecule can absorb or emit as a consequence of the variation of its electric dipole. As we have done several times before, we shall look for the Bohr frequencies that can appear in the time evolution of the mean value of $D(r)\cos\theta$. All we must do, then, is find for what values of v', l', m' and v, l, m the matrix element:

$$\langle \varphi_{v',l',m'} | D(r)\cos\theta | \varphi_{v,l,m} \rangle = \int r^2 \, \mathrm{d}r \, \mathrm{d}\Omega \, \varphi^*_{v',l',m'}(r,\theta,\varphi) \, D(r)\cos\theta \, \varphi_{v,l,m}(r,\theta,\varphi) \quad (27)$$

is different from zero. Using expression (25) for the wave functions, we put this matrix

element in the form:

$$\left[\int_0^\infty \mathrm{d}r\, u^*_{v'}(r)\, D(r) u_v(r)\right] \times \left[\int \mathrm{d}\Omega\, Y_{l'}^{m'*}(\theta,\varphi) \cos\theta\, Y_l^m(\theta,\varphi)\right] \tag{28}$$

We thus obtain a product of two integrals which have already been treated in Complements $\mathrm{A_V}$ and $\mathrm{C_{VI}}$. The second integral is different from zero only if:

$$l' - l = +1, -1 \tag{29}$$

As for the first one, if we confine ourselves to the terms in d_0 and d_1 of (26), it is different from zero only if:

$$v' - v = 0, +1, -1 \tag{30}$$

The set of lines corresponding to $v - v' = 0$ constitutes the pure rotational spectrum studied in Complement $\mathrm{C_{VI}}$ (its intensity is proportional to d_0^2). As for the lines $v' - v = \pm 1$, $l' - l = \pm 1$, of intensity proportional to d_1^2, they constitute the vibrational-rotational spectrum which we shall now briefly describe.

Comment:

The selection rule $l' - l = +1$ arises from the angular dependence of the wave functions. It is therefore independent of the approximation used to solve the radial equation (3), while (30) is valid only in the harmonic approximation.

β. Form of the spectrum

Let v' be the larger of the two vibrational quantum numbers under consideration ($v' = v + 1$). The vibrational-rotational lines can be separated into two groups:

– the lines $v' = v + 1$, $l' = l + 1 \leftrightarrow v$, l, of frequencies:

$$\frac{\omega}{2\pi} + B(l+1)(l+2) - Bl(l+1) = \frac{\omega}{2\pi} + 2B(l+1) \tag{31}$$

with $l = 0, 1, 2, \ldots$
(these lines correspond to the transitions indicated by the arrows on the right-hand side of Figure 2).

– the lines $v' = v + 1$, $l' = l - 1 \leftrightarrow v$, l, of frequencies:

$$\frac{\omega}{2\pi} + Bl'(l'+1) - B(l'+1)(l'+2) = \frac{\omega}{2\pi} - 2B(l'+1) \tag{32}$$

with $l' = 0, 1, 2, \ldots$
(transitions indicated by the arrows on the left-hand side of Figure 2).

The vibrational-rotational spectrum therefore has the form shown in Figure 3. It contains two groups of equidistant lines, symmetric with respect to the vibrational frequency $\omega/2\pi$. All these lines together constitute a "band". The group of lines corresponding to frequencies (31) is called the "R branch" of the band, and the one corresponding to frequencies (32), the "P branch" of the band. In each branch, the distance between two adjacent lines is $2B$. The central interval separating the two branches is of width

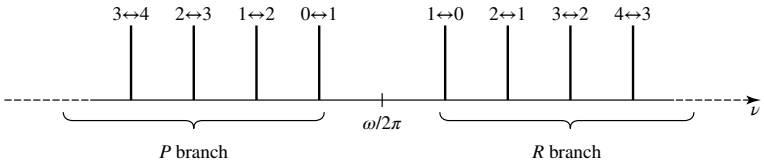


Figure 3: The vibrational-rotational spectrum for a heteropolar molecule. Since transitions between levels of Figure 2 with the same value of l are forbidden by the selection rules, no line has the pure vibrational frequency $\frac{\omega}{2\pi}$. Transitions in which the molecule passes from the level (v', l') to the level $(v = v' - 1,\ l = l' - 1)$ correspond to frequencies $\frac{\omega}{2\pi} + 2B(l+1)$ (lines of the "R branch"). Transitions in which the molecule passes from the level (v', l') to the level $(v = v' - 1,\ l = l' + 1)$ correspond to frequencies $\frac{\omega}{2\pi} - 2B(l'+1)$ (lines of the "P branch"). The different lines are labeled $l' \leftrightarrow l$ in the figure.

$4B$: there is no line at the pure vibrational frequency $\omega/2\pi$ (there is often said to be a "missing line" in the spectrum).

Comment:

The "pure vibrational" spectrum, studied in A_V and composed of a single line at $\omega/2\pi$, therefore does not exist in practice. It is only when one uses a spectroscopic device with low resolution that one can ignore the rotational structure of the vibrational-rotational line and treat the band of Figure 3 like a single line centered at $\omega/2\pi$ (recall that $\omega/2\pi \gg 2B$).

3. Evaluation of some corrections

The calculations of the preceding section are based on the approximation that replaces the centrifugal potential by its value at $r = r_e$ in the radial equation. The effective potential $V_{\text{eff}}(r)$ can then be obtained from $V(r)$ by a simple vertical translation [formula (21)].

In this section, we shall study the corrections that must be performed on the results of § 2 in order to take into account the slow variation of the centrifugal potential about $r = r_e$. To do so, we shall use its expansion in powers of $(r - r_e)$:

$$\frac{l(l+1)\hbar^2}{2\mu r^2} = \frac{l(l+1)\hbar^2}{2\mu r_e^2} - \frac{l(l+1)\hbar^2}{\mu r_e^3}(r - r_e) + \frac{3l(l+1)\hbar^2}{2\mu r_e^4}(r - r_e)^2 + \ldots \tag{33}$$

3-a. More precise study of the form of the effective potential $V_{\text{eff}}(r)$

If we use (5) and (33), the expansion of the effective potential (4) in the neighborhood of $r = r_e$ can be written:

$$V_{\text{eff}}(r) = -V_0 + f(r-r_e)^2 - g(r-r_e)^3 + ... + \frac{l(l+1)\hbar^2}{2\mu r_e^2} - \frac{l(l+1)\hbar^2}{\mu r_e^3}(r-r_e) + \frac{3l(l+1)\hbar^2}{2\mu r_e^4}(r-r_e)^2 + ... \tag{34}$$

We shall see that the variation of the centrifugal potential in the neighborhood of $r = r_e$ produces, for l different from zero, the following effects:

(i) The position $\tilde{r}_e$ of the minimum of $V_{\text{eff}}(r)$ does not coincide exactly with r_e.

(ii) The value $V_{\text{eff}}(\tilde{r}_e)$ of this minimum is slightly different from $-V_0 + Bh\,l(l+1)$.

(iii) The curvature of $V_{\text{eff}}(r)$ at $r = \tilde{r}_e$ [which fixes, as in formula (6), the angular frequency of the equivalent harmonic oscillator] is no longer strictly given by the coefficient f.

We shall evaluate these various effects by using expansion (34). As far as the first two are concerned, we can neglect terms of order higher than 2 in $V(r)$, and those of order higher than 1 in the centrifugal potential , since the distance $\tilde{r}_e - r_e$ which we shall find is very small [it will even be small relative to $(\Delta r)_0$]. In fact, we shall be able to verify *a posteriori* that:

$$g(\tilde{r}_e - r_e) \ll f \tag{35a}$$

$$\frac{3l(l+1)\hbar^2}{2\mu r_e^4}(\tilde{r}_e - r_e) \ll \frac{l(l+1)\hbar^2}{\mu r_e^3} \tag{35b}$$

α. *Position and value of the minimum of* $V_{\text{eff}}(r)$

If, in expansion (34), we keep only the first two terms of $V(r)$ and the first two terms of the centrifugal potential, $\tilde{r}_e$ is given by:

$$2f(\tilde{r}_e - r_e) \simeq \frac{l(l+1)\hbar^2}{\mu r_e^3} \tag{36}$$

that is:

$$\tilde{r}_e - r_e \simeq \frac{l(l+1)\hbar^2}{2\mu f r_e^3} = \frac{Bh\,l(l+1)}{f r_e} \tag{37}$$

According to (6) and (12), we have:

$$\frac{\tilde{r}_e - r_e}{(\Delta r)_0} \simeq \frac{2Bh\,l(l+1)}{\hbar\omega}\frac{(\Delta r)_0}{r_e} \ll 1 \tag{38}$$

which, with (13) and (14) taken into account, proves (35a) and (35b).

Substituting this value of $\tilde{r}_e$ into the expansion of $V_{\text{eff}}(r)$, we find:

$$V_{\text{eff}}(\tilde{r}_e) \simeq -V_0 + Bh\,l(l+1) - Gh[l(l+1)]^2 \tag{39}$$

with:

$$G = \frac{\hbar^3}{8\pi\mu^2 r_e^6 f} \tag{40}$$

β. Curvature of $V_{\text{eff}}(r)$ at its minimum

In the neighborhood of $r = \tilde{r}_e$, we can therefore write $V_{\text{eff}}(r)$ in the form:

$$V_{\text{eff}}(r) = V_{\text{eff}}(\tilde{r}_e) + f'(r - \tilde{r}_e)^2 - g'(r - \tilde{r}_e)^3 + ... \tag{41}$$

The coefficient f' is related to the curvature of $V_{\text{eff}}(r)$ at $r = \tilde{r}_e$:

$$f' = \frac{1}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} V_{\text{eff}}(r)\right]_{r=\tilde{r}_e} \tag{42}$$

To evaluate the difference between f' and f, we must take into account the term in $(r - r_e)^3$ of $V(r)$ in expansion (34) and, consequently, also the term in $(r - r_e)^2$ of the centrifugal potential. A simple calculation then yields, using (37):

$$2f' \simeq 2f + \frac{3l(l+1)\hbar^2}{\mu r_e^4} - \frac{3g\, l(l+1)\hbar^2}{\mu r_e^3 f} \tag{43}$$

The angular frequency ω defined by (6) must therefore be replaced by:

$$\omega' = \sqrt{\frac{2f'}{\mu}} \tag{44}$$

Expanding the square root, we easily find:

$$\omega' = \omega - 2\pi\alpha_e\, l(l+1) \tag{45}$$

with:

$$\alpha_e = \frac{3\hbar^2\omega}{8\pi\mu r_e^3 f}\left[\frac{g}{f} - \frac{1}{r_e}\right] \tag{46}$$

We could carry out an analogous calculation to determine g'. Actually, since the term in $(r - \tilde{r}_e)^3$ of (41) adds only a small correction to the results obtained using the first two terms, we shall neglect the variation of $\dfrac{\mathrm{d}^3}{\mathrm{d}r^3} V_{\text{eff}}(r)$ when we go from r_e to $\tilde{r}_e$, and take $g' \simeq g$.

In conclusion, in the neighborhood of its minimum, we can write $V_{\text{eff}}(r)$ in the form:

$$V_{\text{eff}}(r) \simeq V_{\text{eff}}(\tilde{r}_e) + \frac{1}{2}\mu\omega'^2(r - \tilde{r}_e)^2 - g(r - \tilde{r}_e)^3 \tag{47}$$

where $\tilde{r}_e$, $V_{\text{eff}}(\tilde{r}_e)$, ω' are given by (37), (39) and (45).

3-b. Energy levels and wave functions of the stationary states

With expression (47) for $V_{\text{eff}}(r)$, the radial equation becomes:

$$\left[-\frac{h^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{1}{2}\mu\omega'^2(r - \tilde{r}_e)^2 - g(r - \tilde{r}_e)^3\right] u_{v,l}(r) = [E_{v,l} - V_{\text{eff}}(\tilde{r}_e)]\, u_{v,l}(r) \tag{48}$$

If, as in § 2, we neglect the term in $g(r-\tilde{r}_e)^3$, we recognize the eigenvalue equation of a one-dimensional harmonic oscillator of angular frequency ω' whose equilibrium position is $r=\tilde{r}_e$. From this, we deduce that the only possible values of the term in brackets on the right-hand side are $(v+1/2)\hbar\omega'$, with $v=0,1,2,$ ... According to (39), we therefore have:

$$E_{v,l}=-V_0+\left(v+\frac{1}{2}\right)\hbar\omega'+Bh\,l(l+1)-Gh[l(l+1)]^2 \tag{49}$$

As for the wave functions of the stationary states, they have the same form as in (25). All we need to do in expression (8) for the radial function is replace r_e by $\tilde{r}_e$ and β by:

$$\beta'=\sqrt{\frac{\mu\omega'}{\hbar}} \tag{50}$$

We have taken into account the term in $g(r-r_e)^3$ in the calculation of the new angular frequency ω'. For the calculation to be consistent, it is then necessary to evaluate the corrections in the eigenvalues and eigenfunctions of the radial equation due to the presence of this term on the left-hand side of (48). We shall do this in Complement A_{XI}, using perturbation theory. Here, we shall content ourselves with stating the result concerning the eigenvalues: we must add to expression (49) for the energy the term:

$$\xi\hbar\omega'\left(v+\frac{1}{2}\right)^2+\frac{7}{60}\xi\hbar\omega' \tag{51}$$

where:

$$\xi=-\frac{15}{4}\frac{g^2\hbar}{\mu^3\omega'^5} \tag{52}$$

is a dimensionless quantity much smaller than 1 (hence ω' may be replaced by ω in this corrective term).

3-c. Interpretation of the various corrections

α. Centrifugal distortion of the molecule

The discussion of § 3-a-α shows that the distance between the two nuclei increases when the molecule rotates. According to (37), this increase in distance becomes larger when $l(l+1)$ becomes larger, that is, when the molecule rotates faster. This is quite comprehensible: in classical terms, one would say that the "centrifugal force" tends to separate the two nuclei until it is balanced by the restoring force $2f(\tilde{r}_e-r_e)$ due to the potential $V(r)$.

The molecule is therefore not really a "rigid rotator". The variation $\tilde{r}_e-r_e$ of the average distance between the nuclei produces an increase in the moment of inertia of the molecule and, consequently, a decrease (at constant angular momentum) in the rotational energy. This decrease is only partially compensated by the increase in the potential energy $V(\tilde{r}_e)-V(r_e)$. This is the physical origin of the energy correction $-Gh\,l^2(l+1)^2$ which appears in (49). This correction, whose sign is negative, increases

much faster with l than the rotational energy $Bh\,l(l+1)$. This can be seen experimentally: the lines of the pure rotational spectrum are not rigorously equidistant; the separation of the lines decreases when l increases.

β. Vibrational-rotational coupling

We shall group the second and third terms of (49) and replace ω' by its expression (45). We obtain:

$$\left(v+\frac{1}{2}\right)\hbar\omega' + Bh\,l(l+1) = \left(v+\frac{1}{2}\right)\hbar\omega + Bh\,l(l+1) - \alpha_e h\,l(l+1)\left(v+\frac{1}{2}\right) \quad (53)$$

The first two terms on the right-hand side of (53) are the vibrational and rotational energies calculated in Complements A_V and C_{VI}. The third term, which depends on the two quantum numbers v and l, represents the effects of the coupling of the vibrational and rotational degrees of freedom.

We can rewrite (53) in the form:

$$\left(v+\frac{1}{2}\right)\hbar\omega + B_v h\,l(l+1) \quad (54)$$

with:

$$B_v = B - \alpha_e\left(v+\frac{1}{2}\right) \quad (55)$$

It looks as if each vibrational level had an effective rotational constant B_v depending on the v associated.

To explain this coupling of the vibration and rotation of the molecule, we shall argue in classical terms. The rotational constant B is proportional to $1/r^2$ [formula (16)]. When the molecule vibrates, r varies, and, consequently, so does B. Since the vibrational frequencies are much higher than the rotational ones, we can define an effective rotational constant of the molecule in a given vibrational state: this will be the average of B taken over a time interval which is much longer than the vibrational period. We must therefore take the time average of $1/r^2$ in the vibrational state under consideration.

In this way, we can interpret the two terms of opposite sign that appear in expression (46) for α_e. The first of these terms, which is proportional to g, is due to the anharmonicity of the potential $V(r)$, which increases with the amplitude of vibration (that is, in fact, with v). Given the asymmetric form of $V(r)$ (Fig. 1), the molecule "spends more time" in the region $r > r_e$ than in the region $r < r_e$. It follows that the average value of $1/r^2$ is less than $1/r_e^2$: the anharmonicity decreases the effective rotational constant. This can be seen in formulas (55) and (46). Actually, even if the vibrational motion were perfectly symmetric with respect to r_e (that is, if g were zero), the average value of $1/r^2$ would not be equal to $1/r_e^2$, since:

$$\left\langle\frac{1}{r^2}\right\rangle \neq \frac{1}{\langle r\rangle^2} \quad (56)$$

This is the origin of the second term of expression (46): when the average of $1/r^2$ is taken, small values of r are favorized, so that $\langle 1/r^2\rangle$ is greater than $1/\langle r\rangle^2$; hence the sign of this second correction.

The overall sign of α_e results from the competition between the two preceding effects. In general, the anharmonicity term dominates, so that α_e is positive and B_v is less than B.

Comments:

(i) Vibrational-rotational coupling exists even in the vibrational ground state $v = 0$:

$$B_0 = B - \frac{1}{2}\alpha_e \tag{57}$$

This is another manifestation of the finite extension $(\Delta r)_0$ of the wave function of the $v = 0$ state.

(ii) Experimentally, vibrational-rotational coupling appears in the following way: if α_e is positive, the rotational structure is slightly more compact in the higher vibrational state v' than in the lower vibrational state $v = v' - 1$. It is easy to show that the P and R branches of Figure 3 are affected differently. Adjacent lines are no longer completely equidistant and are, on the average, closer together in the R branch than in the P branch.

To sum up, the energy of a vibrational-rotational level of a diatomic molecule, labeled by the quantum numbers v and l, is given by:

$$E_{v,l} = -V_0 + \left(v + \frac{1}{2}\right)\hbar\omega + \left[B - \alpha_e\left(v + \frac{1}{2}\right)\right] h\, l(l+1) - Gh\, l^2(l+1)^2 + \xi\left(v + \frac{1}{2}\right)^2 \hbar\omega + \frac{7}{60}\xi\hbar\omega \tag{58}$$

V_0: dissociation energy of the molecule;
$\omega/2\pi$: vibrational frequency;
B: rotational constant given by (16);
G, α_e, ξ: dimensionless constants given by (40), (46)) and (52).

References and suggestions for further reading:

Molecular spectra: Eisberg and Resnick (1.3), Chap. 12; Pauling and Wilson (1.9), Chap. X; Karplus and Porter (12.1), Chap. 7; Herzberg (12.4), Vol. I, Chap. III, §§ 2 b and 2 c; Landau and Lifshitz (1.19), Chaps. XI and XIII.

Nuclear vibration and rotation: Valentin (16.1), § VII-2.

Complement G_{VII}

Exercises

1. Particle in a cylindrically symmetric potential

Let ρ, φ, z be the cylindrical coordinates of a spinless particle ($x = \rho\cos\varphi$, $y = \rho\sin\varphi$; $\rho \geq 0$, $0 \leq \varphi < 2\pi$). Assume that the potential energy of this particle depends only on ρ, and not on φ and z. Recall that:

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial \varphi^2}$$

a. Write, in cylindrical coordinates, the differential operator associated with the Hamiltonian. Show that H commutes with L_z and P_z. Show that this allows writing the wave functions associated with the stationary states of the particle as:

$$\varphi_{n,m,k}(\rho,\varphi,z) = f_{n,m}(\rho)\ \mathrm{e}^{im\varphi}\,\mathrm{e}^{ikz}$$

where the values that can be taken on by the indices m and k are to be specified.

b. Write, in cylindrical coordinates, the eigenvalue equation of the Hamiltonian H of the particle. Derive from it the differential equation that $f_{n,m}(\rho)$ obeys.

c. Let Σ_y be the operator whose action, in the $\{|\mathbf{r}\rangle\}$ representation, is to change y to $-y$ (reflection with respect to the xOz plane). Does Σ_y commute with H? Show that Σ_y anticommutes with L_z, and show that, as a result, $\Sigma_y|\varphi_{n,m,k}\rangle$ is an eigenvector of L_z. What is the corresponding eigenvalue? What can be concluded concerning the degeneracy of the energy levels of the particle? Could this result be predicted directly from the differential equation established in (*b*)?

2. Three-dimensional harmonic oscillator in a uniform magnetic field

N.B. The object of this exercise is to study a simple physical system for which the effect of a uniform magnetic field can be calculated exactly. This will allow comparing precisely the relative importance of the "paramagnetic" and "diamagnetic" terms. The modification of the wave function of the ground state due to the effect of the diamagnetic term is also detailed (the reader may wish to refer to Complements D_{VI} and B_{VII}).

Consider a particle of mass μ, whose Hamiltonian is:

$$H_0 = \frac{\mathbf{P}^2}{2\mu} + \frac{1}{2}\mu\omega_0^2\ \mathbf{R}^2$$

(an isotropic three-dimensional harmonic oscillator), where ω_0 is a given positive constant.

a. Find the energy levels of the particle and their degrees of degeneracy. Is it possible to construct a basis of eigenstates common to H_0, $\mathbf{L}^2$, L_z?

b. Now, assume that the particle, which has a charge q, is placed in a uniform magnetic field $\mathbf{B}$ parallel to Oz. We set $\omega_L = -qB/2\mu$. If we choose to use the gauge $\mathbf{A} = -\frac{1}{2}\mathbf{r}\times\mathbf{B}$, the Hamiltonian H of the particle is:

$$H = H_0 + H_1(\omega_L)$$

where H_1 is the sum of an operator which is linearly dependent on ω_L (the paramagnetic term) and an operator which is quadratically dependent on ω_L (the diamagnetic term). Show that the new stationary states of the system and their degrees of degeneracy can be determined exactly.

c. Show that if ω_L is much smaller than ω_0, the effect of the diamagnetic term is negligible compared to that of the paramagnetic term.

d. We now consider the first excited state of the oscillator, that is, the states whose energies approach $5\hbar\omega_0/2$ when $\omega_L \longrightarrow 0$. To first order in ω_L/ω_0, what are the energy levels in the presence of the field $\mathbf{B}$ and their degrees of degeneracy (the Zeeman effect for a three-dimensional harmonic oscillator)? Same questions for the second excited state.

e. Now consider the ground state. How does its energy vary as a function of ω_L (the diamagnetic effect on the ground state)? Calculate the magnetic susceptibility χ of this state. Is the ground state, in the presence of the field $\mathbf{B}$, an eigenvector of $\mathbf{L}^2$? of L_z? of L_x? Give the form of its wave function and the corresponding probability current. Show that the effect of the field $\mathbf{B}$ is to compress the wave function about Oz (in a ratio $[1 + (\omega_L/\omega_0)^2]^{1/4}$) and to induce a current.

Absorption
and emission of photons, 2073
collision with, 971
of a quantum, a photon, 1311, 1353
of field, 2149
of several photons, 1368
rates, 1334
Acceptor (electron acceptor), 1495
Acetylene (molecule), 878
Action, 341, 1539, 1980
Addition
of angular momenta, 1015, 1043
of spherical harmonics, 1059
of two spins 1/2, 1019
Adiabatic
branching of the potential, 932
Adjoint
matrix, 123
operator, 112
Algebra (commutators), 165
Allowed energy band, 381, 1481, 1491
Ammonia (molecule), 469, 873
Amplitude
scattering amplitude, 929, 953
Angle (quantum), 2258
Angular momentum
addition of momenta, 1015, 1043
and rotations, 717
classical, 1529
commutation relations, 669, 725
conservation, 668, 736, 1016
coupling, 1016
electromagnetic field, 1968, 2043
half-integral, 987
of identical particles, 1497(ex.)
of photons, 1370
orbital, 667, 669, 685
quantization, 394
quantum, 667
spin, 987, 991
standard representation, 677, 691
two coupled momenta, 1091
Anharmonic oscillator, 502, 1135
Annihilation operator, 504, 513, 514, 1597
Annihilation-creation (pair), 1831, 1878
Anomalous
average value, 1828, 1852
dispersion, 2149
Zeeman effect, 987
Anti-normal correlation function, 1782, 1789
Anti-resonant term, 1312
Anti-Stokes (Raman line), 532, 752
Antibunching (photon), 2121
Anticommutation, 1599
field operator, 1754
Anticrossing of levels, 415, 482
Antisymmetric ket, state, 1428, 1431
Antisymmetrizer, 1428, 1431
Applications of the perturbation theory, 1231
Approximation
central field approximation, 1459
secular approximation, 1374
Argument (EPR), 2205
Atom(s), *see* helium, hydrogenoid
donor, 837
dressed, 2129, 2133
many-electron atoms, 1459, 1467
mirrors for atoms, 2153
muonic atom, 541
single atom fluorescence, 2121
Atomic
beam (deceleration), 2025
orbital, 869, 1496(ex.)
parameters, 41
Attractive bosons, 1747
Autler-Townes
doublet, 2144
effect, 1410
Autoionization, 1468
Average value (anomalous), 1828
Azimuthal
quantum number, 811

Band (energy), 381
Bardeen-Cooper-Schrieffer, 1889
Barrier (potential barrier), 68, 367, 373
Basis

Quantum Mechanics, Volume I, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

change of bases, 174
characteristic relations, 101, 119
continuous basis in the space of states, 99
mixed basis in the space of states, 99
BCHSH inequalities, 2209, 2210
BCS, 1889
broken pairs and excited pairs, 1920
coherent length, 1909
distribution functions, 1899
elementary excitations, 1923
excited states, 1919
gap, 1894, 1896, 1923
pairs (wave function of), 1901
phase locking, 1893, 1914, 1916
physical mechanism, 1914
two-particle distribution, 1901
Bell's
inequality, 2208
theorem, 2204, 2208
Benzene (molecule), 417, 495
Bessel
Bessel-Parseval relation, 1507
spherical Bessel function, 944
spherical equation, 961
spherical function, 966
Biorthonormal decomposition, 2194
Bitter, 2059
Blackbody radiation, 651
Bloch
equations, 463, 1358, 1361
theorem, 659
Bogolubov
excitations, 1661
Hamiltonian, 1952
operator method, 1950
phonons, spectrum, 1660
transformation, 1950
Bogolubov-Valatin transformation, 1836, 1919
Bohr, 2207
electronic magneton, 856
frequencies, 249
magneton, *see* front cover pages
model, 40, 819
nuclear magneton, 1237
radius, 820
Boltzmann
constant, *see* front cover pages
distribution, 1630
Born
approximation, 938, 977, 1320
Born-Oppenheimer approximation, 528, 1177, 1190
Born-von Karman conditions, 1490
Bose-Einstein
condensation, 1446, 1638, 1940
condensation (repulsive bosons), 1933
condensation of pairs, 1857
distribution, 652, 1630
statistics, 1446
Bosons, 1434
at non-zero temperature, 1745
attractive, 1747
attractive instability, 1745
condensed, 1638
in a Fock state, 1775
paired, 1881
Boundary conditions (periodic), 1489
Bra, 103, 104, 119
Bragg reflection, 382
Brillouin
formula, 452
zone, 614
Broadband
detector, 2165
optical excitation, 1332
Broadening (radiative), 2138
Broken pairs and excited pairs (BCS), 1920
Brossel, 2059
Bunching of bosons, 1777

C.S.C.O., 133, 137, 153, 236
Canonical
commutation relations, 142, 223, 1984
ensemble, 2289
Hamilton-Jacobi canonical equations, 214
Hamilton-Jacobi equations, 1532
Cauchy principal part, 1517
Center of mass, 812, 1528
Center of mass frame, 814
Central

field approximation, 1459
potential, 1533
Central potential, 803, 841
scattering, 941
stationary states, 804
Centrifugal potential, 809, 888, 893
Chain (von Neumann), 2201
Chain of coupled harmonic oscillators, 611
Change
of bases, 124, 174, 1601
of representation, 124
Characteristic equation, 129
Characteristic relation of an orthonormal basis, 116
Charged harmonic oscillator in an electric field, 575
Charged particle
in an electromagnetic field, 1536
Charged particle in a magnetic field, 240, 321, 771
Chemical bond, 417, 869, 1189, 1210
Chemical potential, 1486, 2287
Circular quanta, 761, 783
Classical
electrodynamics, 1957
histories, 2272
Clebsch-Gordan coefficients, 1038, 1051
Closure relation, 93, 117
Coefficients
Clebsch-Gordan, 1038
Einstein, 1334, 2083
Coherences (of the density matrix), 307
Coherent length (BCS), 1909
Coherent state (field), 2008
Coherent superposition of states, 253, 301, 307
Collision, 923
between identical particles, 1454, 1497(ex.)
between identical particles in classical mechanics, 1420
between two identical particles, 1450
cross section, 926
scattering states, 928
total scattering cross section, 926
with absorption, 971
Combination
of atomic orbitals, 1172
Commutation, 1599
canonical relations, 142, 223
field operator, 1754
of pair field operators, 1861
relations, 1984
Commutation relations
angular momentum, 669, 725
field, 1989, 1996
Commutator algebra, 165
Commutator(s), 91, 167, 171, 187
of functions of operators, 168
Compatibility of observables, 232
Complementarity, 45
Complete set of commuting observables (C.S.C.O.), 133, 137, 236
Complex variables (Lagrangian), 1982
Compton wavelength of the electron, 825, 1235
Condensates
relative phase, 2237
with spins, 2254
Condensation
BCS condensation energy, 1917
Bose-Einstein, 1446, 1857, 1933
Condensed bosons, 1638
Conduction band, 1492
Conductivity (solid), 1492
Configurations, 1467
Conjugate momentum, 214, 323, 1531, 1983, 1987, 1995
Conjugation (Hermitian), 111
Conservation
local conservation of probability, 238
of angular momentum, 668, 736, 1016
of energy, 248
of probability, 237
Conservative systems, 245, 315
Constants of the motion, 248, 317
Contact term, 1273
Contact term (Fermi), 1238, 1247
Contextuality, 2231
Continuous
spectrum, 133, 219, 264, 1316
variables (in a Lagrangian), 1984
Continuum of final states, 1316, 1378, 1380
Contractions, 1802

Convolution product of two functions, 1510
Cooling
Doppler, 2026
down atoms, 2025
evaporative, 2034
Sisyphus, 2034
sub-Doppler, 2155
subrecoil, 2034
Cooper model, 1927
Cooper pairs, 1927
Cooperative effects (BCS), 1916
Correlation functions, 1781, 1804
anti-normal, 1782, 1789
dipole and field, 2113
for one-photon processes, 2084
normal, 1782, 1787
of the field, spatial, 1758
Correlations, 2231
between two dipoles, 1157
between two physical systems, 296
classical and quantum, 2221
introduced by a collision, 1104
Coulomb
field, 1962
gauge, 1965
Coulomb potential
cross section, 979
Coupling
between angular momenta, 1016
between two angular momenta, 1091
between two states, 412
effect on the eigenvalues, 438
spin-orbit coupling, 1234, 1241
Creation and annihilation operators, 504, 513, 514, 1596, 1990
Creation operator (pair of particles), 1813, 1846
Critical velocity, 1671
Cross section
and phase shifts, 951
scattering cross section, 926, 933, 953, 972
Current
metastable current in superfluid, 1667
of particles, 1758
of probability, 240
probability current in hydrogen atom, 851
Cylindrical symmetry, 899(ex.)

Darwin term, 1235, 1279
De Broglie
relation, 10
wavelength, *see* front cover pages, 11, 35
Decay of a discrete state, 1378
Deceleration of an atomic beam, 2025
Decoherence, 2199
Decomposition (Schmidt), 2193
Decoupling (fine or hyperfine structure), 1262, 1291
Degeneracy
essential, 811, 825, 845
exchange degeneracy, 1423
exchange degeneracy removal, 1435
lifted by a perturbation, 1125
rotation invariance, 1072
systematic and accidental, 203
Degenerate eigenvalue, 127, 203, 217, 260
Degereracy
lifted by a perturbation, 1117
parity, 199
Delta Dirac function, 1515
potential well and barriers, 83–85(ex.)
use in quantum mechanics, 97, 106, 280
Density
Lagrangian, 1986
of probability, 264
of states, 389, 1316, 1484, 1488
operator, 449, 1391
operator and matrix, 299
particle density operator, 1756
Density functions
one and two-particle, 1502(ex.)
Depletion (quantum), 1940
Derivative of an operator, 169
Detection probability amplitude (photon), 2166
Detectors (photon), 2165
Determinant
Slater determinant, 1438, 1679
Deuterium, 834, 1107(ex.)
Diagonalization

of a 2×2 matrix, 429
of an operator, 128
Diagram (dressed-atom), 2133
Diamagnetism, 855
Diatomic molecules
rotation, 739
Diffusion (momentum), 2030
Dipole
-dipole interaction, 1142, 1153
-dipole magnetic interaction, 1237
electric dipole transition, 863
electric moment, 1080
Hamiltonian, 2011
magnetic dipole moment, 1084
magnetic term, 1272
trap, 2151
Dirac, *see* Fermi
delta function, 97, 106, 280, 1515
equation, 1233
notation, 102
Direct
and exchange terms, 1613, 1632, 1634, 1646, 1650
term, 1447, 1453
Discrete
bases of the state space, 91
spectrum, 132, 217
Dispersion (anomalous), 2149
Dispersion and absorption (field), 2147
Distribution
Boltzmann, 1630
Bose-Einstein, 1630
Fermi-Dirac, 1630
function (bosons), 1629
function (fermions), 1629
functions, 1625, 1733
functions (BCS), 1899
Distribution law
Bose-Einstein, 652
Divergence (energy), 2007
Donor atom, 837, 1495
Doppler
cooling, 2026
effect, 2022
effect (relativistic), 2022
free spectroscopy, 2105
temperature, 2033
Double
condensate, 2237
resonance method, 2059
spin condensate, 2254
Doublet (Autler-Townes), 2144
Down-conversion (parametric), 2181
Dressed
states and energies, 2133
Dressed-atom, 2129, 2133
diagram, 2133
strong coupling, 2141
weak coupling, 2137

E.P.R., 1225(ex.)
Eckart (Wigner-Eckart theorem), *see* Wigner
Effect
Autler-Townes, 2144
Mössbauer, 2040
photoelectric, 2110
Effective Hamiltonian, 2141
Ehrenfest theorem, 242, 319, 522
Eigenresult, 9
Eigenstate, 217, 232
Eigenvalue, 11, 25, 176, 216
degenerate, 217, 260
equation, 126, 429
of an operator, 126
Eigenvector, 176
of an operator, 126
Einstein, 2110
coefficients, 1334, 1356, 2083
EPR argument, 297, 1104
model, 534, 653
Planck-Einstein relations, 3
temperature, 659
Einstein-Podolsky-Rosen, 2204, 2261
Elastic
scattering, 925
scattering (photon), 2086
scattering, form factor, 1411(ex.)
total cross section, 972
Elastically bound electron model, 1350
Electric
conductivity of a solid, 1492
Electric dipole
Hamiltonian, 2011
interaction, 1342

matrix elements, 1344
moment, 1080
selection rules, 1345
transition and selection rules, 863
transitions, 2056
Electric field (quantized), 2000, 2005
Electric polarisability
NH_3, 484
Electric polarizability
of the 1s state in Hydrogen, 1299
Electric quadrupole
Hamiltonian, 1347
moment, 1082
transitions, 1348
Electric susceptibility
bound electron, 577
of an atom, 1351
Electrical
susceptibility, 1223(ex.)
Electrodynamics
classical, 1957
quantum, 1997
Electromagnetic field
and harmonic oscillators, 1968
and potentials, 321
angular momentum, 1968, 2043
energy, 1966
Lagrangian, 1986, 1992
momentum, 1967, 2019
polarization, 1970
quantization, 631, 637
Electromagnetic interaction of an atom with a wave, 1340
Electromagnetism
fields and potentials, 1536
Electron spin, 393, 985
Electron(s)
configurations, 1463
gas in solids, 1491
in solids, 1177, 1481
mass and charge, *see* front cover pages
Electronic
configuration, 1459
paramagnetic resonance, 1225(ex.)
shell, 827
Elements of reality, 2205
Emergence of a relative phase, 2248, 2253
Emission
of a quantum, 1311
photon, 2080
spontaneous, 2081, 2135
stimulated (or induced), 2081
Energy, *see* Conservation, Uncertainty
and momentum of the transverse electromagnetic field, 1973
band, 381
bands in solids, 1177, 1481
conservation, 248
electromagnetic field, 1966
Fermi energy, 1772
fine structure energy levels, 986
free energy, 2290
levels, 359
levels of harmonic oscillator, 509
levels of hydrogen, 823
of a paired state, 1869
recoil energy, 2023
Ensemble
canonical, 2289
grand canonical, 2291
microcanonical, 2285
statistical ensembles, 2295
Entanglement
quantum, 2187, 2193, 2203, 2242
swapping, 2232
Entropy, 2286
EPR, 2204, 2261
elements of reality, 2205
EPRB, 2205
paradox/argument, 1104
Equation of state
ideal quantum gas, 1640
repulsive bosons, 1745
Equation(s)
Bloch, 1361
Hamilton-Jacobi, 1982, 1983, 1988
Lagrange, 1982, 1993
Lorentz, 1959
Maxwell, 1959
Schrödinger, 11, 12, 306
von Neumann, 306
Essential degeneracy, 811, 825
Ethane (molecule), 1223
Ethylene (molecule), 536, 881

Evanescent wave, 29, 67, 70, 78, 285
Evaporative cooling, 2034
Even operators, 196
Evolution
  field operator, 1765
  of quantum systems, 223
  of the mean value, 241
  operator, 313, 2069
  operator (expansion), 2070
  operator (integral equation), 2069
Exchange, 1611
  degeneracy, 1423
  degeneracy removal, 1435
  energy, 1469
  hole, 1774
  integral, 1474
  term, 1447, 1451, 1453
Excitations
  BCS, 1923
  Bogolubov, 1661
  vacuum, 1623
Excited states (BCS), 1919
Exciton, 838
Exclusion principle (Pauli), 1437, 1444, 1463, 1484
Extensive (or intensive) variables, 2292

Fermi
  contact term, 1238
  energy, 1445, 1481, 1486, 1772
  gas, 1481
  golden rule, 1318
  level, 1486, 1621
  radius, 1621
  surface (modified), 1914
  , *see* Fermi-Dirac
Fermi level
  and electric conductivity, 1492
Fermi-Dirac
  distribution, 1486, 1630, 1717
  statistics, 1446
Fermions, 1434
  in a Fock state, 1771
  paired, 1874
Ferromagnetism, 1477
Feynman
  path, 2267
  postulates, 341
Fictitious spin, 435, 1359
Field
  absorption, 2149
  commutation relations, 1989, 1996
  dispersion and absorption, 2147
  intense laser, 2126
  interaction energy, 1764
  kinetic energy, 1763
  normal variables, 1971
  operator, 1752
  operator (evolution), 1763, 1765
  pair field operator, 1861
  potential energy, 1764
  quantization, 1765, 1999
  quasi-classical state, 2008
  spatial correlation functions, 1758
Final states continuum, 1378, 1380
Fine and hyperfine structure, 1231
Fine structure
  constant, *see* front cover pages, 825
  energy levels, 1478
  Hamiltonian, 1233, 1276, 1478
  Helium atom, 1478
  Hydrogen, 1238
  of spectral lines, 986
  of the states $1s$, $2s$ et $2p$, 1276
Fletcher, 2111
Fluctuations
  boson occupation number, 1633
  intensity, 2125
  vacuum, 644, 2007
Fluorescence (single atom), 2121
Fluorescence triplet, 2144
Fock
  space, 1593, 2004
  state, 1593, 1614, 1769, 2103
Forbidden, *see* Band
  energy band, 381, 390, 1481
  transition, 1345
Forces
  van der Waals, 1151
Form factor
  elastic scattering, 1411(ex.)
Forward scattering (direct and exchange), 1874
Fourier

series and transforms, 1505
Fragmentation (condensate), 1654, 1776
Free
electrons in a box, 1481
energy, 2290
particle, 14
quantum field (Fock space), 2004
spherical wave, 941, 944, 961
spherical waves and plane waves, 967
Free particle
stationary states with well-defined angular momentum, 959
stationary states with well-defined momentum, 19
wave packet, 14, 57, 347
Frequency
Bohr, 249
components of the field (positive and negative), 2072
Rabi's frequency, 1325
Friction (coefficient), 2028
Function
of operators, 166
periodic functions, 1505
step functions, 1521
Fundamental state, 41

Gap (BCS), 1894, 1896, 1923
Gauge, 1343, 1536, 1960, 1963
Coulomb, 1965
invariance, 321
Lorenz, 1965
Gaussian
wave packet, 57, 292, 2305
Generalized velocities, 214, 1530
Geometric quantization, 2311
Gerlach, *see* Stern
GHZ state, 2222, 2227
Gibbs-Duhem relation, 2296
Golden rule (Fermi), 1318
Good quantum numbers, 248
Grand canonical, 1626, 2291
Grand potential, 1627, 1721, 2292
Green's function, 337, 936, 1781, 1786, 1789
evolution, 1785
Greenberger-Horne-Zeilinger, 2227
Groenewold's formula, 2315
Gross-Pitaevskii equation, 1643, 1657
Ground state, 363
harmonic oscillator, 509, 520
Hydrogen atom, 1228(ex.)
Group velocity, 55, 60, 614
Gyromagnetic ratio, 396, 455
orbital, 860
spin, 988

H_2^+ molecular ion, 85(ex.), 417, 1189
Hadronic atoms, 840
Hall effect, 1493
Hamilton
function, 1532
function and equations, 1531
Hamilton-Jacobi canonical equations, 214, 1532, 1982, 1983, 1988
Hamiltonian, 223, 245, 1527, 1983, 1988, 1995
classical, 1531
effective, 2141
electric dipole, 1342, 2011
electric quadrupole, 1347
fine structure, 1233, 1276
hyperfine, 1237, 1267
magnetic dipolar, 1347
of a charged particle in a vector potential, 1539
of a particle in a central potential, 806, 1533
of a particle in a scalar potential, 225
of a particle in a vector potential, 225, 323, 328
Hanbury Brown and Twiss, 2120
Hanle effect, 1372(ex.)
Hard sphere
scattering, 980, 981(ex.)
Harmonic oscillator, 497
in an electric field, 575
in one dimension, 527, 1131
in three dimensions, 569
in two dimensions, 755
infinite chain of coupled oscillators, 611
quasiclassical states, 583
thermodynamic equilibrium, 647

three-dimensional, 841, 899(ex.)
two coupled oscillators, 599
Hartree-Fock
approximation, 1677, 1701
density operator (one-particle), 1691
equations, 1686, 1731
for electrons, 1695
mean field, 1677, 1693
potential, 1706
thermal equilibrium, 1711, 1733
time-dependent, 1701, 1708
Healing length, 1652
Heaviside step function, 1521
Heisenberg
picture, 317, 1763
relations, 19, 39, 41, 45, 55, 232, 290
Helicity (photon), 2051
Helium
energy levels, 1467
ion, 838
isotopes, 1480
isotopes ^{3}He and ^{4}He, 1435, 1446
solidification, 535
Hermite polynomials, 516, 547, 561
Hermitian
conjugation, 111
matrix, 124
operator, 115, 124, 130
Histories (classical), 2272
Hole
creation and annihilation, 1622
exchange, 1774
Holes, 1621
Hybridization of atomic orbitals, 869
Hydrogen, 645
atom, 803
atom in a magnetic field, 853, 855, 862
atom, relativistic energies, 1245
Bohr model, 40, 819
energy levels, 823
fine and hyperfine stucture, 1231
ionisation energy, *see* front cover pages
ionization energy, 820
maser, 1251
molecular ion, 85(ex.), 417, 1189
quantum theory, 41
radial equation, 821
Stark effect, 1298
stationary states, 851
stationary wave functions, 830
Hydrogen-like systems in solid state physics, 837
Hydrogenoid systems, 833
Hyperfine
decoupling, 1262
Hamiltonian, 1237, 1267
Hyperfine structure, *see* Hydrogen, muonium, positronium, Zeeman effect, 1231
Muonium, 1281

Ideal gas, 1625, 1787, 1791, 1804
correlations, 1769
Identical particles, 1419, 1591
Induced
emission, 1334, 1366, 2081
emission of a quantum, 1311
emission of photons, 1355
Inequality (Bell's), 2208
Infinite one-dimensional well, 271
Infinite potential well, 74
in two dimensions, 201
Infinitesimal unitary operator, 178
Insulator, 1492
Integral
exchange integral, 1474
scattering equation, 935
Intense laser fields, 2126
Intensive (or extensive) variables, 2292
Interaction
between magnetic dipoles, 1141
dipole-dipole interaction, 1141, 1153
electromagnetic interaction of an atom with a wave, 1340
field and particles, 2009
field and atom, 2010
magnetic dipole-dipole interaction, 1237
picture, 353, 1393, 2070
tensor interaction, 1141
Interference
photons, 2167
two-photon, 2170, 2183
Ion H_2^+, 1189

Ionization
photo-ionization, 2109
tunnel ionization, 2126
Isotropic radiation, 2079

Jacobi, *see* Hamilton

Kastler, 2059, 2062
Ket, *see* state, 103, 119
for identical particles, 1436
Kuhn, *see* Thomas

Lagrange
equations, 1530, 1982, 1993
fonction and equations, 214
multipliers, 2281
Lagrangian, 1530, 1980
densities, 1986
electromagnetic field, 1986, 1992
formulation of quantum mechanics, 339
of a charged particle in an electromagnetic field, 1538
particle in an electromagnetic field, 323
Laguerre-Gaussian beams, 2065
Lamb shift, 645, 1245, 1388, 2008
Landau levels, 771
Landé factor, 1072, 1107(ex.), 1256, 1292
Laplacian, 1527
of $1/r$, 1524
of $Y_l^m(\theta,\varphi)/r^{l+1}$, 1526
Larmor
angular frequency, 857
precession, 394, 396, 410, 455, 857, 1071
Laser, 1359, 1365
Raman laser, 2093
saturation, 1370
trap, 2151
Lattices (optical), 2153
Least action
principle of, 1539
Legendre
associated function, 714
polynomial, 713
Length (healing), 1652
Level
anticrossing, 415, 482
Fermi level, 1621
Lifetime, 343, 485, 645
of a discrete state, 1386
radiative, 2081
Lifting of degeneracy by a perturbation, 1125
Light
quanta, 3
shifts, 1334, 2138, 2151, 2156
Linear, *see* operator
combination of atomic orbitals, 1172
operators, 90, 108, 163
response, 1350, 1357, 1364
superposition of states, 253
susceptibility, 1365
Local conservation of probability, 238
Local realism, 2209, 2230
Longitudinal
fields, 1961
relaxation, 1400
relaxation time, 1401
Lorentz equations, 1959
Lorenz (gauge), 1965

Magnetic
dipole term, 1272
dipole-dipole interaction, 1237
effect of a magnetic field on the levels of the Hydrogen atom, 1251
hyperfine Hamiltonian, 1267
interactions, 1232, 1237
quantum number, 811
resonance, 455
susceptibility, 1224, 1487
Magnetic dipole
Hamiltonian, 1347
transitions and selection rules, 1084, 1098, 1348
Magnetic dipoles
interactions between two dipoles, 1141
Magnetic field
and vector potential, 321
charged particle in a, 240, 771
effects on hydrogen atom, 853, 855
harmonic oscillator in a, 899(ex.)
Hydrogen atom in a magnetic field, 1263, 1289

multiplets, 1074
quantized, 2000, 2005
Magnetism (spontaneous), 1737
Many-electron atoms, 1459
Maser, 477, 1359, 1365
hydrogen, 1251
Mass correction (relativistic), 1234
Master equation, 1358
Matrice(s), 119, 121
diagonalization of a 2×2 matrix, 429
Pauli matrices, 425
unitary matrix, 176
Maxwell's equations, 1959
Mean field (Hartree-Fock), 1693, 1708, 1725
Mean value of an observable, 228
evolution, 241
Measurement
general postulates, 216, 226
ideal von Neumann measurement, 2196
of a spin 1/2, 394
of observables, 216
on a part of a physical system, 293
state after measurement, 221, 227
Mendeleev's table, 1463
Metastable superfluid flow, 1671
Methane (molecule), 883
Microcanonical ensemble, 2285
Millikan, 2111
Minimal wave packet, 290, 520, 591
Mirrors for atoms, 2153
Mixing of states, 1121, 1137
Model
Cooper model, 1927
Einstein model, 534
elastically bound electron, 1350
vector model of atom, 1071
Modes
vibrational modes, 599, 611
Modes (radiation), 1974, 1975
Molecular ion, 417
Molecule(s)
chemical bond, 417, 869, 873, 878, 883, 1189
rotation, 796
vibration, 527, 1137
vibration-rotation, 885
Mollow, 2144
Moment
quadrupole electric moment, 1225(ex.)
Momentum, 1539
conjugate, 214, 323, 1983, 1987, 1995
diffusion, 2030
electromagnetic field, 1967, 2019
mechanical momentum, 328
Monogamy (quantum), 2221
Mössbauer effect, 1415, 2040
Motional narrowing, 1323
condition, 1323, 1398, 1408
Multiphoton transition, 1368, 2040, 2097
Multiplets, 1072, 1074, 1467
Multipliers (Lagrange), 2281
Multipolar waves, 2052
Multipole
moments, 1077
Multipole operators
introduction, 1077, 1083
parity, 1082
Muon, 527, 541, 1281
Muonic atom, 541, 839
Muonium, 835
hyperfine structure, 1281
Zeeman effect, 1281

Narrowing (motional), 1323, 1408
condition, 1398
Natural width, 345, 1388
Need for a quantum treatment, 2118, 2120
Neumann
spherical function, 967
Neutron mass, *see* front cover pages
Non-destructive detection of a photon, 2159
Non-diagonal order (BCS), 1912
Non-locality, 2204
Non-resonant excitation, 1350
Non-separability, 2207
Nonlinear
response, 1357, 1368
susceptibility, 1369
Norm
conservation, 238
of a state vector, 104, 237
of a wave function, 13, 90, 99

Normal
correlation function, 1782, 1787
variables, 602, 616, 631, 633
variables (field), 1971
Nuclear
multipole moments, 1088
Bohr magneton, 1237
Nucleus
spin, 1088
volume effect, 1162, 1268
Number
occupation number, 1439, 1593
photon number, 2135
total number of particles in an ideal gas, 1635

Observable(s), 130
C.S.C.O., 133, 137
commutation, 232
compatibility, 232
for identical particles, 1429, 1441
mean value, 228
measurement of, 216, 226
quantization rules, 223
symmetric observables, 1441
transformation by permutation, 1434
whose commutator is $i\hbar$, 187, 289
Occupation number, 1439, 1593
operator, 1598
Odd operators, 196
One-particle
Hartree-Fock density operator, 1691
operators, 1603, 1605, 1628, 1756
Operator(s)
adjoint operator, 112
annihilation operator, 504, 513, 514, 1597
creation and annihilation, 1990
creation operator, 504, 513, 514, 1596
derivative of an operator, 169
diagonalization, 126, 128
even and odd operators, 196
evolution operator, 313, 2069
field, 1752
function of, 166
Hermitian operators, 115
linear operators, 90, 108, 163
occupation number, 1598
one-particle operator, 1603, 1605, 1628, 1756
parity operator, 193
particle density operator, 1756
permutation operators, 1425, 1430
potential, 168
product of, 90
reduced to a single particle, 1607
representation, 121
restriction, 165
restriction of, 1125
rotation operator, 1001
symmetric, 1628, 1755
translation operator, 190
two-particle operator, 1608, 1610, 1631, 1756
unitary operators, 173
Weyl operator, 2300
Oppenheimer, *see* Born, 1177, 1190
Optical
excitation (broadband), 1332
lattices, 2153
pumping, 2062, 2140
Orbital
angular momentum (of radiation), 2052
atomic orbital, 1496(ex.)
hybridization, 869
linear combination of atomic orbitals, 1172
quantum number, 1463
state space, 988
Order parameter for pairs, 1851
Orthonormal basis, 91, 99, 101, 133
characteristic relation, 116
Orthonormalization
and closure relations, 101, 140
relation, 116
Oscillation(s)
between two discrete states, 1374
between two quantum states, 418
Rabi, 2134
Oscillator
anharmonic, 502
harmonic, 497
strength, 1352

Pair(s)

annihilation-creation of pairs, 1831, 1874, 1887
BCS, wave function, 1909
Cooper, 1927
of particles (creation operator), 1813, 1846
pair field (commutation), 1861
pair field operator, 1845
pair wave function, 1851
Paired
bosons, 1881
fermions, 1874
state energy, 1869
states, 1811
states (building), 1818
Pairing term, 1878
Paramagnetism, 855
Parametric down-conversion, 2181
Parity, 2106
degeneracy, 199
of a permutation operator, 1431
of multipole operators, 1082
operator, 193
Parseval
Parseval-Plancherel equality, 20
Parseval-Plancherel formula, 1511, 1521
Partial
reflection, 79
trace of an operator, 309
waves in the potential, 948
waves method, 941
Particle (current), 1758
Particles and holes, 1621
Partition function, 1626, 1627, 1717
Path
integral, 2267
space-time path, 339
Pauli
exclusion principle, 1437, 1444, 1463, 1481
Hamiltonian, 1009(ex.)
matrices, 425, 991
spin theory, 986
spinor, 993
Penetrating orbit, 1463
Penrose-Onsager criterion, 1776, 1860, 1947
Peres, 2212
Periodic
boundary conditions, 1489
classification of elements, 1463
functions, 1505
potential (one-dimensional), 375
Permutation operators, 1425, 1430
Perturbation
applications of the perturbation theory, 1231
lifting of a degeneracy, 1125
one-dimensional harmonic oscillator, 1131
random perturbation, 1320, 1325, 1390
sinusoidal, 1311
stationary perturbation theory, 1115
Perturbation theory
time dependent, 1303
Phase
locking (BCS), 1893, 1916
locking (bosons), 1938, 1944
relative phase between condensates, 2237, 2248
velocity, 37
Phase shift (collision), 951, 1497(ex.)
with imaginary part, 971
Phase velocity, 21
Phonons, 611, 626
Bogolubov phonons, 1660
Photodetection
double, 2172, 2184
single, 2169, 2171
Photoelectric effect, 1412(ex.), 2110
Photoionization, 2109, 2165
rate, 2115, 2124
two-photon, 2123
Photon, 3, 631, 651, 2004, 2005, 2110
absorption and emission, 2067
angular momentum, 1370
antibunching, 2121
detectors, 2165
non-destructive detection, 2159
number, 2135
scattering (elastic), 2086
scattering by an atom, 2085
vacuum, 2007
, *see* Absorption, Emission
Picture

Heisenberg, 317, 1763
interaction, 1393, 2070
Pitaevskii (Gross-Pitaevskii equation), 1643, 1657
Plancherel, *see* Parseval
Planck
constant, *see* front cover pages, 3
law , 2083
Planck-Einstein relations, 3, 10
Plane wave, 14, 19, 95, 943
Podolsky (EPR argument), 297, 1104
Pointer states, 2199
Polarizability
of the 1s state in Hydrogen, 1299
Polarization
electromagnetic field, 1970
of Zeeman components, 1295
space-dependent, 2156
Polynomial method (harmonic oscillator), 555, 842
Polynomials
Hermite polynomials, 516, 547, 561
Position and momentum representations, 181
Positive and negative frequency components, 2072
Positron, 1281
Positronium, 836
hyperfine structure, 1281
Zeeman effect, 1281
Postulate (von Neumann projection), 2202
Postulates of quantum mechanics, 215
Potential
adiabatic branching, 932
barrier, 26, 68, 367, 373
centrifugal potential, 809, 888, 893
Coulomb potential, cross section, 979
cylindrically symmetric, 899(ex.)
Hartree-Fock, 1706
infinite one-dimensional well, 74
operator, 168
scalar and vector potentials, 1536, 1960, 1963
scattering by a, 923
self-consistent potential, 1461
square potential, 63
square well, 29
step, 28, 65, 75, 284
well, 71, 367
well (arbitrary shape), 359
well (infinite one-dimensional), 271
well (infinite two-dimensional, 201
Yukawa potential, 977
Precession
Larmor precession, 396, 1071
Thomas precession, 1235
Preparation of a state, 235
Pressure (ideal quantum gas), 1640
Principal part, 1517
Principal quantum number, 827
Principle
of least action, 1539, 1980
of spectral decomposition, 11, 216
of superposition, 237
Probability
amplitude, 11, 253, 259
conservation, 237
current, 240, 283, 333, 349, 932
current in hydrogen atom, 851
density, 11, 264
fluid, 932
of photon absorption, 2076
of the measurement results, 9, 11
transition probability, 439
Process (pair annihilation-creation), 1878, 1887
Product
convolution product of functions, 1510
of matrices, 122
of operators, 90
scalar product, 101, 141, 149, 161
state (tensor product), 311
tensor product, 147
tensor product, applications, 441
Projection theorem, 1070
Projector, 109, 133, 165, 218, 222, 1108(ex.)
Propagator
for the Schrödinger equation, 335
of a particle, 2267, 2272
Proper result, 9
Proton
mass, *see* front cover pages
spin and magnetic moment, 1237, 1274
Pumping, 1358

Pure (state or case), 301

Quadrupolar electric moment, 1082, 1225(ex.)
Quanta (circular), 761, 783
Quantization
  electrodynamics, 1997
  electromagnetic field, 631, 637, 1997
  of a field, 1765
  of angular momentum, 394, 677
  of energy, 3, 11, 71, 359
  of measurement results, 9, 216, 398
  of the measurement results, 405
  rules, 11, 223, 226, 2274
Quantum
  angle, 2258
  electrodynamics, 1245, 1282, 1997
  entanglement, 2187, 2193
  monogamy, 2221
  number
    orbital, 1463
    principal quantum number, 827
  numbers (good), 248
  resonance, 417
  treatment needed, 2118, 2120
Quasi-classical
  field states, 2008
  states, 765, 791, 801
  states of the harmonic oscillator, 583
Quasi-particles, 1736, 1840
  Bogolubov phonons, 1954
  Quasi-particle vacuum, 1836

Rabi
  formula, 440, 460, 1324, 1376
  formula), 419
  frequency, 1325
  oscillation, 2134
Radial
  equation, 842
  equation (Hydrogen), 821
  equation in a central potential, 808
  integral, 1277
  quantum number, 811
Radiation
  isotropic, 2079
  pressure, 2024
Radiative
  broadening, 2138
  cascade of the dressed atom, 2145
Raman
  effect, 532, 740, 1373(ex.)
  laser, 2093
  scattering, 2091
  scattering (stimulated), 2093
Random perturbation, 1320, 1325, 1390
Rank (Schmidt), 2196
Rate (photoionization), 2115, 2124
Rayleigh
  line, 752
  scattering, 532, 2089
Realism (local), 2205, 2209
Recoil
  blocking, 2036
  effect of the nucleus, 834
  energy, 1415, 2023
  free atom, 2020
  suppression, 2040
Reduced
  density operator, 1607
  mass, 813
Reduction of the wave packet, 221, 279
Reflection on a potential step, 285
Refractive index, 2149
Reiche, *see* Thomas
Relation (Gibbs-Duhem), 2296
Relative
  motion, 814
  particle, 814
  phase between condensates, 2248, 2258
  phase between spin condensates, 2253
Relativistic
  corrections, 1233, 1478
  Doppler effect, 2022
  mass correction, 1234
Relaxation, 465, 1358, 1390, 1413, 1414(ex.)
  general equations, 1397
  longitudinal, 1400
  longitudinal relaxation time, 1401
  transverse, 1403
  transverse relaxation time, 1406
Relay state, 2086, 2098, 2106
Renormalization, 2007
Representation(s)
  change of, 124
  in the state space, 116

of operators, 121
position and momentum, 139, 181
Schrödinger equation, 183–185
Repulsion between electrons, 1469
Resonance
magnetic resonance, 455
quantum resonance, 417, 1158
scattering resonance, 69, 954, 983(ex.)
two resonnaces with a sinusoidal excitation, 1365
width, 1312
with sinusoidal perturbation, 1311
Restriction of an operator, 165, 1125
Rigid rotator, 740, 1222(ex.)
Ritz theorem, 1170
Root mean square deviation
general definition, 230
Rosen (EPR argument), 297, 1104
Rotating frame, 459
Rotation(s)
and angular momentum, 717
invariance and degeneracy, 734
of diatomic molecules, 739
of molecules, 796, 885
operator(s), 720, 1001
rotation invariance, 1478
rotation invariance and degeneracy, 1072
Rotator
rigid rotator, 740, 1222(ex.)
Rules
quantization rules, 2274
selection rules, 197
Rutherford's formula, 979
Rydberg constant, *see* front cover pages

Saturation
of linear response, 1368
of the susceptibility, 1369
Scalar
and vector potentials, 321, 1536
interaction between two angular momenta, 1091
observable, operator, 732, 737
potential, 225
product, 89, 92, 101, 141, 149, 161
product of two coherent states, 593
Scattering
amplitude, 929, 953
by a central potential, 941
by a hard sphere, 980, 981(ex.)
by a potential, 923
cross section, 933, 953, 972
cross section and phase shifts, 951
inelastic, 2091
integral equation, 935
of particles with spin, 1102
of spin 1/2 particles, 1108(ex.)
photon, 2086
Raman, 2091
Rayleigh, 532, 2089
resonance, 954, 983(ex.)
resonant, 2089
stationary scattering states, 951
stationary states, 928
stimulated Raman, 2093
Schmidt
decomposition, 2193
rank, 2196
Schottky anomaly, 654
Schrödinger, 2190
equation, 11, 12, 223, 306
equation in momentum representation, 184
equation in position representation, 183
equation, physical implications, 237
equation, resolution for conservative systems, 245
picture, 317
Schwarz inequality, 161
Second
quantization, 1766
harmonic generation, 1368
Secular approximation, 1316, 1374
Selection rules, 197, 863, 2014, 2056
electric quadrupolar, 1348
magnetic dipolar, 1098, 1348
Self-consistent potential, 1461
Semiconductor, 837, 1493
Separability, 2207, 2223
Separable density operator, 2223
Shell (electronic), 827
Shift

light shift, 2138
of a discrete state, 1387
Singlet, 1024, 1474
Sinusoidal perturbation, 1311, 1374
Sisyphus
cooling, 2034
effect, 2155
Slater determinant, 1438, 1679
Slowing down atoms, 2025
Solids
electronic bands, 1177
energy bands of electrons, 1491
energy bands of electrons in solids, 381
hydrogen-like systems in solid state physics, 837
Space (Fock), 1593
Space-dependent polarization, 2156
Space-time path, 339, 1539
Spatial correlations (ideal gas), 1769
Specific heat
of an electron gas, 1484
of metals, 1487
of solids, 653
two level system, 654
Spectral
decomposition principle, 7, 11, 216
function, 1795
terms, 1469
Spectroscopy (Doppler free), 2105
Spectrum
BCS elementary excitation, 1923
continuous, 219, 264
discrete, 132, 217
of an observable, 126, 216
Spherical
Bessel equation, 961
Bessel function, 944, 966
free spherical waves, 961
free wave, 944
Neumann function, 967
wave, 941
waves and plane waves, 967
Spherical harmonics, 689, 705
addition of, 1059
expression for $l = 0, 1, 2$, 709
general expression, 707
Spin
and magnetic moment of the proton, 1237
angular momentum, 987
electron, 985, 1289
fictitious, 435
gyromagnetic ratio, 396, 455, 988
nuclear, 1088
of the electron, 393
Pauli theory, 986, 988
quantum description, 985, 991
rotation operator, 1001
scattering of particles with spin, 1102
spin 1 and radiation, 2044, 2049, 2050
system of two spins, 441
Spin 1/2
density operator, 449
ensemble of, 1358
fictitious, 1359
interaction between two spins, 1141
preparation and measurement, 401
scattering of spin 1/2 particles, 1108(ex.)
Spin-orbit coupling, 1018, 1234, 1241, 1279
Spin-statistics theorem, 1434
Spinor, 993
rotation, 1005
Spontaneous
emission, 343, 645, 1301, 2081, 2135
emission of photons, 1356
magnetism of fermions, 1737
Spreading of a wave packet, 59, 348
Square
barrier of potential, 26, 68
potential, 26, 63, 75, 283
potential well, 71, 271
spherical well, 982(ex.)
Standard representation (angular momentum), 677, 691
Stark effect in Hydrogen atom, 1298
State(s), *see* Density operator
density of, 389, 1316, 1484, 1488
Fock, 1593, 1614, 1769, 2103
ground state, 363
mixing of states by a perturbation, 1121
orbital state space, 988
paired, 1811

pointer states, 2199
quasi-classical states, 583, 765, 791, 801
relay state, 2086, 2098, 2106
stable and unstable states, 485
state after measurement, 221
state preparation, 235
stationary, 63, 359, 375
stationary state, 24, 246
stationary states in a central potential, 804
unstable, 343
vacuum state, 1595
vector, 102, 215
Stationary
perturbation theory, 1115
phase condition, 18, 54
scattering states, 928, 951
states, 24, 63, 246, 359
states in a periodic potential, 375
states with well-defined angular momentum, 944, 959
states with well-defined momentum, 943
Statistical
entropy, 2217
mechanics (review of), 2285
mixture of states, 253, 299, 304, 450
Statistics
Bose-Einstein, 1446
Fermi-Dirac, 1446
Step
function, 1521
potential, 28, 65, 75, 284
Stern-Gerlach experiment, 394
Stimulated
(or induced) emission, 1334, 1366, 2081
Raman scattering, 2093
Stokes Raman line, 532, 752
Stoner (spontaneous magnetism), 1737
Strong coupling (dressed-atom), 2141
Subrecoil cooling, 2034
Sum rule (Thomas-Reiche-Kuhn), 1352
Superfluidity, 1667, 1674
Superposition
of states, 253
principle, 7, 237
principle and physical predictions, 253
Surface (modified Fermi surface), 1914
Susceptibility, *see* Linear, nonlinear, tensor
electric susceptibility of an atom, 1351
electrical susceptibility, 577, e1223
electrical susceptibility of NH_3, 484
magnetic susceptibility, 1224
tensor, 1224, 1410(ex.)
Swapping (entanglement), 2232
Symmetric
ket, state, 1428, 1431
observables, 1429, 1441
operators, 1603, 1605, 1608, 1610, 1628, 1631, 1755
Symmetrization
of observables, 224
postulate, 1434
Symmetrizer, 1428, 1431
System
time evolution of a quantum system, 223
two-level system, 435
Systematic
and accidental degeneracies, 203
degeneracy, 845

Temperature (Doppler), 2033
Tensor
interaction, 1141
product, 147, 441
product of operators, 149
product state, 295, 311
product, applications, 201
susceptibility tensor, 1224
Term
direct and exchange terms, 1613, 1632, 1634, 1646, 1650
pairing, 1878
spectral terms, 1467, 1469
Theorem
Bell, 2204, 2208
Bloch, 659
projection, 1070
Ritz, 1170
Wick, 1799, 1804

Wigner-Eckart, 1065, 1085, 1254
Thermal wavelength, 1635
Thermodynamic equilibrium, 308
harmonic oscillator, 647
ideal quantum gas, 1625
spin 1/2, 452
Thermodynamic potential (minimization), 1715
Thomas precession, 1235
Thomas-Reiche-Kuhn sum rule, 1352
Three-dimensional harmonic oscillator, 569, 841, 899(ex.)
Three-level system, 1409(ex.)
Three-photon transition, 1370
Time evolution of quantum systems, 223
Time-correlations (fluorescent photons), 2145
Time-dependent
Gross-Pitaevskii equation, 1657
perturbation theory, 1303
Time-energy uncertainty relation, 250, 279, 345, 1312, 1389
Torsional oscillations, 536
Torus (flow in a), 1667
Total
elastic scattering cross section, 972
reflection, 67, 75
scattering cross section (collision), 926
Townes
Autler-Townes effect, 1410
Trace
of an operator, 163
partial trace of an operator, 309
Transform (Wigner), 2297
Transformation
Bogolubov, 1950
Bogolubov-Valatin, 1836, 1919
Gauge, 1960
of observables by permutation, 1434
Transition, *see* Probability, Forbidden, Electric dipole, Magnetic dipole, Quadrupole electric dipole, 2056
magnetic dipole transition, 1098
probability, 439, 1308, 1321, 1355
probability per unit time, 1319
probability, spin 1/2, 460
three-photon transition, 1370
two-photon, 2097
virtual, 2100
Translation operator, 190, 579, 791
Transpositions, 1431
Transverse
fields, 1961
relaxation, 1403
relaxation time, 1406
Trap
dipolar, 2151
laser, 2151
Triplet, 1024, 1474
fluorescence triplet, 2144
Tunnel
effect, 29, 70, 365, 476, 540, 1177
ionization, 2126
Two coupled harmonic oscillators, 599
Two-dimensional
harmonic oscillator, 755
infinite potential well, 201
wave packets, 49
Two-level system, 393, 411, 435, 1357
Two-particle operators, 1608, 1610, 1631, 1756
Two-photon
absorption, 1373(ex.)
interference, 2170, 2183
transition, 1409(ex.), 2097

Uncertainty
relation, 19, 39, 41, 45, 232, 290
time-energy uncertainty relation, 1312
Uniqueness of the measurement result, 2201
Unitary
matrix, 125, 176
operator, 173, 314
transformation of operators, 177
Unstable states, 343

Vacuum
electromagnetism, 644, 2007
excitations, 1623
fluctuations, 2007
photon vacuum, 2007
quasi-particule vacuum, 1836
state, 1595
Valence band, 1493

Van der Waals forces, 1151
Variables
intensive or extensive, 2292
normal variables, 602, 616, 631, 633
Variational method, 1169, 1190, 1228(ex.)
Vector
model, 1091
model of the atom, 1071, 1256
observable, operator, 732
operator, 1065
potential, 225
potential of a magnetic dipole, 1268
Velocity
critical, 1671
generalized velocities, 214, 1530
group velocity, 23, 614
phase velocity, 21, 37
Vibration(s)
modes, 599, 611
modes of a continuous system, 631
of molecules, 885, 1137
of nuclei in a crystal, 534, 611, 653
of the nuclei in a molecule, 527
Violations of Bell's inequalities, 2210, 2265
Virial theorem, 350, 1210
Virtual transition, 2100
Volume effect, 544, 840, 1162, 1268
Von Neumann
chain, 2201
equation, 306
ideal measurement, 2196
reduction postulate, 2202
statistical entropy, 2217
Vortex in a superfluid, 1667

Water (molecule), 873, 874
Wave (evanescent), 67
Wave function, 88, 140, 226
BCS pairs, 1901, 1909
Hydrogen, 830
norm, 90
pair wave functions, 1851
particle, 11
Wave packet(s)
Gaussian, 57, 2305
in a potential step, 75
in three dimensions, 53
minimal, 290, 520, 591
motion in a harmonic potential, 596
one-photon, 2168
particle, 13
photon, 2163
propagation, 20, 57, 242, 398
reduction, 221, 227, 265, 279
spreading, 57, 59, 347, 348(ex.)
two-dimension, 49
two-photons, 2181
Wave(s)
de Broglie wavelength, 10, 35
evanescent, 29
free spherical waves, 961
multipolar, 2052
partial waves, 948
plane, 14, 19, 943
wave function, 11, 88, 140, 226
Wave-particle duality, 3, 45
Wavelength
Compton wavelength, 1235
de Broglie, 10
Weak coupling (dressed-atom), 2137
Well
potential square well, 29
potential well, 367
Weyl
operator, 2300
quantization, 2311
Which path type of experiments, 2202
Wick's theorem, 1799, 1804
Wigner transform, 2297
Wigner-Eckart theorem, 1065, 1085, 1254

Young (double slit experiment), 4
Yukawa potential, 977

Zeeman
components, polarizations, 865
effect, 855, 862, 987, 1251, 1253, 1257, 1261, 1281
polarization of the components, 1295
slower, 2025
Zeeman effect
Hydrogen, 1289
in muonium, 1281
in positronium, 1281
Muonium, 1284

Zone (Brillouin zone), 614

WILEY-VCH

This new edition of the unrivalled textbook introduces the fundamental concepts of quantum mechanics such as waves, particles and probability before explaining the postulates of quantum mechanics in detail. In the proven didactic manner, the textbook then covers the classical scope of introductory quantum mechanics, namely simple two-level systems, the one-dimensional harmonic oscillator, the quantized angular momentum and particles in a central potential. The entire book has been revised to take into account new developments in quantum mechanics curricula.

The textbook retains its original style in this new edition: fundamental concepts are explained in *chapters,* each followed by *complements* that provide more detailed discussions, examples and applications.

- The quantum mechanics classic in a new edition: written by 1997 Nobel laureate Claude Cohen-Tannoudji and his colleagues Bernard Diu and Franck Laloë
- As easily comprehensible as possible: all steps of the physical background and its mathematical representation are spelled out explicitly
- Comprehensive: in addition to the fundamentals themselves, the book contains more than 350 worked examples plus exercises

Claude Cohen-Tannoudji *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris where he also studied and received his PhD in 1962. In 1973 he became Professor of atomic and molecular physics at the Collège des France. His main research interests were optical pumping, quantum optics and atom-photon interactions. In 1997, Claude Cohen-Tannoudji, together with Steven Chu and William D. Phillips, was awarded the Nobel Prize in Physics for his research on laser cooling and trapping of neutral atoms.*

Bernard Diu *was Professor at the Denis Diderot University (Paris VII). He was engaged in research at the Laboratory of Theoretical Physics and High Energy where his focus was on strong interactions physics and statistical mechanics.*

Franck Laloë *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris. His first assignment was with the University of Paris VI before he was appointed to the CNRS, the French National Research Center. His research was focused on optical pumping, statistical mechanics of quantum gases, musical acoustics and the foundations of quantum mechanics.*

Cover Design: Wiley
Cover Image: © antishock/Getty Images

ISBN 978-3-527-34553-3

www.wiley-vch.de

Also available as an e-book

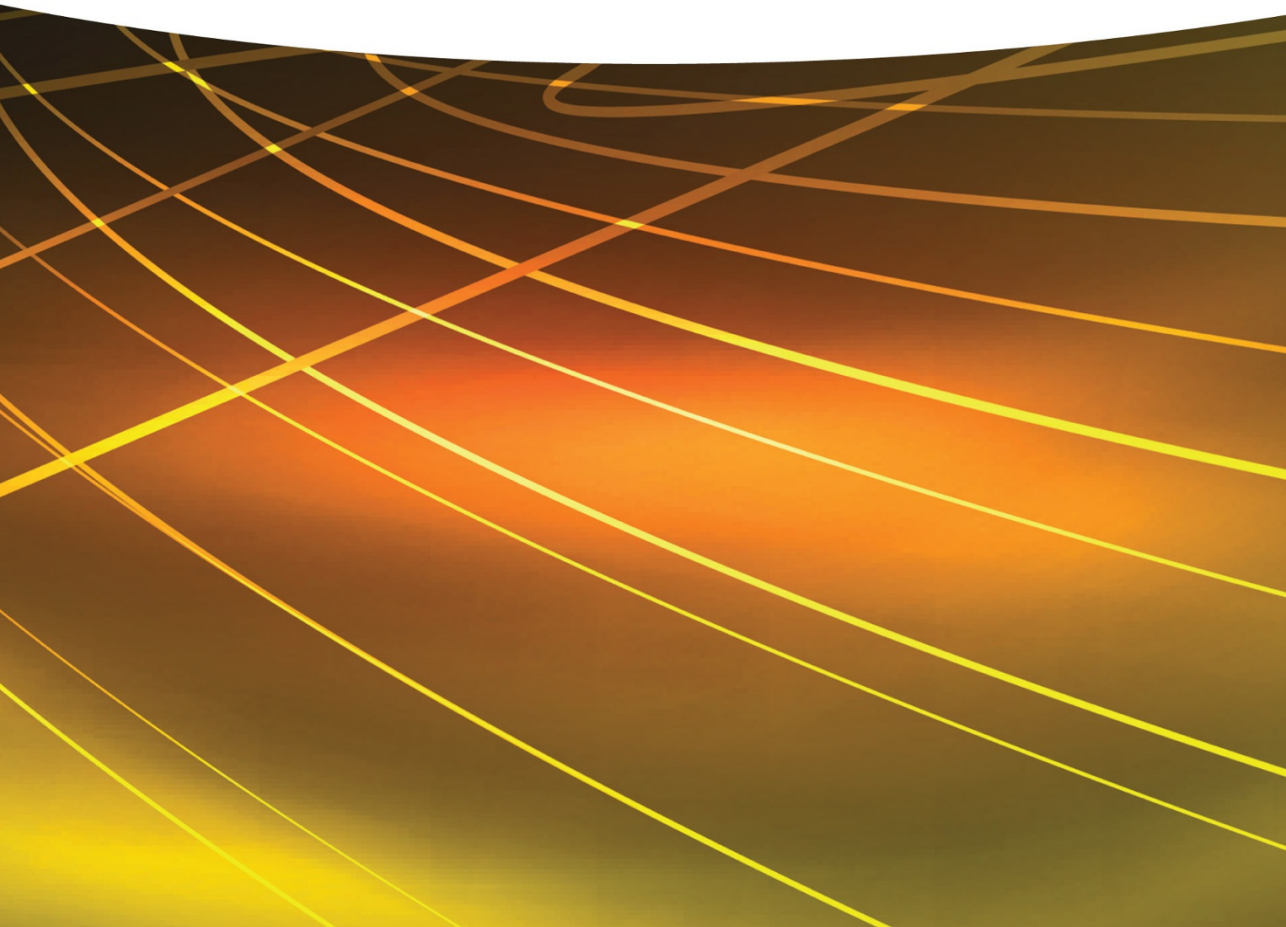
WILEY-VCH

Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë

Quantum Mechanics

Volume II: Angular Momentum, Spin, and Approximation Methods

Second Edition



QUANTUM MECHANICS

Volume II

Angular Momentum, Spin, and Approximation Methods

Claude Cohen-Tannoudji, Bernard Diu,
and Franck Laloë

Translated from the French by **Susan Reid Hemley, Nicole Ostrowsky, and Dan Ostrowsky**

Authors

Prof. Dr. Claude Cohen-Tannoudji
Laboratoire Kastler Brossel (ENS)
24 rue Lhomond
75231 Paris Cedex 05
France

Prof. Dr. Bernard Diu
4 rue du Docteur Roux
91440 Boures-sur-Yvette
France

Prof. Dr. Frank Laloë
Laboratoire Kastler Brossel (ENS)
24 rue Lhomond
75231 Paris Cedex 05
France

Cover Image
© antishock/Getty Images

Second Edition

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data:
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de.

© 2020 WILEY-VCH Verlag GmbH & Co. KGaA, Boschstr. 12, 69469 Weinheim, Germany

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Print ISBN 978-3-527-34554-0
ePDF ISBN 978-3-527-82272-0
ePub ISBN 978-3-527-82273-7

Cover Design Tata Consulting Services
Printing and Binding CPI Ebner & Spiegel

Printed on acid-free paper.

Directions for Use

This book is composed of chapters and their complements:

– *The chapters* contain the fundamental concepts. Except for a few additions and variations, they correspond to a course given in the last year of a typical undergraduate physics program (Volume I) or of a graduate program (Volumes II and III). The 21 chapters are *complete in themselves* and can be studied independently of the complements.

– *The complements* follow the corresponding chapter. Each is labelled by a letter followed by a subscript, which gives the number of the chapter (for example, the complements of Chapter V are, in order, A_V, B_V, C_V, etc.). They can be recognized immediately by the symbol • that appears at the top of each of their pages.

The complements vary in character. Some are intended to expand the treatment of the corresponding chapter or to provide a more detailed discussion of certain points. Others describe concrete examples or introduce various physical concepts. One of the complements (usually the last one) is a collection of exercises.

The *difficulty* of the complements varies. Some are very simple examples or extensions of the chapter. Others are more difficult and at the graduate level or close to current research. In any case, the reader should have studied the material in the chapter before using the complements.

The complements are generally independent of one another. *The student should not try to study all the complements of a chapter at once.* In accordance with his/her aims and interests, he/she should choose a small number of them (two or three, for example), plus a few exercises. The other complements can be left for later study. To help with the choise, the complements are listed at the end of each chapter in a "*reader's guide*", which discusses the difficulty and importance of each.

Some passages within the book have been set in small type, and these can be omitted on a first reading.

Foreword

Quantum mechanics is a branch of physics whose importance has continually increased over the last decades. It is essential for understanding the structure and dynamics of microscopic objects such as atoms, molecules and their interactions with electromagnetic radiation. It is also the basis for understanding the functioning of numerous new systems with countless practical applications. This includes lasers (in communications, medicine, milling, etc.), atomic clocks (essential in particular for the GPS), transistors (communications, computers), magnetic resonance imaging, energy production (solar panels, nuclear reactors), etc. Quantum mechanics also permits understanding surprising physical properties such as superfluidity or supraconductivity. There is currently a great interest in entangled quantum states whose non-intuitive properties of nonlocality and nonseparability permit conceiving remarkable applications in the emerging field of quantum information. Our civilization is increasingly impacted by technological applications based on quantum concepts. This why a particular effort should be made in the teaching of quantum mechanics, which is the object of these three volumes.

The first contact with quantum mechanics can be disconcerting. Our work grew out of the authors' experiences while teaching quantum mechanics for many years. It was conceived with the objective of easing a first approach, and then aiding the reader to progress to a more advance level of quantum mechanics. The first two volumes, first published more than forty years ago, have been used throughout the world. They remain however at an intermediate level. They have now been completed with a third volume treating more advanced subjects. Throughout we have used a progressive approach to problems, where no difficulty goes untreated and each aspect of the diverse questions is discussed in detail (often starting with a classical review).

This willingness to go further "without cheating or taking shortcuts" is built into the book structure, using two distinct linked texts: *chapters* and *complements*. As we just outlined in the "Directions for use", the chapters present the general ideas and basic concepts, whereas the complements illustrate both the methods and concepts just exposed.

Volume I presents a general introduction of the subject, followed by a second chapter describing the basic mathematical tools used in quantum mechanics. While this chapter can appear long and dense, the teaching experience of the authors has shown that such a presentation is the most efficient. In the third chapter the postulates are announced and illustrated in many of the complements. We then go on to certain important applications of quantum mechanics, such as the harmonic oscillator, which lead to numerous applications (molecular vibrations, phonons, etc.). Many of these are the object of specific complements.

Volume II pursues this development, while expanding its scope at a slightly higher level. It treats collision theory, spin, addition of angular momenta, and both time-dependent and time-independent perturbation theory. It also presents a first approach to the study of identical particles. In this volume as in the previous one, each theoretical concept is immediately illustrated by diverse applications presented in the complements. Both volumes I and II have benefited from several recent corrections, but there have also been additions. Chapter XIII now contains two sections §§ D and E that treat random perturbations, and a complement concerning relaxation has been added.

Volume III extends the two volumes at a slightly higher level. It is based on the use of the creation and annihilation operator formalism (second quantization), which is commonly used in quantum field theory. We start with a study of systems of identical particles, fermions or bosons. The properties of ideal gases in thermal equilibrium are presented. For fermions, the Hartree-Fock method is developed in detail. It is the base of many studies in chemistry, atomic physics and solid state physics, etc. For bosons, the Gross-Pitaevskii equation and the Bogolubov theory are discussed. An original presentation that treats the pairing effect of both fermions and bosons permits obtaining the BCS (Bardeen-Cooper-Schrieffer) and Bogolubov theories in a unified framework. The second part of volume III treats quantum electrodynamics, its general introduction, the study of interactions between atoms and photons, and various applications (spontaneous emission, multiphoton transitions, optical pumping, etc.). The dressed atom method is presented and illustrated for concrete cases. A final chapter discusses the notion of quantum entanglement and certain fundamental aspects of quantum mechanics, in particular the Bell inequalities and their violations.

Finally note that we have not treated either the philosophical implications of quantum mechanics, or the diverse interpretations of this theory, despite the great interest of these subjects. We have in fact limited ourselves to presenting what is commonly called the "orthodox point of view". It is only in Chapter XXI that we touch on certain questions concerning the foundations of quantum mechanics (nonlocality, etc.). We have made this choice because we feel that one can address such questions more efficiently after mastering the manipulation of the quantum mechanical formalism as well as its numerous applications. These subjects are addressed in the book *Do we really understand quantum mechanics?* (F. Laloë, Cambridge University Press, 2019); see also section 5 of the bibliography of volumes I and II.

Acknowledgments:

Volumes I and II:

The teaching experience out of which this text grew were group efforts, pursued over several years. We wish to thank all the members of the various groups and particularly Jacques Dupont-Roc and Serge Haroche, for their friendly collaboration, for the fruitful discussions we have had in our weekly meetings and for the ideas for problems and exercises that they have suggested. Without their enthusiasm and valuable help, we would never have been able to undertake and carry out the writing of this book.

Nor can we forget what we owe to the physicists who introduced us to research, Alfred Kastler and Jean Brossel for two of us and Maurice Levy for the third. It was in the context of their laboratories that we discovered the beauty and power of quantum mechanics. Neither have we forgotten the importance to us of the modern physics taught at the C.E.A. by Albert Messiah, Claude Bloch and Anatole Abragam, at a time when graduate studies were not yet incorporated into French university programs.

We wish to express our gratitude to Ms. Aucher, Baudrit, Boy, Brodschi, Emo, Heywaerts, Lemirre, Touzeau for preparation of the mansucript.

Volume III:

We are very grateful to Nicole and Daniel Ostrowsky, who, as they translated this Volume from French into English, proposed numerous improvements and clarifications. More recently, Carsten Henkel also made many useful suggestions during his translation of the text into German; we are very grateful for the improvements of the text that resulted from this exchange. There are actually many colleagues and friends who greatly contributed, each in his own way, to finalizing this book. All their complementary remarks and suggestions have been very helpful and we are in particular thankful to:

Pierre-François Cohadon
Jean Dalibard
Sébastien Gleyzes
Markus Holzmann
Thibaut Jacqmin
Philippe Jacquier
Amaury Mouchet
Jean-Michel Raimond
Félix Werner

Some delicate aspects of Latex typography have been resolved thanks to Marco Picco, Pierre Cladé and Jean Hare. Roger Balian, Edouard Brézin and William Mullin have offered useful advice and suggestions. Finally, our sincere thanks go to Geneviève Tastevin, Pierre-François Cohadon and Samuel Deléglise for their help with a number of figures.

Volume I

Table of contents vii

I WAVES AND PARTICLES. INTRODUCTION TO THE BASIC IDEAS OF QUANTUM MECHANICS 1

READER'S GUIDE FOR COMPLEMENTS 33

A_I Order of magnitude of the wavelengths associated with material particles 35

B_I Constraints imposed by the uncertainty relations 39

C_I Heisenberg relation and atomic parameters 41

D_I An experiment illustrating the Heisenberg relations 45

E_I A simple treatment of a two-dimensional wave packet 49

F_I The relationship between one- and three-dimensional problems 53

G_I One-dimensional Gaussian wave packet: spreading of the wave packet 57

H_I Stationary states of a particle in one-dimensional square potentials 63

J_I Behavior of a wave packet at a potential step 75

K_I Exercises 83

****

II THE MATHEMATICAL TOOLS OF QUANTUM MECHANICS 87

READER'S GUIDE FOR COMPLEMENTS 159

A_{II} The Schwarz inequality 161

B_{II} Review of some useful properties of linear operators 163

C_{II} Unitary operators 173

D_{II} A more detailed study of the r and p representations 181

E_{II} Some general properties of two observables, Q and P, whose commutator is equal to $i\hbar$ 187

F_{II} The parity operator 193

G_{II} An application of the properties of the tensor product: the two-dimensional infinite well **201**

H_{II} Exercises **205**

III THE POSTULATES OF QUANTUM MECHANICS **213**

READER'S GUIDE FOR COMPLEMENTS **267**

A_{III} Particle in an infinite one-dimensional potential well **271**

B_{III} Study of the probability current in some special cases **283**

C_{III} Root mean square deviations of two conjugate observables **289**

D_{III} Measurements bearing on only one part of a physical system **293**

E_{III} The density operator **299**

F_{III} The evolution operator **313**

G_{III} The Schrödinger and Heisenberg pictures **317**

H_{III} Gauge invariance **321**

J_{III} Propagator for the Schrödinger equation **335**

K_{III} Unstable states. Lifetime **343**

L_{III} Exercises **347**

M_{III} Bound states in a "potential well" of arbitrary shape **359**

N_{III} Unbound states of a particle in the presence of a potential well or barrier **367**

O_{III} Quantum properties of a particle in a one-dimensional periodic structure **375**

****

IV APPLICATIONS OF THE POSTULATES TO SIMPLE CASES: SPIN 1/2 AND TWO-LEVEL SYSTEMS **393**

READER'S GUIDE FOR COMPLEMENTS **423**

A_{IV} The Pauli matrices **425**

B_{IV} Diagonalization of a 2×2 Hermitian matrix **429**

C_{IV} Fictitious spin 1/2 associated with a two-level system **435**

D_{IV} System of two spin 1/2 particles **441**

E_{IV} Spin $1/2$ density matrix **449**

F_{IV} Spin 1/2 particle in a static and a rotating magnetic fields: magnetic resonance **455**

G_{IV} A simple model of the ammonia molecule **469**

H_{IV} Effects of a coupling between a stable state and an unstable state **485**

J_{IV} Exercises **491**

****

V THE ONE-DIMENSIONAL HARMONIC OSCILLATOR **497**

READER'S GUIDE FOR COMPLEMENTS **525**

A_V Some examples of harmonic oscillators **527**

B_V Study of the stationary states in the x representation. Hermite polynomials **547**

C_V Solving the eigenvalue equation of the harmonic oscillator by the polynomial method **555**

D_V Study of the stationary states in the momentum representation **563**

E_V The isotropic three-dimensional harmonic oscillator **569**

F_V A charged harmonic oscillator in a uniform electric field **575**

G_V Coherent "quasi-classical" states of the harmonic oscillator **583**

H_V Normal vibrational modes of two coupled harmonic oscillators **599**

J_V Vibrational modes of an infinite linear chain of coupled harmonic oscillators; phonons **611**

K_V Vibrational modes of a continuous physical system. Photons **631**

L_V One-dimensional harmonic oscillator in thermodynamic equilibrium at a temperature T **647**

M_V Exercises **661**

****

VI GENERAL PROPERTIES OF ANGULAR MOMENTUM IN QUANTUM MECHANICS **667**

READER'S GUIDE FOR COMPLEMENTS 703

A_{VI} Spherical harmonics 705

B_{VI} Angular momentum and rotations 717

C_{VI} Rotation of diatomic molecules 739

D_{VI} Angular momentum of stationary states of a two-dimensional harmonic oscillator 755

E_{VI} A charged particle in a magnetic field: Landau levels 771

F_{VI} Exercises 795

****

VII PARTICLE IN A CENTRAL POTENTIAL, HYDROGEN ATOM 803

READER'S GUIDE FOR COMPLEMENTS 831

A_{VII} Hydrogen-like systems 833

B_{VII} A soluble example of a central potential: The isotropic three-dimensional harmonic oscillator 841

C_{VII} Probability currents associated with the stationary states of the hydrogen atom 851

D_{VII} The hydrogen atom placed in a uniform magnetic field. Paramagnetism and diamagnetism. The Zeeman effect 855

E_{VII} Some atomic orbitals. Hybrid orbitals 869

F_{VII} Vibrational-rotational levels of diatomic molecules 885

G_{VII} Exercises 899

INDEX 901

****

Volume II

VOLUME II **923**

Table of contents **v**

VIII AN ELEMENTARY APPROACH TO THE QUANTUM THEORY OF SCATTERING BY A POTENTIAL **923**
A Introduction . . . 923
B Stationary scattering states. Calculation of the cross section . . . 928
C Scattering by a central potential. Method of partial waves . . . 941

READER'S GUIDE FOR COMPLEMENTS **957**

A_{VIII} The free particle: stationary states with well-defined angular momentum **959**
1 The radial equation . . . 959
2 Free spherical waves . . . 961
3 Relation between free spherical waves and plane waves . . . 967

B_{VIII} Phenomenological description of collisions with absorption **971**
1 Principle involved . . . 971
2 Calculation of the cross sections . . . 972

C_{VIII} Some simple applications of scattering theory **977**
1 The Born approximation for a Yukawa potential . . . 977
2 Low energy scattering by a hard sphere . . . 980
3 Exercises . . . 981

****

IX ELECTRON SPIN **985**
A Introduction of electron spin . . . 986
B Special properties of an angular momentum 1/2 . . . 990
C Non-relativistic description of a spin 1/2 particle . . . 991

READER'S GUIDE FOR COMPLEMENTS **999**

A_{IX} Rotation operators for a spin 1/2 particle **1001**
1 Rotation operators in state space . . . 1001
2 Rotation of spin states . . . 1002
3 Rotation of two-component spinors . . . 1005

B_{IX} Exercises **1009**

****

X ADDITION OF ANGULAR MOMENTA 1015
A Introduction 1015
B Addition of two spin 1/2's. Elementary method 1019
C Addition of two arbitrary angular momenta. General method 1025

READER'S GUIDE FOR COMPLEMENTS 1041

A_X Examples of addition of angular momenta 1043
1 Addition of $\mathbf{j}_1 = 1$ and $\mathbf{j}_2 = 1$ 1043
2 Addition of an integral orbital angular momentum l and a spin 1/2 1046

B_X Clebsch-Gordan coefficients 1051
1 General properties of Clebsch-Gordan coefficients 1051
2 Phase conventions. Reality of Clebsch-Gordan coefficients 1054
3 Some useful relations 1056

C_X Addition of spherical harmonics 1059
1 The functions $\Phi_J^M(\Omega_1;\Omega_2)$ 1059
2 The functions $\mathbf{F}_l^m(\Omega)$ 1060
3 Expansion of a product of spherical harmonics; the integral of a product of three spherical harmonics 1062

D_X Vector operators: the Wigner-Eckart theorem 1065
1 Definition of vector operators; examples 1066
2 The Wigner-Eckart theorem for vector operators 1067
3 Application: calculation of the Landé g_J factor of an atomic level 1072

E_X Electric multipole moments 1077
1 Definition of multipole moments 1077
2 Matrix elements of electric multipole moments 1084

F_X Two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ coupled by an interaction $a\mathbf{J}_1 \cdot \mathbf{J}_2$ 1091
1 Classical review 1092
2 Quantum mechanical evolution of the average values $\langle \mathbf{J}_1 \rangle$ and $\langle \mathbf{J}_2 \rangle$ 1093
3 The special case of two spin 1/2's 1095
4 Study of a simple model for the collision of two spin 1/2 particles 1102

G_X Exercises 1107

****

XI STATIONARY PERTURBATION THEORY 1115
A Description of the method 1116
B Perturbation of a non-degenerate level 1120
C Perturbation of a degenerate state 1124

READER'S GUIDE FOR COMPLEMENTS 1129

A_{XI} A one-dimensional harmonic oscillator subjected to a perturbing potential in x, x^2, x^3 1131
1 Perturbation by a linear potential . . . 1131
2 Perturbation by a quadratic potential . . . 1133
3 Perturbation by a potential in x^3 . . . 1135

B_{XI} Interaction between the magnetic dipoles of two spin 1/2 particles 1141
1 The interaction Hamiltonian W . . . 1141
2 Effects of the dipole-dipole interaction on the Zeeman sublevels of two fixed particles . . . 1144
3 Effects of the interaction in a bound state . . . 1149

C_{XI} Van der Waals forces 1151
1 The electrostatic interaction Hamiltonian for two hydrogen atoms . . . 1152
2 Van der Waals forces between two hydrogen atoms in the $1s$ ground state . . . 1154
3 Van der Waals forces between a hydrogen atom in the $1s$ state and a hydrogen atom in the $2p$ state . . . 1158
4 Interaction of a hydrogen atom in the ground state with a conducting wall . . . 1160

D_{XI} The volume effect: the influence of the spatial extension of the nucleus on the atomic levels 1162
1 First-order energy correction . . . 1164
2 Application to some hydrogen-like systems . . . 1166

E_{XI} The variational method 1169
1 Principle of the method . . . 1169
2 Application to a simple example . . . 1172
3 Discussion . . . 1174

F_{XI} Energy bands of electrons in solids: a simple model 1177
1 A first approach to the problem: qualitative discussion . . . 1178
2 A more precise study using a simple model . . . 1181

G_{XI} A simple example of the chemical bond: the H_2^+ ion 1189
1 Introduction . . . 1189
2 The variational calculation of the energies . . . 1192
3 Critique of the preceding model. Possible improvements . . . 1201
4 Other molecular orbitals of the H_2^+ ion . . . 1204
5 The origin of the chemical bond; the virial theorem . . . 1210

H_{XI} Exercises 1221

****

XII AN APPLICATION OF PERTURBATION THEORY: THE FINE AND HYPERFINE STRUCTURE OF THE HYDROGEN ATOM 1231
A Introduction . . . 1232
B Additional terms in the Hamiltonian . . . 1233
C The fine structure of the $n = 2$ level . . . 1238
D The hyperfine structure of the $n = 1$ level . . . 1246
E The Zeeman effect of the $1s$ ground state hyperfine structure . . . 1251

READER'S GUIDE FOR COMPLEMENTS 1265

A_{XII} The magnetic hyperfine Hamiltonian 1267
1 Interaction of the electron with the scalar and vector potentials created by the proton . . . 1267
2 The detailed form of the hyperfine Hamiltonian . . . 1268
3 Conclusion: the hyperfine-structure Hamiltonian . . . 1274

B_{XII} Calculation of the average values of the fine-structure Hamiltonian in the $1s$, $2s$ and $2p$ states 1276
1 Calculation of $\langle 1/R \rangle$, $\langle 1/R^2 \rangle$ and $\langle 1/R^3 \rangle$ . . . 1276
2 The average values $\langle W_{mv} \rangle$ . . . 1278
3 The average values $\langle W_D \rangle$ . . . 1279
4 Calculation of the coefficient ξ_{2p} associated with W_{SO} in the $2p$ level . . . 1279

C_{XII} The hyperfine structure and the Zeeman effect for muonium and positronium 1281
1 The hyperfine structure of the $1s$ ground state . . . 1281
2 The Zeeman effect in the $1s$ ground state . . . 1282

D_{XII} The influence of the electronic spin on the Zeeman effect of the hydrogen resonance line 1289
1 Introduction . . . 1289
2 The Zeeman diagrams of the $1s$ and $2s$ levels . . . 1290
3 The Zeeman diagram of the $2p$ level . . . 1290
4 The Zeeman effect of the resonance line . . . 1293

E_{XII} The Stark effect for the hydrogen atom 1298
1 The Stark effect on the $n = 1$ level . . . 1298
2 The Stark effect on the $n = 2$ level . . . 1300

****

XIII APPROXIMATION METHODS FOR TIME-DEPENDENT PROBLEMS 1303
A Statement of the problem . . . 1303
B Approximate solution of the Schrödinger equation . . . 1305
C An important special case: a sinusoidal or constant perturbation . . . 1309
D Random perturbation . . . 1320
E Long-time behavior for a two-level atom . . . 1324

READER'S GUIDE FOR COMPLEMENTS 1337

A_{XIII} Interaction of an atom with an electromagnetic wave 1339
1 The interaction Hamiltonian. Selection rules 1340
2 Non-resonant excitation. Comparison with the elastically bound electron model 1350
3 Resonant excitation. Absorption and induced emission 1353

B_{XIII} Linear and non-linear responses of a two-level system subject to a sinusoidal perturbation 1357
1 Description of the model 1358
2 The approximate solution of the Bloch equations of the system 1360
3 Discussion 1364
4 Exercises: applications of this complement 1372

C_{XIII} Oscillations of a system between two discrete states under the effect of a sinusoidal resonant perturbation 1374
1 The method: secular approximation 1374
2 Solution of the system of equations 1375
3 Discussion 1376

D_{XIII} Decay of a discrete state resonantly coupled to a continuum of final states 1378
1 Statement of the problem 1378
2 Description of the model 1379
3 Short-time approximation. Relation to first-order perturbation theory 1383
4 Another approximate method for solving the Schrödinger equation 1385
5 Discussion 1386

E_{XIII} Time-dependent random perturbation, relaxation 1390
1 Evolution of the density operator 1391
2 Relaxation of an ensemble of spin 1/2's 1398
3 Conclusion 1408

F_{XIII} Exercises 1409

****

XIV SYSTEMS OF IDENTICAL PARTICLES 1419
A Statement of the problem 1420
B Permutation operators 1425
C The symmetrization postulate 1434
D Discussion 1443

READER'S GUIDE FOR COMPLEMENTS 1457

A_{XIV} Many-electron atoms. Electronic configurations 1459
1 The central-field approximation 1459
2 Electron configurations of various elements 1463

B_{XIV} Energy levels of the helium atom. Configurations, terms, multiplets **1467**
1 The central-field approximation. Configurations 1467
2 The effect of the inter-electron electrostatic repulsion: exchange energy, spectral terms 1469
3 Fine-structure levels; multiplets 1478

C_{XIV} Physical properties of an electron gas. Application to solids **1481**
1 Free electrons enclosed in a box 1481
2 Electrons in solids 1491

D_{XIV} Exercises **1496**

****

APPENDICES **1505**

I Fourier series and Fourier transforms **1505**
1 Fourier series 1505
2 Fourier transforms 1508

II The Dirac δ-"function" **1515**
1 Introduction; principal properties 1515
2 The δ-"function" and the Fourier transform 1520
3 Integral and derivatives of the δ-"function" 1521
4 The δ-"function" in three-dimensional space 1524

III Lagrangian and Hamiltonian in classical mechanics **1527**
1 Review of Newton's laws 1527
2 The Lagrangian and Lagrange's equations 1530
3 The classical Hamiltonian and the canonical equations 1531
4 Applications of the Hamiltonian formalism 1533
5 The principle of least action 1539

BIBLIOGRAPHY OF VOLUMES I AND II **1545**

INDEX **1569**

****

Volume III

VOLUME III **1591**

Table of contents **v**

XV CREATION AND ANNIHILATION OPERATORS FOR IDENTICAL PARTICLES **1591**

READER'S GUIDE FOR COMPLEMENTS **1617**

A_{XV} Particles and holes **1621**

B_{XV} Ideal gas in thermal equilibrium; quantum distribution functions **1625**

C_{XV} Condensed boson system, Gross-Pitaevskii equation **1643**

D_{XV} Time-dependent Gross-Pitaevskii equation **1657**

E_{XV} Fermion system, Hartree-Fock approximation **1677**

F_{XV} Fermions, time-dependent Hartree-Fock approximation **1701**

G_{XV} Fermions or Bosons: Mean field thermal equilibrium **1711**

H_{XV} Applications of the mean field method for non-zero temperature **1733**

****

XVI FIELD OPERATOR **1751**

READER'S GUIDE FOR COMPLEMENTS **1767**

A_{XVI} Spatial correlations in an ideal gas of bosons or fermions **1769**

B_{XVI} Spatio-temporal correlation functions, Green's functions **1781**

C_{XVI} Wick's theorem **1799**

****

XVII PAIRED STATES OF IDENTICAL PARTICLES **1811**

READER'S GUIDE FOR COMPLEMENTS **1843**

A_{XVII} Pair field operator for identical particles 1845

B_{XVII} Average energy in a paired state 1869

C_{XVII} Fermion pairing, BCS theory 1889

D_{XVII} Cooper pairs 1927

E_{XVII} Condensed repulsive bosons 1933

****

XVIII REVIEW OF CLASSICAL ELECTRODYNAMICS 1957

READER'S GUIDE FOR COMPLEMENTS 1977

A_{XVIII} Lagrangian formulation of electrodynamics 1979

****

XIX QUANTIZATION OF ELECTROMAGNETIC RADIATION 1997

READER'S GUIDE FOR COMPLEMENTS 2017

A_{XIX} Momentum exchange between atoms and photons 2019

B_{XIX} Angular momentum of radiation 2043

C_{XIX} Angular momentum exchange between atoms and photons 2055

****

XX ABSORPTION, EMISSION AND SCATTERING OF PHOTONS BY ATOMS 2067

READER'S GUIDE FOR COMPLEMENTS 2095

A_{XX} A multiphoton process: two-photon absorption 2097

B_{XX} Photoionization 2109

C_{XX} Two-level atom in a monochromatic field. Dressed-atom method 2129

D_{XX} Light shifts: a tool for manipulating atoms and fields 2151

E_{XX} Detection of one- or two-photon wave packets, interference 2163

****

XXI QUANTUM ENTANGLEMENT, MEASUREMENTS, BELL'S INEQUALITIES 2187

READER'S GUIDE FOR COMPLEMENTS 2215

A_{XXI} Density operator and correlations; separability 2217

B_{XXI} GHZ states, entanglement swapping 2227

C_{XXI} Measurement induced relative phase between two condensates 2237

D_{XXI} Emergence of a relative phase with spin condensates; macroscopic non-locality and the EPR argument 2253

****

APPENDICES 2267

IV Feynman path integral 2267

V Lagrange multipliers 2281

VI Brief review of Quantum Statistical Mechanics 2285

VII Wigner transform 2297

BIBLIOGRAPHY OF VOLUME III 2325

INDEX 2333

Chapter VIII

An elementary approach to the quantum theory of scattering by a potential

A Introduction 923
A-1 Importance of collision phenomena 923
A-2 Scattering by a potential 925
A-3 Definition of the scattering cross section 926
A-4 Organization of this chapter 927
B Stationary scattering states. Calculation of the cross section 928
B-1 Definition of stationary scattering states 929
B-2 Calculation of the cross section using probability currents . . 932
B-3 Integral scattering equation 935
B-4 The Born approximation 938
C Scattering by a central potential. Method of partial waves 941
C-1 Principle of the method of partial waves 941
C-2 Stationary states of a free particle 942
C-3 Partial waves in the potential $V(r)$ 948
C-4 Expression of the cross section in terms of phase shifts 951

A. Introduction

A-1. Importance of collision phenomena

Many experiments in physics, especially in high energy physics, consist of directing a beam of particles (1) (produced for example by an accelerator) onto a target composed

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

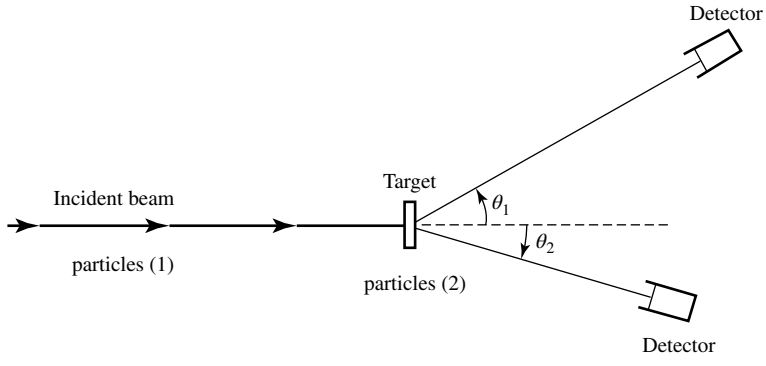


Figure 1: Diagram of a collision experiment involving the particles (1) of an incident beam and the particles (2) of a target. The two detectors represented in the figure measure the number of particles scattered through angles θ_1 and θ_2 with respect to the incident beam.

---

of particles (2), and studying the resulting collisions: the various particles[1] constituting the final state of the system – that is, the state after the collision (*cf.* Fig. 1) – are detected and their characteristics (direction of emission, energy, etc.) are measured. Obviously, the aim of such a study is to determine the interactions that occur between the various particles entering into the collision.

The phenomena observed are sometimes very complex. For example, if particles (1) and (2) are in fact composed of more elementary components (protons and neutrons in the case of nuclei), the latter can, during the collision, redistribute themselves amongst two or several final composite particles which are different from the initial particles; in this case, one speaks of "rearrangement collisions". Moreover, at high energies, the relativistic possibility of the "materialization" of part of the energy appears: new particles are then created and the final state can include a great number of them (the higher the energy of the incident beam, the greater the number). Broadly speaking, one says that collisions give rise to *reactions*, which are described most often as in chemistry:

$$(1) + (2) \rightarrow (3) + (4) + (5) + ... \tag{A-1}$$

Amongst all the reactions possible[2] under given conditions, *scattering* reactions are defined as those in which the final state and the initial state are composed of the same particles (1) and (2). In addition, a scattering reaction is said to be elastic when none of the particles' internal states change during the collision.

[1] In practice, it is not always possible to detect all the particles emitted, and one must often be satisfied with partial information about the final system.

[2] Since the processes studied occur on a quantum level, it is not generally possible to predict with certainty what final state will result from a given collision; one merely attempts to predict the probabilities of the various possible states.

A-2. Scattering by a potential

We shall confine ourselves in this chapter to the study of the elastic scattering of the incident particles (1) by the target particles (2). If the laws of classical mechanics were applicable, solving this problem would involve determining the deviations in the incident particles' trajectories due to the forces exerted by particles (2). For processes occurring on an atomic or nuclear scale, it is clearly out of the question to use classical mechanics to resolve the problem; we must study the evolution of the wave function associated with the incident particles under the influence of their interactions with the target particles [which is why we speak of the "scattering" of particles (1) by particles (2)]. Rather than attack this question in its most general form, we shall introduce the following simplifying hypotheses:

(*i*) We shall suppose that particles (1) and (2) have no spin. This simplifies the theory considerably but should not be taken to imply that the spin of particles is unimportant in scattering phenomena.

(*ii*) We shall not take into account the possible internal structure of particles (1) and (2). The following arguments are therefore not applicable to "inelastic" scattering phenomena, where part of the kinetic energy of (1) is absorbed in the final state by the internal degrees of freedom of (1) and (2) (*cf.* for example, the experiment of Franck and Hertz). We shall confine ourselves to the case of *elastic scattering*, which does not affect the internal structure of the particles.

(*iii*) We shall assume that the target is thin enough to enable us to neglect multiple scattering processes; that is, processes during which a particular incident particle is scattered several times before leaving the target.

(*iv*) We shall neglect any possibility of coherence between the waves scattered by the different particles which make up the target. This simplification is justified when the spread of the wave packets associated with particles (1) is small compared to the average distance between particles (2). Therefore we shall concern ourselves only with the elementary process of the scattering of a particle (1) of the beam by a particle (2) of the target. This excludes a certain number of phenomena which are nevertheless very interesting, such as coherent scattering by a crystal (Bragg diffraction) or scattering of slow neutrons by the phonons of a solid, which provide valuable information about the structure and dynamics of crystal lattices. When these coherence effects can be neglected, the flux of particles detected is simply the sum of the fluxes scattered by each of the $\mathcal{N}$ target particles, that is, $\mathcal{N}$ times the flux scattered by any one of them (the exact position of the scattering particle inside the target is unimportant since the target dimensions are much smaller than the distance between the target and the detector).

(*v*) We shall assume that the interactions between particles (1) and (2) can be described by a potential energy $V(\mathbf{r}_1 - \mathbf{r}_2)$, which depends only on the relative position $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ of the particles. If we follow the reasoning of § B, Chapter VII, then, in the center-of-mass reference frame[3] of the two particles (1) and (2), the problem

[3] In order to interpret the results obtained in scattering experiments, it is clearly necessary to return to the laboratory reference frame. Going from one frame of reference to another is a simple kinematic problem that we will not consider here. See for example Messiah (1.17), vol. I. Chap. X, § 7.

reduces to the study of *the scattering of a single particle by the potential* $V(\mathbf{r})$. The mass μ of this "relative particle" is related to the masses m_1 and m_2 of (1) and (2) by the formula:

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \tag{A-2}$$

A-3. Definition of the scattering cross section

Let Oz be the direction of the incident particles of mass μ (fig. 2). The potential $V(\mathbf{r})$ is localized around the origin O of the coordinate system [which is in fact the center of mass of the two real particles (1) and (2)]. We shall designate by F_i the flux of particles in the incident beam, that is, the number of particles per unit time which traverse a unit surface perpendicular to Oz in the region where z takes on very large negative values. (The flux F_i is assumed to be weak enough to allow us to neglect interactions between different particles of the incident beam.)

We place a detector far from the region under the influence of the potential and in the direction fixed by the polar angles θ and φ, with an opening facing O and subtending the solid angle $\mathrm{d}\Omega$ (the detector is situated at a distance from O which is large compared to the linear dimensions of the potential's zone of influence). We can thus count the number $\mathrm{d}n$ of particles scattered per unit time into the solid angle $\mathrm{d}\Omega$ about the direction (θ, φ). The differential $\mathrm{d}n$ is obviously proportional to $\mathrm{d}\Omega$ and to the incident flux F_i. We shall define $\sigma(\theta, \varphi)$ to be the coefficient of proportionality between $\mathrm{d}n$ and $F_i\ \mathrm{d}\Omega$:

$$\boxed{\mathrm{d}n = F_i\, \sigma(\theta, \varphi)\, \mathrm{d}\Omega} \tag{A-3}$$

The dimensions of $\mathrm{d}n$ and F_i are, respectively, T^{-1} and $(L^2T)^{-1}$, $\sigma(\theta, \varphi)$ therefore has the dimensions of a surface; it is called the *differential scattering cross section* in the direction (θ, φ). Cross sections are frequently measured in barns and submultiples of barns:

$$1 \text{ barn} = 10^{-24} \text{ cm}^2 \tag{A-4}$$

The definition (A-3) can be interpreted in the following way: the number of particles per unit time which reach the detector is equal to the number of particles which would cross a surface $\sigma(\theta, \varphi)\ \mathrm{d}\Omega$ placed perpendicular to Oz in the incident beam.

Similarly, the *total scattering cross section* σ is defined by the formula:

$$\sigma = \int \sigma(\theta, \varphi)\, \mathrm{d}\Omega \tag{A-5}$$

Comments:

(i) Definition (A-3), in which $\mathrm{d}n$ is proportional to $\mathrm{d}\Omega$, implies that only the scattered particles are taken into consideration. The flux of these particles reaching a given detector D [of fixed surface and placed in the direction

(θ, φ)] is inversely proportional to the square of the distance between D and O (this property is characteristic of a scattered flux). In practice, the incident beam is laterally bounded [although its width remains much larger than the extent of the zone of influence of $V(\mathbf{r})$], and the detector is placed outside its trajectory so that it receives only the scattered particles. Of course, such an arrangement does not permit the measurement of the cross section in the direction $\theta = 0$ (the forward direction), which can only be obtained by extrapolation from the values of $\sigma(\theta, \varphi)$ for small θ.

(*ii*) The concept of a cross section is not limited to the case of elastic scattering: reaction cross sections are defined in an analogous manner.

A-4. Organization of this chapter

§ B is devoted to a brief study of scattering by an arbitrary potential $V(\mathbf{r})$ (decreasing however faster than $1/r$ as r tends toward infinity). First of all, in § B-1, we introduce the fundamental concepts of a stationary scattering state and a scattering amplitude. We then show, in § B-2, how knowledge of the asymptotic behavior of the wave functions associated with stationary scattering states enables us to obtain scattering cross sections. Afterwards, in § B-3, we discuss in a more precise way, using the integral scattering equation, the existence of these stationary scattering states. Finally (in § B-4), we derive an approximate solution of this equation, valid for weak potentials. This leads us to the Born approximation, in which the cross section is very simply related to the Fourier transform of the potential.

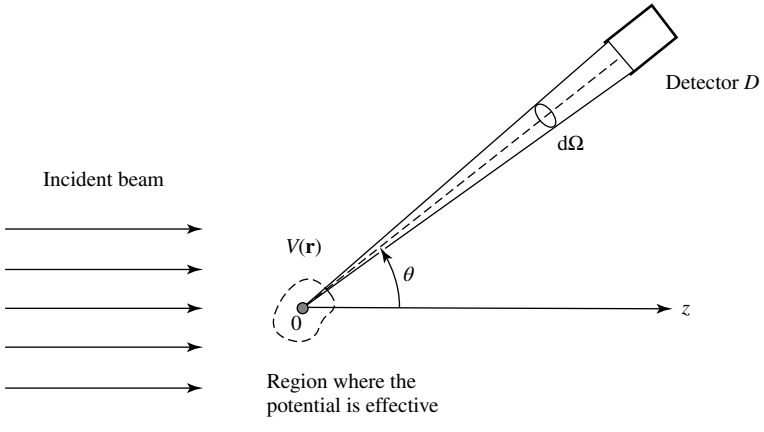


Figure 2: The incident beam, whose flux of particles is F_i, is parallel to the axis Oz; it is assumed to be much wider than the zone of influence of the potential $V(\mathbf{r})$, which is centered at O. Far from this zone of influence, a detector D measures the number dn of particles scattered per unit time into the solid angle $d\Omega$, centered around the direction defined by the polar angles θ and φ. The number dn is proportional to F_i and to $d\Omega$; the coefficient of proportionality $\sigma(\theta, \varphi)$ is, by definition, the scattering "cross section" in the direction (θ, φ).

For a central potential $V(\mathbf{r})$, the general methods described in § B clearly remain applicable, but the method of partial waves, set forth in § C, is usually considered preferable. This method is based (§ C-1) on the comparison of the stationary states with well-defined angular momentum in the presence of the potential $V(r)$ (which we shall call "partial waves") and their analogues in the absence of the potential ("free spherical waves"). Therefore, we begin by studying, in § C-2, the essential properties of the stationary states of a free particle, and more particularly those of free spherical waves. Afterwards (§ C-3), we show that the difference between a partial wave in the potential $V(r)$ and a free spherical wave with the same angular momentum l is characterized by a "phase shift" δ_l. Thus, it is only necessary to know how stationary scattering states can be constructed from partial waves in order to obtain the expression of cross sections in terms of phase shifts (§ C-4).

B. Stationary scattering states. Calculation of the cross section

In order to describe in quantum mechanical terms the scattering of a given incident particle by the potential $V(\mathbf{r})$, it is necessary to study the time evolution of the wave packet representing the state of the particle. The characteristics of this wave packet are assumed to be known for large negative values of the time t, when the particle is in the negative region of the Oz axis, far from and not yet affected by the potential $V(\mathbf{r})$. It is known that the subsequent evolution of the wave packet can be obtained immediately if it is expressed as a superposition of stationary states. This is why we are going to begin by studying the eigenvalue equation of the Hamiltonian:

$$H = H_0 + V(\mathbf{r}) \tag{B-1}$$

where:

$$H_0 = \frac{\mathbf{P}^2}{2\mu} \tag{B-2}$$

describes the particle's kinetic energy.

Actually, to simplify the calculations, we are going to base our reasoning directly on the stationary states and not on wave packets. We have already used this procedure in Chapter I, in the study of "square" one-dimensional potentials (§ D-2 and complement $\mathrm{H_I}$). It consists of considering a stationary state to represent a "probability fluid" in steady flow, and studying the structure of the corresponding probability currents. Naturally, this simplified reasoning is not rigorous: it remains to be shown that it leads to the same results as the correct treatment of the problem, which is based on wave packets. Assuming this will enable us to develop certain general ideas easily, without burying them in complicated calculations[4].

[4]The proof was given in complement $\mathrm{J_I}$, for a particular one-dimensional problem; we verified that the same results are obtained by calculating the probability current associated with a stationary scattering state or by studying the evolution of a wave packet describing a particle which undergoes a collision.

B-1. Definition of stationary scattering states

B-1-a. Eigenvalue equation of the Hamiltonian

Schrödinger's equation describing the evolution of the particle in the potential $V(\mathbf{r})$ is satisfied by solutions associated with a well-defined energy E (stationary states):

$$\psi(\mathbf{r},t) = \varphi(\mathbf{r})\,\mathrm{e}^{-iEt/\hbar} \tag{B-3}$$

where $\varphi(\mathbf{r})$ is a solution of the eigenvalue equation:

$$\left[-\frac{\hbar^2}{2\mu}\Delta + V(\mathbf{r})\right]\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}) \tag{B-4}$$

We are going to assume that the potential $V(\mathbf{r})$ decreases faster than $1/r$ as r approaches infinity. Notice that this hypothesis excludes the Coulomb potential, which demands special treatment; we shall not consider it here.

We shall only be concerned with solutions of (B-4) associated with a positive energy E, equal to the kinetic energy of the incident particle before it reaches the zone of influence of the potential. Defining:

$$E = \frac{\hbar^2 k^2}{2\mu} \tag{B-5}$$

$$V(\mathbf{r}) = \frac{\hbar^2}{2\mu}\,U(\mathbf{r}) \tag{B-6}$$

enables us to write (B-4) in the form:

$$\left[\Delta + k^2 - U(\mathbf{r})\right]\varphi(\mathbf{r}) = 0 \tag{B-7}$$

For each value of k (that is, of the energy E), equation (B-7) can be satisfied by an infinite number of solutions (the positive eigenvalues of the Hamiltonian H are infinitely degenerate). As in "square" one-dimensional potential problems (*cf.* Chap. I, § D-2 and complement $\mathrm{H_I}$), we must choose from amongst these solutions the one that corresponds to the physical problem being studied (for example, when we wanted to determine the probability that a particle with a given energy would cross a one-dimensional potential barrier, we chose the stationary state which, in the region on the other side of the barrier, was composed simply of a transmitted wave). Here, the choice proves to be more complicated, since the particle is moving in three-dimensional space and the potential $V(\mathbf{r})$ has, *a priori*, an arbitrary form. Therefore, we shall specify, using wave packet properties in an intuitive way, the conditions that must be imposed on the solutions of equation (B-7) if they are to be used in the description of a scattering process. We shall call the eigenstates of the Hamiltonian which satisfy these conditions *stationary scattering states*, and we shall designate by $v_k^{(\mathrm{scatt})}(\mathbf{r})$ the associated wave functions.

B-1-b. Asymptotic form of stationary scattering states. Scattering amplitude

For large negative values of t, the incident particle is free [$V(\mathbf{r})$ is practically zero when one is sufficiently far from the point O], and its state is represented by a plane wave packet. Consequently, the stationary wave function that we are looking for must contain a term of the form e^{ikz}, where k is the constant which appears in equation (B-7).

When the wave packet reaches the region which is under the influence of the potential $V(\mathbf{r})$, its structure is profoundly modified and its evolution complicated. Nevertheless, for large positive values of t, it has left this region and once more takes on a simple form: it is now split into a transmitted wave packet which continues to propagate along Oz in the positive direction (hence having the form e^{ikz}) and a scattered wave packet. Consequently, the wave function $v_k^{(\text{scatt})}(\mathbf{r})$, representing the stationary scattering state associated with a given energy $E = \hbar^2 k^2/2\mu$, will be obtained from the superposition of the plane wave e^{ikz} and a scattered wave (we are ignoring the problem of normalization).

The structure of the scattered wave obviously depends on the potential $V(\mathbf{r})$. Yet its asymptotic form (valid far from the zone of influence of the potential) is simple; reasoning by analogy with wave optics, we see that the scattered wave must present the following characteristics for large r:

(i) In a given direction (θ, φ), its radial dependence is of the form e^{ikr}/r. It is a divergent (or "outgoing") wave which has the same energy as the incident wave. The factor $1/r$ results from the fact that there are three spatial dimensions: $(\Delta + k^2)e^{ikr}$ is not zero, while:

$$(\Delta + k^2)\frac{e^{ikr}}{r} = 0 \quad \text{for } r \geq r_0 \text{ where } r_0 \text{ is any positive distance} \tag{B-8}$$

(in optics, the factor $1/r$ insures that the total flux of energy passing through a sphere of radius r is independent of r for large r; in quantum mechanics, it is the probability flux passing through this sphere that does not depend on r).

(ii) Since scattering is not generally isotropic, the amplitude of the outgoing wave depends on the direction (θ, φ) being considered.

Finally, the wave function $v_k^{(\text{scatt})}(\mathbf{r})$ associated with the stationary scattering state is, by definition, the solution of equation (B-7) whose asymptotic behavior is of the form:

$$\boxed{v_k^{(\text{scatt})}(\mathbf{r}) \underset{r\to\infty}{\sim} e^{ikz} + f_k(\theta, \varphi)\ \frac{e^{ikr}}{r}} \tag{B-9}$$

In this expression, only the function $f_k(\theta, \varphi)$, which is called the *scattering amplitude*, depends on the potential $V(\mathbf{r})$. It can be shown (*cf.* § B-3) that equation (B-7) has indeed one and only one solution, for each value of k, that satisfies condition (B-9).

Comments:

(i) We have already pointed out that in order to obtain simply the time evolution of the wave packet representing the state of the incident particle, it is necessary to expand it in terms of eigenstates of the total Hamiltonian H rather than in terms of plane waves. Therefore, let us consider a wave function of the form[5]:

$$\psi(\mathbf{r}, t) = \int_0^\infty dk\, g(k)\, v_k^{(\text{scatt})}(\mathbf{r})\, e^{-iE_k t/\hbar} \tag{B-10}$$

[5] Actually, it is also necessary to superpose the plane waves corresponding to wave vectors $\boldsymbol{k}$ having slightly different orientations, for the incident wave packet is limited in the directions perpendicular to Oz. For the sake of simplicity, we are concerning ourselves here only with the energy dispersion (which limits the spread of the wave packet along Oz)

where:

$$E_k = \frac{\hbar^2 k^2}{2\mu} \tag{B-11}$$

and where the function $g(k)$, taken to be real for the sake of simplicity, has a pronounced peak at $k = k_0$ and practically vanishes elsewhere; $\psi(\mathbf{r}, t)$ is a solution of Schrödinger's equation and therefore correctly describes the time evolution of the particle. It remains to be shown that this function indeed satisfies the boundary conditions imposed by the particular physical problem being considered. According to (B-9), it approaches asymptotically the sum of a plane wave packet and a scattered wave packet:

$$\psi(\mathbf{r}, t) \underset{r\to\infty}{\sim} \int_0^\infty \mathrm{d}k\, g(k)\, \mathrm{e}^{ikz}\, \mathrm{e}^{-iE_k t/\hbar} + \int_0^\infty \mathrm{d}k\, g(k)\, f_k(\theta, \varphi) \frac{\mathrm{e}^{ikr}}{r}\, \mathrm{e}^{-iE_k t/\hbar} \tag{B-12}$$

The position of the maximum of each of these packets can be obtained from the stationary phase condition (*cf.* Chap. I, § C-2). A simple calculation then gives for the plane wave packet:

$$z_M(t) = v_G\, t \tag{B-13}$$

with:

$$v_G = \frac{\hbar k_0}{\mu} \tag{B-14}$$

As for the scattered wave packet, its maximum in the direction (θ, φ) is located at a distance from the point O given by:

$$r_M(\theta, \varphi; t) = -\alpha'_{k_0}(\theta, \varphi) + v_G t \tag{B-15}$$

where $\alpha'_k(\theta, \varphi)$ is the derivative with respect to k of the argument of the scattering amplitude $f_k(\theta, \varphi)$. Note that formulas (B-13) and (B-15) are valid only in the asymptotic region (that is, for large $|t|$).

For large negative values of t, there is no scattered wave packet, as can be seen from (B-15). The waves of which it is composed interfere constructively only for negative values of r, and these values lie outside the domain permitted to r. Therefore, all that we find in this region is the plane wave packet, which, according to (B-13), is making its way towards the interaction region with a group velocity v_G. For large positive values of t, both packets are actually present; the first one moves off along the positive Oz axis, continuing along the path of the incident packet, and the second one diverges in all directions. The scattering process can thus be well described by the asymptotic condition (B-9).

(*ii*) The spatial extension Δz of the wave packet (B-10) is related to the momentum dispersion $\hbar\Delta k$ by the relation:

$$\Delta z \simeq \frac{1}{\Delta k} \tag{B-16}$$

We shall assume that Δk is small enough for Δz to be much larger than the linear dimensions of the potential's zone of influence. Under these conditions, the wave packet moving at a velocity v_G towards the point O (Fig. 3) will take a time:

$$\Delta T \simeq \frac{\Delta z}{v_G} \simeq \frac{1}{v_G \Delta k} \tag{B-17}$$

to cross this zone. Let us fix the time origin at the instant when the center of the incident wave packet reaches point O. Scattered waves exist only for $t \gtrsim -\Delta T/2$, i.e. after the forward edge of the incident wave packet has arrived at the potential's zone of influence. For $t = 0$, the most distant part of the scattered wave packet is at a distance of the order of $\Delta z/2$ from the point O.

Let us now consider an *a priori* different problem, where we have a time-dependent potential, obtained by multiplying $V(\mathbf{r})$ by a function $f(t)$ that increases slowly from 0 to 1 between $t = -\Delta T/2$ and $t = 0$. For t much less than $-\Delta T/2$, the potential is zero and we shall assume that the state of the particle is represented by a plane wave (extending throughout all space). This plane wave begins to be modified only at $t \simeq -\Delta T/2$, and at the instant $t = 0$ the scattered waves look like those in the preceding case.

Thus we see that there is a certain similarity between the two different problems that we have just described. On the one hand, we have scattering by a constant potential of an incident wave packet whose amplitude at the point O increases smoothly between the times $-\Delta T/2$ and zero; on the other hand, we have scattering of a plane wave of constant amplitude by a potential that is gradually "turned on" over the same time interval $[-\Delta T/2, 0]$.

If $\Delta k \to 0$, the wave packet (B-10) tends toward a stationary scattering state [$g(k)$ tends toward $\delta(k - k_0)$]; in addition, according to (B-17), ΔT becomes infinite and the turning on of the potential associated with the function $f(t)$ becomes infinitely slow (for this reason, it is often said to be "adiabatic"). The preceding discussion, although qualitative, thus makes it possible to describe a stationary scattering state as the result of adiabatically imposing a scattering potential on a free plane wave. We could make this interpretation more precise by studying in a more detailed way the evolution of the initial plane wave under the influence of the potential $f(t)V(\mathbf{r})$.

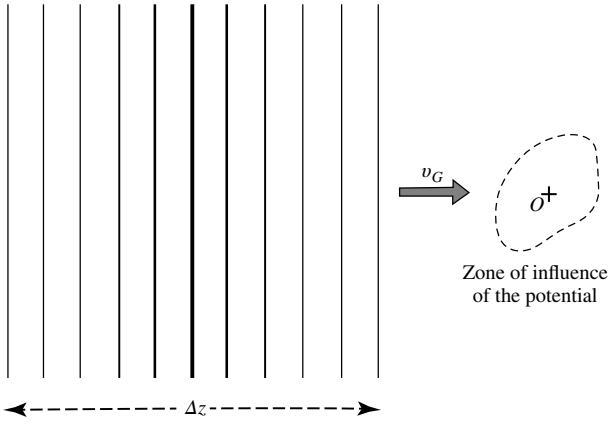


Figure 3: The incident wave packet of length Δz moves at a velocity v_G towards the potential $V(\mathbf{r})$; it interacts with the potential during a time of the order of $\Delta T = \Delta z/v_G$ (assuming the size of the potential's zone of influence to be negligible compared to Δz).

B-2. Calculation of the cross section using probability currents

B-2-a. Probability fluid associated with a stationary scattering state

In order to determine the cross section, one should study the scattering of an incident wave packet by the potential $V(\mathbf{r})$. However, we can obtain the result much more simply by basing our reasoning on the stationary scattering states; we consider

such a state to describe a *probability fluid in steady flow* and we calculate the cross section from the incident and scattered currents. As we have already pointed out, this method is analogous to the one we used in one-dimensional "square" barrier problems: in those problems, the ratio between the reflected (or transmitted) current and the incident current yielded the reflection (or transmission) coefficient directly.

Hence we shall calculate the contributions of the incident wave and the scattered wave to the probability current in a stationary scattering state. We recall that the expression for the current $\mathbf{J}(\mathbf{r})$ associated with a wave function $\varphi(\mathbf{r})$ is:

$$\mathbf{J}(\mathbf{r}) = \frac{1}{\mu}\ \mathrm{Re}\left[\varphi^*(\mathbf{r})\frac{\hbar}{i}\boldsymbol{\nabla}\varphi(\mathbf{r})\right] \tag{B-18}$$

B-2-b. Incident current and scattered current

The incident current $\mathbf{J}_i$ is obtained from (B-18) by replacing $\varphi(\mathbf{r})$ by the plane wave e^{ikz}; $\mathbf{J}_i$ is therefore directed along the Oz axis in the positive direction, and its modulus is:

$$|\mathbf{J}_i| = \frac{\hbar k}{\mu} \tag{B-19}$$

Since the scattered wave is expressed in spherical coordinates in formula (B-9), we shall calculate the components of the scattered current $\mathbf{J}_d$ along the local axes defined by this coordinate system. Recall that the corresponding components of the operator $\boldsymbol{\nabla}$ are:

$$\begin{aligned}(\boldsymbol{\nabla})_r &= \frac{\partial}{\partial r}\\ (\boldsymbol{\nabla})_\theta &= \frac{1}{r}\frac{\partial}{\partial\theta}\\ (\boldsymbol{\nabla})_\varphi &= \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}\end{aligned} \tag{B-20}$$

If we replace $\varphi(\mathbf{r})$ in formula (B-18) by the function $f_k(\theta,\varphi)\mathrm{e}^{ikr}/r$, we can easily obtain the scattered current in the asymptotic region:

$$\begin{aligned}(\mathbf{J}_d)_r &= \frac{\hbar k}{\mu}\frac{1}{r^2}|f_k(\theta,\varphi)|^2\\ (\mathbf{J}_d)_\theta &= \frac{\hbar}{\mu}\frac{1}{r^3}\ \mathrm{Re}\left[\frac{1}{i}f_k^*(\theta,\varphi)\frac{\partial}{\partial\theta}f_k(\theta,\varphi)\right]\\ (\mathbf{J}_d)_\varphi &= \frac{\hbar}{\mu}\frac{1}{r^3\sin\theta}\ \mathrm{Re}\left[\frac{1}{i}f_k^*(\theta,\varphi)\frac{\partial}{\partial\varphi}f_k(\theta,\varphi)\right]\end{aligned} \tag{B-21}$$

Since r is large, $(\mathbf{J}_d)_\theta$ and $(\mathbf{J}_d)_\varphi$ are negligible compared to $(\mathbf{J}_d)_r$, and the scattered current is practically radial.

B-2-c. Expression for the cross section

The incident beam is composed of independent particles, all of which are assumed to be prepared in the same way. Sending a great number of these particles amounts to

repeating the same experiment a great number of times with one particle whose state is always the same. If this state is $v_k^{(\text{scatt})}(\mathbf{r})$, it is clear that the incident flux F_i (that is, the number of particles of the incident beam that cross a unit surface perpendicular to Oz per unit time) is proportional to the flux of the vector $\mathbf{J}_i$ across this surface; that is, according to (B-19) :

$$F_i = C|\mathbf{J}_i| = C\frac{\hbar k}{\mu} \tag{B-22}$$

Similarly, the number $\mathrm{d}n$ of particles that strike the opening of the detector (Fig. 2) per unit time is proportional to the flux of the vector $\mathbf{J}_\mathrm{d}$ across the surface $\mathrm{d}S$ of this opening [the proportionality constant C is the same as in (B-22)]:

$$\begin{aligned}\mathrm{d}n &= C\ \mathbf{J}_d \cdot \mathrm{d}\mathbf{S} = C\ (\mathbf{J}_d)_r\ r^2\,\mathrm{d}\Omega \\ &= C\ \frac{\hbar k}{\mu}\ |f_k(\theta,\varphi)|^2\ \mathrm{d}\Omega\end{aligned} \tag{B-23}$$

We see that $\mathrm{d}n$ is independent of r if r is sufficiently large.

If we substitute formulas (B-22) and (B-23) into the definition (A-3) of the differential cross section $\sigma(\theta,\varphi)$, we obtain:

$$\boxed{\sigma(\theta,\varphi) = |f_k(\theta,\varphi)|^2} \tag{B-24}$$

The differential cross section is thus simply the square of the modulus of the scattering amplitude.

B-2-d. Interference between the incident and the scattered waves

In the preceding sections, we have neglected a contribution to the current associated with $v_k^{(\text{scatt})}(\mathbf{r})$ in the asymptotic region: the one that arises from interference between the plane wave e^{ikz} and the scattered wave, and which is obtained by replacing $\varphi^*(\mathbf{r})$ in (B-18) by e^{-ikz} and $\varphi(\mathbf{r})$ by $f_k(\theta,\varphi)\mathrm{e}^{ikr}/r$, and vice versa.

Nevertheless, we can convince ourselves that these interference terms do not appear when we are concerned with scattering in directions other than the forward direction ($\theta = 0$). In order to see this, let us go back to the description of the collision in terms of wave packets (Fig. 4), and let us take into consideration the fact that in practice the wave packet always has a finite lateral spread. Initially, the incident wave packet is moving towards the zone of influence of $V(\mathbf{r})$ (Fig. 4-a). After the collision (Fig. 4-b), we find two wave packets: a plane one which results from the propagation of the incident wave packet (as if there were no scattering potential) and a scattered one moving away from the point O in all directions. The transmitted wave thus results from the interference between these two wave packets. In general, however, we place the detector D outside the beam, so that it is not struck by transmitted particles; thus we observe only the scattered wave packet and it is not necessary to take into consideration the interference terms that we have just mentioned.

Yet it follows from Figure 4-b that interference between the plane and scattered wave packets cannot be neglected in the forward direction, where they occupy the same region of space. The transmitted wave packet results from this interference. It must have a smaller amplitude than the incident packet because of conservation of total probability (that is, conservation of the number

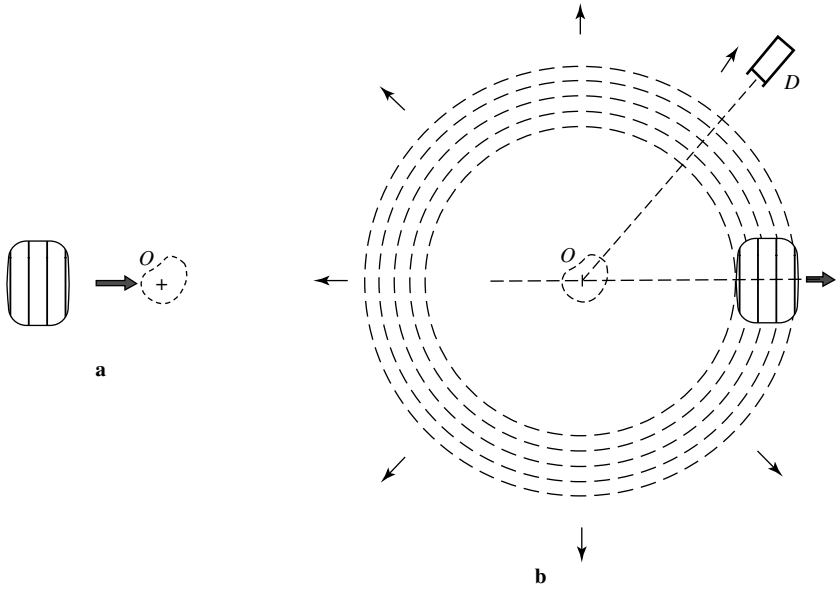


Figure 4: Before the collision (fig. a), the incident wave packet is moving towards the zone of influence of the potential. After the collision (fig. b), we observe a plane wave packet and a spherical wave packet scattered by the potential (dashed lines in the figure). The plane and scattered waves interfere in the forward direction in a destructive way (conservation of total probability); the detector D is placed in a lateral direction and can only see the scattered waves.

---

of particles: particles scattered in all directions of space other than the forward direction leave the beam, whose intensity is therefore attenuated after it has passed the target). It is thus the destructive interference between the plane and forward-scattered wave packets that insures the global conservation of the total number of particles.

B-3. Integral scattering equation

We propose to show now, in a more precise way than in § B-1-b, how one can demonstrate the existence of stationary wave functions whose asymptotic behavior is of the form (B-9). In order to do so, we shall introduce the integral scattering equation, whose solutions are precisely these stationary scattering state wave functions.

Let us go back to the eigenvalue equation of H [formula (B-7)] and put it in the form:

$$(\Delta + k^2)\,\varphi(\mathbf{r}) = U(\mathbf{r})\,\varphi(\mathbf{r}) \tag{B-25}$$

Suppose (we shall see later that this is in fact the case) that there exists a function $G(\mathbf{r})$ such that:

$$(\Delta + k^2)\,G(\mathbf{r}) = \delta(\mathbf{r}) \tag{B-26}$$

[$G(\mathbf{r})$ is called the "Green's function" of the operator $\Delta + k^2$]. Then any function $\varphi(\mathbf{r})$ which satisfies:

$$\varphi(\mathbf{r}) = \varphi_0(\mathbf{r}) + \int \mathrm{d}^3r'\ G(\mathbf{r}-\mathbf{r}')\ U(\mathbf{r}')\ \varphi(\mathbf{r}') \tag{B-27}$$

where $\varphi_0(\mathbf{r})$ is a solution of the homogeneous equation:

$$(\Delta + k^2)\varphi_0(\mathbf{r}) = 0 \tag{B-28}$$

obeys the differential equation (B-25). To show this, we apply the operator $\Delta + k^2$ to both sides of equation (B-27); taking (B-28) into account, we obtain:

$$(\Delta + k^2)\varphi(\mathbf{r}) = (\Delta + k^2)\int \mathrm{d}^3r'\ G(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,\varphi(\mathbf{r}') \tag{B-29}$$

Assuming we can move the operator inside the integral, it will act only on the variable $\mathbf{r}$, and we shall have, according to (B-26):

$$\begin{aligned}(\Delta + k^2)\varphi(\mathbf{r}) &= \int \mathrm{d}^3r'\ \delta(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,\varphi(\mathbf{r}') \\ &= U(\mathbf{r})\,\varphi(\mathbf{r})\end{aligned} \tag{B-30}$$

Inversely, it can be shown that any solution of (B-25) satisfies (B-27)[6]. The differential equation (B-25) can thus be replaced by the integral equation (B-27).

We shall see that it is often easier to base our reasoning on the integral equation. Its principal advantage derives from the fact that by choosing $\varphi_0(\mathbf{r})$ and $G(\mathbf{r})$ correctly, one can incorporate into the equation the desired asymptotic behavior. Thus, one single integral equation, called the *integral scattering equation*, becomes the equivalent of the differential equation (B-25) and the asymptotic condition (B-9).

To begin with, let us consider (B-26). It implies that $(\Delta + k^2)G(\mathbf{r})$ must be identically equal to zero in any region which does not include the origin [which, according to (B-8), is the case when $G(\mathbf{r})$ is equal to e^{ikr}/r]. Moreover, according to relation (61) of Appendix II, $G(\mathbf{r})$ must behave like $-1/4\pi r$ when r approaches zero. In fact, it is easy to show that the functions:

$$G_\pm(\mathbf{r}) = -\frac{1}{4\pi}\,\frac{\mathrm{e}^{\pm ikr}}{r} \tag{B-31}$$

are solutions of equation (B-26). We may write:

$$\begin{aligned}\Delta G_\pm(\mathbf{r}) = \mathrm{e}^{\pm ikr}\Delta\left(-\frac{1}{4\pi r}\right) - \frac{1}{4\pi r}\Delta\left(\mathrm{e}^{\pm ikr}\right) \\ + 2\left[\boldsymbol{\nabla}\left(-\frac{1}{4\pi r}\right)\right]\cdot\left[\boldsymbol{\nabla}\mathrm{e}^{\pm ikr}\right]\end{aligned} \tag{B-32}$$

A simple calculation then gives (*cf.* Appendix II):

$$\Delta G_\pm(\mathbf{r}) = -k^2 G_\pm(\mathbf{r}) + \delta(\mathbf{r}) \tag{B-33}$$

[6]This can be seen intuitively if one considers $U(\mathbf{r})\varphi(\mathbf{r})$ to be the right-hand side of a differential equation: the general solution of (B-25) is then obtained by adding to the general solution of the homogeneous equation a particular solution of the complete equation [second term of (B-27)].

which is what we wished to prove. G_+ and G_- are called, respectively, "outgoing and incoming Green's functions".

The actual form of the desired asymptotic behavior (B-9) suggests the choice of the incident plane wave e^{ikz} for $\varphi_0(\mathbf{r})$ and the choice of the outgoing Green's function $G_+(\mathbf{r})$ for $G(\mathbf{r})$. In fact, we are going to show that the integral scattering equation can be written:

$$v_k^{(\text{scatt})}(\mathbf{r}) = e^{ikz} + \int d^3r' \; G_+(\mathbf{r}-\mathbf{r}')\, U(\mathbf{r}')\, v_k^{(\text{scatt})}(\mathbf{r}') \tag{B-34}$$

whose solutions present the asymptotic behaviour given by (B-9).

To do this, let us place ourselves at a point M (position $\mathbf{r}$), very far from the various points P (position $\mathbf{r}'$) of the zone of influence of the potential[7], whose linear dimensions are of the order of L (Fig. 5):

$$\begin{aligned} &r \gg L \\ &r' \lesssim L \end{aligned} \tag{B-35}$$

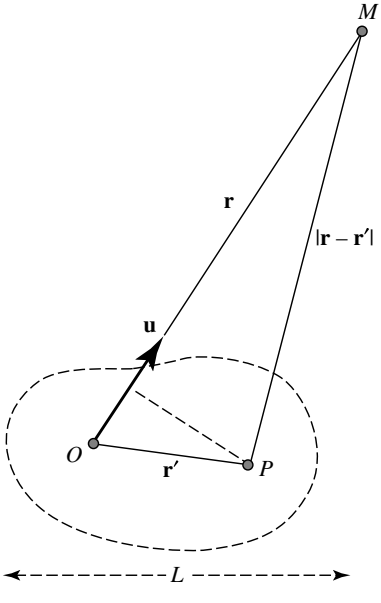


Figure 5: Approximate calculation of the distance $|\mathbf{r}-\mathbf{r}'|$ between a point M very far from O and a point P situated in the zone of influence of the potential (the dimensions of this zone of influence are of the order of L).

Since the angle between MO and MP is very small, the length MP (that is, $|\mathbf{r}-\mathbf{r}'|$) is equal, to within a good approximation, to the projection of MP on MO:

$$|\mathbf{r}-\mathbf{r}'| \simeq r - \mathbf{u}\cdot\mathbf{r}' \tag{B-36}$$

where $\mathbf{u}$ is the unit vector in the $\mathbf{r}$ direction. It follows that, for large r:

$$G_+(\mathbf{r}-\mathbf{r}') = -\frac{1}{4\pi}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \underset{r\to\infty}{\sim} -\frac{1}{4\pi}\frac{e^{ikr}}{r}\, e^{-ik\mathbf{u}\cdot\mathbf{r}'} \tag{B-37}$$

[7]Recall that we have explicitly assumed that $U(\mathbf{r})$ decreases at infinity faster than $1/r$.

Substituting this expression back into equation (B-34), we obtain the asymptotic behavior of $v_k^{(\text{scatt})}(\mathbf{r})$:

$$v_k^{(\text{scatt})}(\mathbf{r}) \underset{r\to\infty}{\sim} \mathrm{e}^{ikz} - \frac{1}{4\pi}\frac{\mathrm{e}^{ikr}}{r}\int \mathrm{d}^3r'\ \mathrm{e}^{-ik\,\mathbf{u}\cdot\mathbf{r}'}U(\mathbf{r}')\,v_k^{(\text{scatt})}(\mathbf{r}') \tag{B-38}$$

which is indeed of the form (B-9), since the integral is no longer a function of the distance $r = OM$ but only (through the unit vector $\mathbf{u}$) of the polar angles θ and φ which fix the direction of the vector $\mathbf{OM}$. Thus, by setting:

$$f_k(\theta,\varphi) = -\frac{1}{4\pi}\int \mathrm{d}^3r'\ \mathrm{e}^{-ik\,\mathbf{u}\cdot\mathbf{r}'}\,U(\mathbf{r}')\,v_k^{(\text{scatt})}(\mathbf{r}') \tag{B-39}$$

we are led to an expression which is identical to (B-9).

It is therefore clear that the solutions of the integral scattering equation (B-34) are indeed the stationary scattering states[8].

Comment:

It is often convenient to define the *incident wave vector* $\mathbf{k}_i$ as a vector of modulus k directed along the Oz axis of the beam such that:

$$\mathrm{e}^{ikz} = \mathrm{e}^{i\mathbf{k}_i\cdot\mathbf{r}} \tag{B-40}$$

In the same way, the vector $\mathbf{k}_s$ which has the same modulus k as the incident wave vector but whose direction is fixed by the angles θ and φ is called the *scattered wave vector* in the direction (θ,φ):

$$\mathbf{k}_s = k\,\mathbf{u} \tag{B-41}$$

Finally, the *scattering (or transferred) wave vector* in the direction (θ,φ) is the difference between $\mathbf{k}_s$ and $\mathbf{k}_i$ (Fig. 6):

$$\mathbf{K} = \mathbf{k}_s - \mathbf{k}_i \tag{B-42}$$

B-4. The Born approximation

B-4-a. Approximate solution of the integral scattering equation

If we take (B-40) into account, we can write the integral scattering equation in the form:

$$v_k^{(\text{scatt})}(\mathbf{r}) = \mathrm{e}^{i\mathbf{k}_i\cdot\mathbf{r}} + \int \mathrm{d}^3r'\ G_+(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,v_k^{(\text{scatt})}(\mathbf{r}') \tag{B-43}$$

We are going to try to solve this equation by iteration.

[8] In order to prove the existence of stationary scattering states rigorously, it would thus be sufficient to demonstrate that equation (B-34) admits a solution.

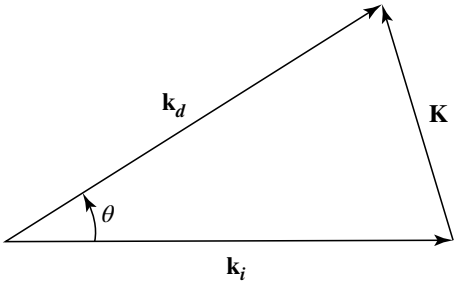


Figure 6: Incident wave vector $\mathbf{k}_i$, scattered wave vector $\mathbf{k}_s$ and transferred wave vector $\mathbf{K}$.

A simple change of notation ($\mathbf{r} \Longrightarrow \mathbf{r}'$; $\mathbf{r}' \Longrightarrow \mathbf{r}''$) permits us to write:

$$v_k^{(\mathrm{scatt})}(\mathbf{r}') = \mathrm{e}^{i\,\mathbf{k}_i\cdot\mathbf{r}'} + \int \mathrm{d}^3 r''\; G_+(\mathbf{r}'-\mathbf{r}'')\,U(\mathbf{r}'')\,v_k^{(\mathrm{scatt})}(\mathbf{r}'') \tag{B-44}$$

Inserting this expression in (B-43), we obtain:

$$\begin{aligned} v_k^{(\mathrm{scatt})}(\mathbf{r}) = \mathrm{e}^{i\,\mathbf{k}_i\cdot\mathbf{r}} &+ \int \mathrm{d}^3 r'\; G_+(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,\mathrm{e}^{i\,\mathbf{k}_i\cdot\mathbf{r}'} \\ &+ \int \mathrm{d}^3 r' \int \mathrm{d}^3 r''\; G_+(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,G_+(\mathbf{r}'-\mathbf{r}'')\;U(\mathbf{r}'')\,v_k^{(\mathrm{scatt})}(\mathbf{r}'') \end{aligned} \tag{B-45}$$

The first two terms on the right-hand side of (B-45) are known; only the third one contains the unknown function $v_k^{(\mathrm{scatt})}(\mathbf{r})$. This procedure can be repeated: changing $\mathbf{r}$ to $\mathbf{r}''$ and $\mathbf{r}'$ to $\mathbf{r}'''$ in (B-43) gives $v_k^{(\mathrm{scatt})}(\mathbf{r}'')$, which can be reinserted in (B-45). We then have:

$$\begin{aligned} v_k^{(\mathrm{scatt})}(\mathbf{r}) = \mathrm{e}^{i\,\mathbf{k}_i\cdot\mathbf{r}} &+ \int \mathrm{d}^3 r'\; G_+(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,\mathrm{e}^{i\,\mathbf{k}_i\cdot\mathbf{r}'} \\ &+ \int \mathrm{d}^3 r' \int \mathrm{d}^3 r''\; G_+(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,G_+(\mathbf{r}'-\mathbf{r}'')\,U(\mathbf{r}'')\,\mathrm{e}^{i\,\mathbf{k}_i\cdot\mathbf{r}''} \\ &+ \int \mathrm{d}^3 r' \int \mathrm{d}^3 r'' \int \mathrm{d}^3 r'''\; G_+(\mathbf{r}-\mathbf{r}')\,U(\mathbf{r}')\,G_+(\mathbf{r}'-\mathbf{r}'')\,U(\mathbf{r}'') \\ &\qquad \times G_+(\mathbf{r}''-\mathbf{r}''')\,U(\mathbf{r}''')\,v_k^{(\mathrm{scatt})}(\mathbf{r}''') \end{aligned} \tag{B-46}$$

where the first three terms are known; the unknown function $v_k^{(\mathrm{scatt})}(\mathbf{r})$ has been pushed back into the fourth term.

Thus we can construct, step by step, what is called the *Born expansion* of the stationary scattering wave function. Note that each term of this expansion brings in one higher power of the potential than the preceding one. Thus, if the potential is weak, each successive term is smaller than the preceding one. If we push the expansion far enough, we can neglect the last term on the right-hand side and thus obtain $v_k^{(\mathrm{scatt})}(\mathbf{r})$ entirely in terms of known quantities.

If we substitute this expansion of $v_k^{(\mathrm{scatt})}(\mathbf{r})$ into expression (B-39), we obtain the Born expansion of the scattering amplitude. In particular, if we limit ourselves to first

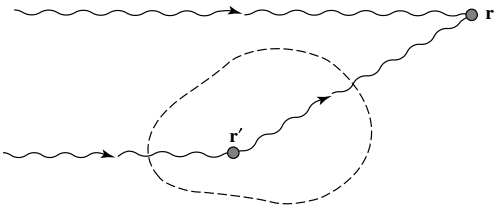


Figure 7: Schematic representation of the Born approximation: we only consider the incident wave and the waves scattered by one interaction with the potential.

order in U, all we need to do is replace $v_k^{(\mathrm{scatt})}(\mathbf{r}')$ by $\mathrm{e}^{i\mathbf{k}_i\cdot\mathbf{r}'}$ on the right-hand side of (B-39). This is the *Born approximation*:

$$\begin{aligned} f_k^{(B)}(\theta,\varphi) &= -\frac{1}{4\pi}\int \mathrm{d}^3r'\,\mathrm{e}^{-ik\,\mathbf{u}\cdot\mathbf{r}'}\,U(\mathbf{r}')\,\mathrm{e}^{i\,\mathbf{k}_i\cdot\mathbf{r}'} \\ &= -\frac{1}{4\pi}\int \mathrm{d}^3r'\,\mathrm{e}^{-i(\mathbf{k}_s-\mathbf{k}_i)\cdot\mathbf{r}'}\,U(\mathbf{r}') \\ &= -\frac{1}{4\pi}\int \mathrm{d}^3r'\,\mathrm{e}^{-i\mathbf{K}\cdot\mathbf{r}'}\,U(\mathbf{r}') \end{aligned} \tag{B-47}$$

where $\mathbf{K}$ is the scattering wave vector defined in (B-42). The scattering cross section, in the Born approximation, is thus very simply related to the Fourier transform of the potential, since, using (B-24) and (B-6), (B-47) implies:

$$\sigma_k^{(B)}(\theta,\varphi) = \frac{\mu^2}{4\pi^2\hbar^4}\left|\int \mathrm{d}^3r\,\mathrm{e}^{-i\mathbf{K}\cdot\mathbf{r}}\,V(\mathbf{r})\right|^2 \tag{B-48}$$

According to Figure 6, the direction and modulus of the scattering wave vector $\mathbf{K}$ depend both on the modulus k of $\mathbf{k}_i$ and $\mathbf{k}_s$ and on the scattering direction (θ,φ). Thus, for a given θ and φ, the Born cross section varies with k, that is, with the energy of the incident beam. Similarly, for a given energy, $\sigma^{(B)}$ varies with θ and φ. We thus see, within the simple framework of the Born approximation, how studying the variation of the differential cross section in terms of the scattering direction and the incident energy gives us information about the potential $V(\mathbf{r})$.

B-4-b. Interpretation of the formulas

We can give formula (B-45) a physical interpretation which brings out very clearly the formal analogy between quantum mechanics and wave optics.

Let us consider the zone of influence of the potential to be a scattering medium whose density is proportional to $U(\mathbf{r})$. The function $G_+(\mathbf{r}-\mathbf{r}')$ [formula (B-31)] represents the amplitude at the point $\mathbf{r}$ of a wave radiated by a point source situated at $\mathbf{r}'$. Consequently, the first two terms of formula (B-45) describe the total wave at the point $\mathbf{r}$ as the result of the superposition of the incident wave $\mathrm{e}^{i\mathbf{k}_i\cdot\mathbf{r}}$ and an infinite number of waves coming from *secondary sources* induced in the scattering medium by the incident wave. The amplitude of each of these sources is indeed proportional to the incident wave ($\mathrm{e}^{i\mathbf{k}_i\cdot\mathbf{r}'}$) and the density of the scattering material [$U(\mathbf{r}')$], evaluated at the corresponding point $\mathbf{r}'$. This interpretation, symbolized by Figure 7, recalls *Huygens' principle* in wave optics.

Actually, formula (B-45) includes a third term. However, we can interpret in an analogous fashion the successive terms of the Born expansion. Since the scattering medium extends over a certain area, a given secondary source is excited not only by the incident wave but also by scattered waves coming from other secondary sources. Figure 8 represents symbolically the third term of the Born expansion [*cf.* formula (B-46)]. If the scattering medium has a very low density [$U(\mathbf{r})$ very small], we can neglect the influence of secondary sources on each other.

Comment:

The interpretation that we have just given for higher-order terms in the Born expansion has nothing to do with the multiple scattering processes that can occur inside a thick target: we are only concerned, here, with the scattering of one particle of the beam by a single particle of the target, while multiple scattering brings in the successive interactions of the same incident particle with several different particles of the target.

C. Scattering by a central potential. Method of partial waves

C-1. Principle of the method of partial waves

In the special case of a central potential $V(r)$, the orbital angular momentum $\mathbf{L}$ of the particle is a constant of the motion. Therefore, there exist stationary states with well-defined angular momentum: that is, eigenstates common to H, $\mathbf{L}^2$ and L_z. We shall call the wave functions associated with these states *partial waves* and we shall write them $\varphi_{k,l,m}(\mathbf{r})$. The corresponding eigenvalues of H, $\mathbf{L}^2$ and L_z are, respectively, $\hbar^2k^2/2\mu$, $l(l+1)\hbar^2$ and $m\hbar$. Their angular dependence is always given by the spherical harmonics $Y_l^m(\theta,\varphi)$; the potential $V(r)$ influences only their radial dependence.

We expect that, for large r, the partial waves will be very close to the common eigenfunctions of H_0, $\mathbf{L}^2$ and L_z, where H_0 is the free Hamiltonian [formula (B-2)]. This is why we are first going to study, in § C-2, the stationary states of a free particle, and, in particular, those which have a well-defined angular momentum. The corresponding wave functions $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ are *free spherical waves*: their angular dependence is, of course, that of a spherical harmonic and we shall see that the asymptotic expression for their radial function is the superposition of an incoming wave e^{-ikr}/r and an outgoing wave e^{ikr}/r with a well-determined phase difference.

The asymptotic expression for the partial wave $\varphi_{k,l,m}(\mathbf{r})$ in the potential $V(r)$ is also (§ C-3) the superposition of an incoming wave and an outgoing wave. However, the

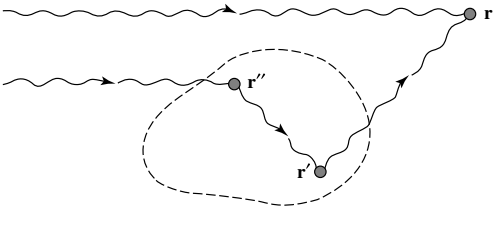


Figure 8: Schematic representation of the second-order term in U in the Born expansion: here we consider waves which are scattered twice by the potential.

phase difference between these two waves is different from the one that characterizes the corresponding free spherical wave: the potential $V(r)$ introduces a supplementary *phase shift* δ_l. This phase shift constitutes the only difference between the asymptotic behavior of $\varphi_{k,l,m}$ and that of $\varphi^{(0)}_{k,l,m}$. Consequently, for fixed k, the phase shifts δ_l for all values of l are all we need to know to be able to calculate the cross section.

In order to carry out this calculation, we shall express (§ C-4) the stationary scattering state $v_k^{(\text{scatt})}(\mathbf{r})$ as a linear combination of partial waves $\varphi_{k,l,m}(\mathbf{r})$ having the same energy but different angular momenta. Simple physical arguments suggest that the coefficients of this linear combination should be the same as those of the free spherical wave expansion of the plane wave e^{ikz}; this is in fact confirmed by an explicit calculation.

The use of partial waves thus permits us to express the scattering amplitude, and hence the cross section, in terms of the phase shifts δ_l. This method is particularly attractive when the range of the potential is not much longer than the wavelength associated with the particle's motion, for, in this case, only a small number of phase shifts are involved (§ C-3-b-β).

C-2. Stationary states of a free particle

In classical mechanics, a free particle of mass μ moves along a uniform linear trajectory. Its momentum $\mathbf{p}$, its energy $E = \mathbf{p}^2/2\mu$ and its angular momentum $\boldsymbol{\mathscr{L}} = \mathbf{r} \times \mathbf{p}$ relative to the origin of the coordinate system are constants of the motion.

In quantum mechanics, the observables $\mathbf{P}$ and $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ do not commute. Hence they represent incompatible quantities: it is impossible to measure the momentum and the angular momentum of a particle simultaneously.

The quantum mechanical Hamiltonian H_0 is written:

$$H_0 = \frac{1}{2\mu}\,\mathbf{P}^2 \tag{C-1}$$

H_0 does not constitute by itself a C.S.C.O.: its eigenvalues are infinitely degenerate (§ 2-a). On the other hand, the four observables:

$$H_0,\ P_x,\ P_y,\ P_z \tag{C-2}$$

form a C.S.C.O. Their common eigenstates are stationary states of well defined momentum. A free particle may also be considered as being placed in a zero central potential. The results of Chap. VII then indicate that the three observables:

$$H_0,\ \boldsymbol{L}^2,\ L_z \tag{C-3}$$

form a C.S.C.O. The corresponding eigenstates are stationay states with well-defined angular momentum (more precisely, $\mathbf{L}^2$ and L_z have well-defined values, but L_x and L_y do not).

The bases of the state space defined by the C.S.C.O.'s (C-2) and (C-3) are distinct, since $\mathbf{P}$ and $\mathbf{L}$ are incompatible quantities. We are going to study these two bases and show how one can pass from one to the other.

C-2-a. Stationary states with well-defined momentum. Plane waves

We already know (*cf.* Chap. II, § E-2-d) that the three observables P_x, P_y and P_z form a C.S.C.O. (for a spinless particle). Their common eigenstates give a basis for the $\{|\mathbf{p}\rangle\}$ representation:

$$\mathbf{P}\,|\mathbf{p}\rangle = \mathbf{p}\,|\mathbf{p}\rangle \tag{C-4}$$

Since H_0 commutes with these three observables, the states $|\mathbf{p}\rangle$ are necessarily eigenstates of H_0:

$$H_0\,|\mathbf{p}\rangle = \frac{\mathbf{p}^2}{2\mu}\,|\mathbf{p}\rangle \tag{C-5}$$

The spectrum of H_0 is therefore continuous and includes all positive numbers and zero. Each of these eigenvalues is infinitely degenerate: to a fixed positive energy E there corresponds an infinite number of kets $|\mathbf{p}\rangle$ since there exists an infinite number of ordinary vectors $\mathbf{p}$ whose modulus satisfies:

$$|\mathbf{p}| = \sqrt{2\mu E} \tag{C-6}$$

The wave functions associated with the kets $|\mathbf{p}\rangle$ are the plane waves (*cf.* Chap. II, § E-1-a):

$$\langle\mathbf{r}|\mathbf{p}\rangle = \left(\frac{1}{2\pi\hbar}\right)^{3/2} \mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \tag{C-7}$$

We shall introduce here the wave vector $\mathbf{k}$ to characterize a plane wave:

$$\mathbf{k} = \frac{\mathbf{p}}{\hbar} \tag{C-8}$$

and we shall define:

$$|\mathbf{k}\rangle = (\hbar)^{3/2}\,|\mathbf{p}\rangle \tag{C-9}$$

The kets $|\mathbf{k}\rangle$ are stationary states with well-defined momentum:

$$H_0\,|\mathbf{k}\rangle = \frac{\hbar^2\mathbf{k}^2}{2\mu}\,|\mathbf{k}\rangle \tag{C-10a}$$

$$\mathbf{P}\,|\mathbf{k}\rangle = \hbar\mathbf{k}\,|\mathbf{k}\rangle \tag{C-10b}$$

They are orthonormal in the extended sense:

$$\langle\mathbf{k}|\mathbf{k}'\rangle = \delta(\mathbf{k}-\mathbf{k}') \tag{C-11}$$

and form a basis in the state space:

$$\int \mathrm{d}^3k\,|\mathbf{k}\rangle\,\langle\mathbf{k}| = \mathbb{1} \tag{C-12}$$

The associated wave functions are the plane waves; normalized in a slightly different way:

$$\langle\mathbf{r}|\mathbf{k}\rangle = \left(\frac{1}{2\pi}\right)^{3/2} \mathrm{e}^{i\,\mathbf{k}\cdot\mathbf{r}} \tag{C-13}$$

C-2-b. Stationary states with well-defined angular momentum. Free spherical waves

In order to obtain the eigenfunctions common to H_0, $\mathbf{L}^2$ and L_z, all we have to do is solve the radial equation for an identically zero central potential. The detailed solution of this problem is given in complement A_{VIII}; we shall be satisfied here with giving the results.

Free spherical waves are the wave functions associated with the well-defined angular momentum stationary states $|\varphi^{(0)}_{k,l,m}\rangle$ of a free particle; they are written:

$$\varphi^{(0)}_{k,l,m}(\mathbf{r}) = \sqrt{\frac{2k^2}{\pi}}\, j_l(kr)\, Y_l^m(\theta,\varphi) \tag{C-14}$$

where j_l is a spherical Bessel function defined by:

$$j_l(\rho) = (-1)^l\, \rho^l \left(\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^l \frac{\sin\rho}{\rho} \tag{C-15}$$

The corresponding eigenvalues of H_0, $\mathbf{L}^2$ and L_z are, respectively, $\hbar^2k^2/2\mu$, $l(l+1)\hbar^2$ and $m\hbar$.

The free spherical waves (C-14) are orthonormal in the extended sense:

$$\begin{aligned}\left\langle \varphi^{(0)}_{k,l,m} | \varphi^{(0)}_{k',l',m'} \right\rangle &= \frac{2}{\pi} kk' \int_0^\infty j_l(kr)\, j_{l'}(k'r)\, r^2\, \mathrm{d}r \times \int \mathrm{d}\Omega\, Y_l^{m*}(\theta,\varphi)\, Y_{l'}^{m'}(\theta,\varphi) \\ &= \delta(k-k')\, \delta_{ll'}\, \delta_{mm'}\end{aligned} \tag{C-16}$$

and form a basis in the state space:

$$\int_0^\infty \mathrm{d}k \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \left|\varphi^{(0)}_{k,l,m}\right\rangle \left\langle \varphi^{(0)}_{k,l,m}\right| = \mathbb{1} \tag{C-17}$$

C-2-c. Physical properties of free spherical waves

α. Angular dependence

The angular dependence of the free spherical wave $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ is entirely given by the spherical harmonic $Y_l^m(\theta,\varphi)$. It is thus fixed by the eigenvalues of $\mathbf{L}^2$ and L_z (that is, by the indices l and m) and not by the energy. For example, a free $s(l=0)$ spherical wave is always isotropic.

β. Behavior in the neighborhood of the origin

Let us consider an infinitesimal solid angle $\mathrm{d}\Omega_0$ about the direction (θ_0,φ_0); when the state of the particle is $|\varphi^{(0)}_{k,l,m}\rangle$, the probability of finding the particle in this solid angle between r and $r+\mathrm{d}r$ is proportional to:

$$r^2\, j_l^2(kr)\, |Y_l^m(\theta_0,\varphi_0)|^2\, \mathrm{d}r\, \mathrm{d}\Omega_0 \tag{C-18}$$

It can be shown (complement A_{VIII}, § 2-c-α) that when ρ approaches zero:

$$j_l(\rho) \underset{\rho\to 0}{\sim} \frac{\rho^l}{(2l+1)!!} \tag{C-19}$$

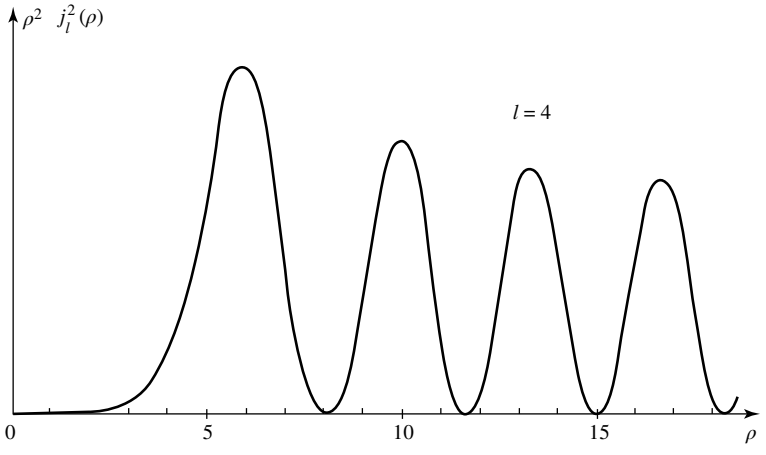


Figure 9: Graph of the function $\rho^2 j_l^2(\rho)$ giving the radial dependence of the probability of finding the particle in the state $|\varphi_{k,l,m}^{(0)}\rangle$. At the origin, this function behaves like ρ^{2l+2}; it remains practically zero as long as $\rho < \sqrt{l(l+1)}$.

This result (which the discussion of Chapter VII, § A-2-c would lead us to expect) implies that the probability (C-18) behaves like r^{2l+2} near the origin; hence, the larger l is, the more slowly it increases.

The shape of the function $\rho^2 j_l^2(\rho)$ is shown in Figure 9. We see that this function remains small as long as:

$$\rho < \sqrt{l(l+1)} \tag{C-20}$$

We may thus assume that the probability (C-18) is practically zero for:

$$r < \frac{1}{k}\sqrt{l(l+1)} \tag{C-21}$$

From a physical point of view, this result is very important for it implies that a particle in the state $|\varphi_{k,l,m}^{(0)}\rangle$ is practically unaffected by what happens inside a sphere centered at O of radius:

$$b_l(k) = \frac{1}{k}\sqrt{l(l+1)} \tag{C-22}$$

We shall return to this point in § C-3-b-β.

Comment:

In classical mechanics, a free particle of momentum $\mathbf{p}$ and angular momentum $\boldsymbol{\mathscr{L}}$ moves in a straight line whose distance b from the point O is given (fig. 10) by:

$$b = \frac{|\boldsymbol{\mathscr{L}}|}{|\mathbf{p}|} \tag{C-23}$$

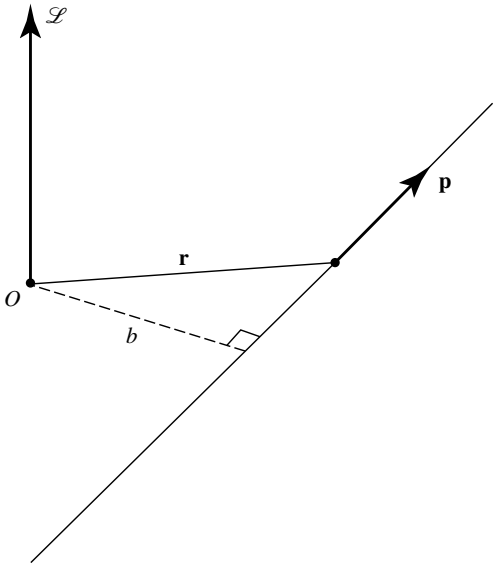


Figure 10: Definition of the classical impact parameter b of a particle of momentum $\mathbf{p}$ *and angular momentum* $\boldsymbol{\mathscr{L}}$ *relative to O.*

b is called the "collision parameter" or "impact parameter" of the particle relative to O; the larger $|\boldsymbol{\mathscr{L}}|$ is and the smaller the momentum (i.e. the energy), the larger b is. If $|\boldsymbol{\mathscr{L}}|$ is replaced by $\hbar\sqrt{l(l+1)}$ and $|\mathbf{p}|$ by $\hbar k$ in (C-23), we again find expression (C-22) for $b_l(k)$, which can thus be interpreted semi-classically.

γ. *Asymptotic behavior*

It can be shown (complement A_{VIII}, § 2-c-β) that as ρ approaches infinity:

$$j_l(\rho) \underset{\rho\to\infty}{\sim} \frac{1}{\rho}\,\sin\left(\rho - l\frac{\pi}{2}\right) \tag{C-24}$$

Consequently, the asymptotic behavior of the free spherical wave $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ is:

$$\varphi^{(0)}_{k,l,m}(r,\theta,\varphi) \underset{\mathbf{r}\to\infty}{\sim} -\sqrt{\frac{2k^2}{\pi}}\,Y_l^m(\theta,\varphi)\frac{\mathrm{e}^{-ikr}\,\mathrm{e}^{il\frac{\pi}{2}} - \mathrm{e}^{ikr}\,\mathrm{e}^{-il\frac{\pi}{2}}}{2ikr} \tag{C-25}$$

At infinity, $\varphi^{(0)}_{k,l,m}$ therefore results from the *superposition of an incoming wave* e^{-ikr}/r *and an outgoing wave* e^{ikr}/r, whose amplitudes differ by a phase difference equal to $l\pi$.

Comment:

Suppose that we construct a packet of free spherical waves, all corresponding to the same values of l and m. A line of reasoning analogous to that of comment (i) of § B-1-b may be applied to it. The following conclusion results: for large negative values of t, only an incoming wave packet exists; while for large positive

values of t, only an outgoing wave packet exists. Therefore, a free spherical wave may be thought of in the following manner: at first we have an incoming wave converging towards O; it becomes distorted as it approaches this point, retraces its steps when it is at a distance of the order of $b_l(k)$ [formula (C-22)], and gives rise to an outgoing wave with a phase shift of $l\pi$.

C-2-d. Expansion of a plane wave in terms of free spherical waves

We thus have two distinct bases formed by eigenstates of H_0: the $\{|\mathbf{k}\rangle\}$ basis associated with the plane waves and the $\{|\varphi^{(0)}_{k,l,m}\rangle\}$ basis associated with the free spherical waves. It is possible to expand any ket of one basis in terms of vectors of the other one.

Let us consider in particular the ket $|0,0,k\rangle$, associated with a plane wave of wave vector $\mathbf{k}$ directed along Oz:

$$\langle \mathbf{r}|0,0,k\rangle = \left(\frac{1}{2\pi}\right)^{3/2} \mathrm{e}^{ikz} \tag{C-26}$$

$|0,0,k\rangle$ represents a state of well-defined energy and momentum ($E = \hbar^2k^2/2\mu$; $\mathbf{p}$ directed along Oz with modulus $\hbar k$). Now:

$$\mathrm{e}^{ikz} = \mathrm{e}^{ikr\cos\theta} \tag{C-27}$$

is independent of φ; since L_z corresponds to $\dfrac{\hbar}{i}\dfrac{\partial}{\partial\varphi}$ in the $\{|\mathbf{r}\rangle\}$ representation, the ket $|0,0,k\rangle$ is also an eigenvector of L_z, with the eigenvalue zero:

$$L_z\,|0,0,k\rangle = 0 \tag{C-28}$$

Using the closure relation (C-17), we can write:

$$|0,0,k\rangle = \int_0^\infty \mathrm{d}k' \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \left|\varphi^{(0)}_{k',l,m}\right\rangle \left\langle\varphi^{(0)}_{k',l,m}\right|0,0,k\Big\rangle \tag{C-29}$$

Since $|0,0,k\rangle$ and $\left|\varphi^{(0)}_{k,l,m}\right\rangle$ are two eigenstates of H_0, they are orthogonal if the corresponding eigenvalues are different; their scalar product is therefore proportional to $\delta(k'-k)$. Similarly, they are both eigenstates of L_z and their scalar product is proportional to δ_{m0} [*cf.* relation (C-28)]. Formula (C-29) therefore takes on the form:

$$|0,0,k\rangle = \sum_{l=0}^{\infty} c_{k,l} \left|\varphi^{(0)}_{k,l,0}\right\rangle \tag{C-30}$$

The coefficients $c_{k,l}$ can be calculated explicitly (complement $\mathrm{A_{VIII}}$, § 3). Thus we obtain:

$$\boxed{\mathrm{e}^{ikz} = \sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\; j_l(kr)\, Y_l^0(\theta)} \tag{C-31}$$

A state of well-defined linear momentum is therefore formed by the superposition of states corresponding to all possible angular momenta.

Comment:

The spherical harmonic $Y_l^0(\theta)$ is proportional to the Legendre polynomial $P_l(\cos\theta)$ (complement A_{VI}, § 2-e-α):

$$Y_l^0(\theta) = \sqrt{\frac{(2l+1)}{4\pi}} P_l(\cos\theta) \tag{C-32}$$

Hence the expansion (C-31) is often written in the form:

$$e^{ikz} = \sum_{l=0}^{\infty} i^l (2l+1)\, j_l(kr)\, P_l(\cos\theta) \tag{C-33}$$

C-3. Partial waves in the potential $V(r)$

We are now going to study the eigenfunctions common to H (the total Hamiltonian), $\mathbf{L}^2$ and L_z; that is, the partial waves $\varphi_{k,l,m}(\mathbf{r})$.

C-3-a. Radial equation. Phase shifts

For any central potential $V(r)$, the partial waves $\varphi_{k,l,m}(\mathbf{r})$ are of the form:

$$\varphi_{k,l,m}(\mathbf{r}) = R_{k,l}(r)\, Y_l^m(\theta,\varphi) = \frac{1}{r}\, u_{k,l}(r)\, Y_l^m(\theta,\varphi) \tag{C-34}$$

where $u_{k,l}(r)$ is the solution of the radial equation:

$$\left[-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r)\right] u_{k,l}(r) = \frac{\hbar^2 k^2}{2\mu}\, u_{k,l}(r) \tag{C-35}$$

satisfying the condition at the origin:

$$u_{k,l}(0) = 0 \tag{C-36}$$

It is just as if we were dealing with a one-dimensional problem, where a particle of mass μ is under the influence of the potential (fig. 11):

$$\begin{aligned} &V_{\text{eff}}(r) = V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2} && \text{for } r > 0 \\ &V_{\text{eff}}(r) \quad \text{infinite} && \text{for } r < 0 \end{aligned} \tag{C-37}$$

For large r, equation (C-35) reduces to:

$$\left[\frac{d^2}{dr^2} + k^2\right] u_{k,l}(r) \underset{r\to\infty}{\simeq} 0 \tag{C-38}$$

whose general solution is:

$$u_{k,l}(r) \underset{r\to\infty}{\sim} A\, e^{ikr} + B\, e^{-ikr} \tag{C-39}$$

Since $u_{k,l}(r)$ must satisfy condition (C-36), the constants A and B cannot be arbitrary. In the equivalent one-dimensional problem [formulas (C-37)], equation (C-36) is related

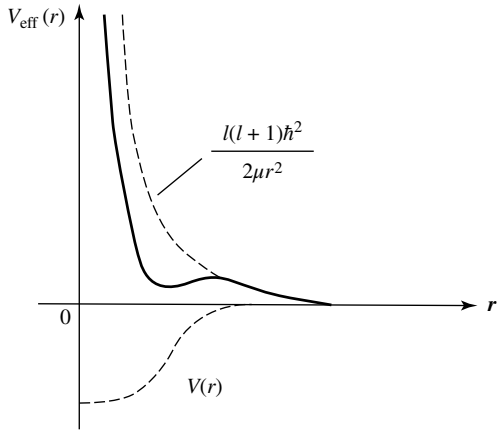


Figure 11: The effective potential $V_{\text{eff}}(r)$ is the sum of the potential $V(r)$ and the centrifugal term $\dfrac{l(l+1)\hbar^2}{2\mu r^2}$.

to the fact that the potential is infinite for negative r, and expression (C-39) represents the superposition of an "incident" plane wave e^{-ikr} coming from the right (along the axis on which the fictitious particle being studied moves) and a "reflected" plane wave e^{ikr} propagating from left to right. Since there can be no "transmitted" wave [as $V(r)$ is infinite on the negative part of the axis], the "reflected" current must be equal to the "incident" current. Thus we see that condition (C-36) implies that, in the asymptotic expression (C-39):

$$|A| = |B| \tag{C-40}$$

Consequently:

$$u_{k,l}(r) \underset{r\to\infty}{\sim} |A| \left[\mathrm{e}^{ikr}\,\mathrm{e}^{i\varphi_A} + \mathrm{e}^{-ikr}\,\mathrm{e}^{i\varphi_B}\right] \tag{C-41}$$

which can be written in the form:

$$u_{k,l}(r) \underset{r\to\infty}{\sim} C\sin\left(kr - \beta_l\right) \tag{C-42}$$

The real phase β_l is completely determined by imposing continuity between (C-42) and the solution of (C-35) which goes to zero at the origin. In the case of an identically null potential $V(r)$, we saw in § C-2-c-γ that β_l is equal to $l\pi/2$. It is convenient to take this value as a point of reference, that is, to write:

$$u_{k,l}(r) \underset{r\to\infty}{\sim} C\sin\left(kr - l\frac{\pi}{2} + \delta_l\right) \tag{C-43}$$

The quantity δ_l defined in this way is called the *phase shift* of the partial wave $\varphi_{k,l,m}(\mathbf{r})$; it obviously depends on k, that is, on the energy.

C-3-b. Physical meaning of phase shifts

α. Comparison between partial waves and free spherical waves

Taking (C-34) and (C-43) into account, we may write the expression for the asymptotic behavior of $\varphi_{k,l,m}(\mathbf{r})$ in the form:

$$\begin{aligned}\varphi_{k,l,m}(\mathbf{r}) &\underset{r\to\infty}{\sim} C\frac{\sin(kr-l\pi/2+\delta_l)}{r}Y_l^m(\theta,\varphi)\\ &\underset{r\to\infty}{\sim} -C\,Y_l^m(\theta,\varphi)\frac{\mathrm{e}^{-ikr}\,\mathrm{e}^{i\left(l\frac{\pi}{2}-\delta_l\right)}-\mathrm{e}^{ikr}\,\mathrm{e}^{-i\left(l\frac{\pi}{2}-\delta_l\right)}}{2ir}\end{aligned} \tag{C-44}$$

We see that the partial wave $\varphi_{k,l,m}(\mathbf{r})$, like a free spherical wave [formula (C-25)], results from the *superposition of an incoming wave and an outgoing wave.*

In order to develop the comparison between partial waves and free spherical waves in detail, we can modify the incoming wave of (C-44) so as to make it identical with the one in (C-25). To do this, we define a new partial wave $\widetilde{\varphi}_{k,l,m}(\mathbf{r})$ by multiplying $\varphi_{k,l,m}(\mathbf{r})$ by $\mathrm{e}^{i\delta_l}$ (this global phase factor has no physical importance) and by choosing the constant C in such a way that:

$$\widetilde{\varphi}_{k,l,m}(\mathbf{r}) \underset{r\to\infty}{\sim} -Y_l^m(\theta,\varphi)\frac{\mathrm{e}^{-ikr}\,\mathrm{e}^{il\pi/2}-\mathrm{e}^{ikr}\,\mathrm{e}^{-il\pi/2}\,\mathrm{e}^{2i\delta_l}}{2ikr} \tag{C-45}$$

This expression can then be interpreted in the following way (*cf.* the comment in § C-2-c-γ): at the outset, we have the same incoming wave as in the case of a free particle (aside from the normalization constant $\sqrt{2k^2/\pi}$). As this incoming wave approaches the zone of influence of the potential, it is more and more perturbed by this potential. When, after turning back, it is transformed into an outgoing wave, *it has accumulated a phase shift of* $2\delta_l$ relative to the free outgoing wave that would have resulted if the potential $V(r)$ had been identically zero. The factor $\mathrm{e}^{2i\delta_l}$ (which varies with l and k) thus summarizes the total effect of the potential on a particle of angular momentum l.

Comment:

Actually, the preceding discussion is only valid if we base our reasoning on a wave packet formed by superposing partial waves $\varphi_{k,l,m}(\mathbf{r})$ with the same l and m, but slightly different k. For large negative values of t, we have only an incoming wave packet; it is the subsequent evolution of this wave packet directed towards the potential's zone of influence that we have analyzed above.

We could also adopt the point of view of comment (*ii*) of § B-1-b; that is, we could study the effect on a stationary free spherical wave of slowly "turning on" the potential $V(r)$. The same type of reasoning would then demonstrate that the partial wave $\varphi_{k,l,m}(\mathbf{r})$ can be obtained from a free spherical wave $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ by adiabatically turning on the potential $V(r)$.

β. Finite-range potentials

Let us suppose that the potential $V(r)$ has a finite range r_0; that is, that:

$$V(r)=0 \qquad \text{for } r>r_0 \tag{C-46}$$

We pointed out earlier (§ C-2-c-β) that the free spherical wave $\varphi^{(0)}_{k,l,m}$ scarcely penetrates a sphere centered at O of radius $b_l(k)$ [formula (C-22)]. Therefore, if we return to the interpretation of formula (C-45) that we have just given, we see that a potential satisfying (C-46) has virtually no effect on waves for which:

$$b_l(k) \gg r_0 \tag{C-47}$$

since the corresponding incoming wave turns back before reaching the zone of influence of $V(r)$. Thus, for each value of the energy, there exists a critical value l_M of the angular momentum, which, according to (C-22), is given approximately by:

$$\sqrt{l_M(l_M+1)} \simeq kr_0 \tag{C-48}$$

The phase shifts δ_l are appreciable only for values of l less than or of the order of l_M.

The shorter the range of the potential and the lower the incident energy, the smaller the value[9] of l_M. Therefore, it may happen that the only non-zero phase shifts are those corresponding to the first few partial waves: the $s(l=0)$ wave at very low energy, followed by s and p waves for slightly greater energies, etc.

C-4. Expression of the cross section in terms of phase shifts

Phase shifts characterize the modifications, caused by the potential, of the asymptotic behavior of stationary states with well-defined angular momentum. Knowing them should therefore allow us to determine the cross section. In order to demonstrate this, all we must do is express the stationary scattering state $v^{(\text{scatt})}_k(\mathbf{r})$ in terms of partial waves[10], and calculate the scattering amplitude in this way.

C-4-a. Construction of the stationary scattering state from partial waves

We must find a linear superposition of partial waves whose asymptotic behavior is of the form (B-9). Since the stationary scattering state is an eigenstate of the Hamiltonian H, the expansion of $v^{(\text{scatt})}_k(\mathbf{r})$ involves only partial waves having the same energy $\hbar^2k^2/2\mu$. Note also that, in the case of a central potential $V(r)$, the scattering problem we are studying is symmetrical with respect to rotation around the Oz axis defined by the incident beam. Consequently, the stationary scattering wave function $v^{(\text{scatt})}_k(\mathbf{r})$ is independent of the azimuthal angle φ, so that its expansion includes only partial waves for which m is zero. Finally, we have an expression of the form:

$$v^{(\text{scatt})}_k(\mathbf{r}) = \sum_{l=0}^{\infty} c_l\, \widetilde{\varphi}_{k,l,0}(\mathbf{r}) \tag{C-49}$$

The problem thus consists of finding the coefficients c_l.

[9] l_M is of the order of kr_0, which is the ratio between the range r_0 of the potential and the wavelength of the incident particle.

[10] If there exist bound states of the particle in the potential $V(r)$ (stationary states with negative energy), the system of partial waves does not constitute a basis of the state space; in order to form such a basis, it is necessary to add the wave functions of the bound states to the partial waves.

α. Intuitive argument

When $V(r)$ is identically zero, the function $v_k^{(\text{scatt})}(\mathbf{r})$ reduces to the plane wave e^{ikz}, and the partial waves become free spherical waves $\varphi_{k,l,m}^{(0)}(\mathbf{r})$. In this case, we already know the expansion (C-49): it is given by formula (C-31).

For non-zero $V(r)$, $v_k^{(\text{scatt})}(\mathbf{r})$ includes a diverging scattered wave as well as a plane wave. Furthermore, we have seen that $\widetilde{\varphi}_{k,l,0}(\mathbf{r})$ differs from $\varphi_{k,l,0}^{(0)}(\mathbf{r})$ in its asymptotic behavior only by the presence of the outgoing wave, which has the same radial dependence as the scattered wave. We should therefore expect that the coefficients c_l of the expansion (C-49) will be the same as those in formula (C-31)[11], that is:

$$v_k^{(\text{scatt})}(\mathbf{r}) = \sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\widetilde{\varphi}_{k,l,0}(\mathbf{r}) \tag{C-50}$$

Comment:

We can also understand (C-50) in terms of the interpretation offered in comment (*ii*) of § B-1-b and the comment in § C-3-b-α. If we have a plane wave whose expansion is given by (C-31) and we turn on the potential $V(r)$ adiabatically, the wave is transformed into a stationary scattering state: the left-hand side of (C-31) must then be replaced by $v_k^{(\text{scatt})}(\mathbf{r})$. In addition, each free spherical wave $j_l(kr)Y_l^0(\theta)$ appearing on the right-hand side of (C-31) is transformed into a partial wave $\widetilde{\varphi}_{k,l,0}(\mathbf{r})$ when the potential is turned on. If we take into account the linearity of Schrödinger's equation, we finally obtain (C-50).

β. Explicit derivation

Let us now consider formula (C-50), which was suggested by a physical approach to the problem, and let us show that it does indeed supply the desired expansion.

First of all, the right-hand side of (C-50) is a superposition of eigenstates of H having the same energy $\hbar^2k^2/2\mu$; consequently, this superposition remains a stationary state.

Therefore, all we must do is make sure that the asymptotic behavior of the sum (C-50) is indeed of type (B-9). In order to do this, we use (C-45):

$$\sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\ \widetilde{\varphi}_{k,l,0}(\mathbf{r}) \underset{r\to\infty}{\sim} -\sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\ Y_l^0(\theta) \times \frac{1}{2ikr}\left[\mathrm{e}^{-ikr}\,\mathrm{e}^{il\frac{\pi}{2}} - \mathrm{e}^{ikr}\,\mathrm{e}^{-il\frac{\pi}{2}}\,\mathrm{e}^{2i\delta_l}\right] \tag{C-51}$$

In order to bring out the asymptotic behavior of expansion (C-31), we write:

$$\mathrm{e}^{2i\delta_l} = 1 + 2i\,\mathrm{e}^{i\delta_l}\sin\delta_l \tag{C-52}$$

[11] Note that the expansion (C-31) brings in $j_l(kr)Y_l^0(\theta)$, that is, the free spherical wave $\varphi_{k,l,0}^{(0)}$ divided by the normalization factor $\sqrt{2k^2/\pi}$; this is why we defined $\widetilde{\varphi}_{k,l,m}(\mathbf{r})$ [formula (C-45)] from expression (C-25) divided by this same factor.

and, regrouping the terms that are independent of δ_l, we have:

$$\sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\ \widetilde{\varphi}_{k,l,0}(\mathbf{r}) \underset{r\to\infty}{\sim} -\sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\, Y_l^0(\theta) \times \left[\frac{\mathrm{e}^{-ikr}\,\mathrm{e}^{il\pi/2} - \mathrm{e}^{ikr}\,\mathrm{e}^{-il\pi/2}}{2ikr} - \frac{\mathrm{e}^{ikr}}{r}\frac{1}{k}\,\mathrm{e}^{-il\frac{\pi}{2}}\,\mathrm{e}^{i\delta_l}\sin\delta_l\right] \quad \text{(C-53)}$$

Taking (C-25) and (C-31) into consideration, we recognize, in the first term of the right-hand side, the asymptotic expansion of the plane wave e^{ikz}, and we obtain finally:

$$\sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\ \widetilde{\varphi}_{k,l,0}(\mathbf{r}) \underset{r\to\infty}{\sim} \mathrm{e}^{ikz} + f_k(\theta)\,\frac{\mathrm{e}^{ikr}}{r} \quad \text{(C-54)}$$

with[12]:

$$\boxed{f_k(\theta) = \frac{1}{k}\sum_{t=0}^{\infty} \sqrt{4\pi(2l+1)}\ \mathrm{e}^{i\delta_l}\ \sin\ \delta_l\, Y_l^0(\theta)} \quad \text{(C-55)}$$

We have thus demonstrated that the expansion of (C-50) is correct and have found at the same time the expression for the scattering amplitude $f_k(\theta)$ in terms of the phase shifts δ_l.

C-4-b. Calculation of the cross section

The differential scattering cross section is then given by formula (B-24):

$$\sigma(\theta) = |f_k(\theta)|^2 = \frac{1}{k^2}\left|\sum_{l=0}^{\infty} \sqrt{4\pi(2l+1)}\,\mathrm{e}^{i\delta_l}\ \sin\ \delta_l\, Y_l^0(\theta)\right|^2 \quad \text{(C-56)}$$

from which we deduce the total scattering cross section by integrating over the angles:

$$\sigma = \int \mathrm{d}\Omega\,\sigma(\theta) = \frac{1}{k^2}\sum_{l,l'} 4\pi\sqrt{(2l+1)(2l'+1)}\,\mathrm{e}^{i(\delta_l-\delta_{l'})}\sin\delta_l\sin\delta_{l'} \times \int \mathrm{d}\Omega\, Y_{l'}^{0*}(\theta)\, Y_l^0(\theta) \quad \text{(C-57)}$$

Since the spherical harmonics are orthonormal [formula (D-23) of Chapter VI], we have finally:

$$\boxed{\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\ \delta_l} \quad \text{(C-58)}$$

This result shows that the terms resulting from interference between waves of different angular momenta disappear from the total cross section. For any potential $V(r)$, the

[12] The factor i^l is compensated by $\mathrm{e}^{-il\frac{\pi}{2}} = (-i)^l = (1/i)^l$.

contribution $(2l+1)(4\pi/k^2)\sin^2\delta_l$ associated with a given value of l is positive and has an upper bound, for a given energy, of $(2l+1)(4\pi/k^2)$.

In theory, formulas (C-56) and (C-58) necessitate knowing all the phase shifts δ_l. Recall (*cf.* § C-3-a) that these phase shifts can be calculated from the radial equation if the potential $V(r)$ is known; this equation must be solved separately for each value of l (most of the time, moreover, this implies resorting to numerical techniques). In other words, the method of partial waves is attractive from a practical point of view only when there is a sufficiently small number of non-zero phase shifts. For a finite-range potential $V(r)$, we saw in § C-3-b-β that the phase shifts δ_l are negligible for $l > l_M$, the critical value l_M being defined by formula (C-48).

When the potential $V(r)$ is unknown at the outset, we attempt to reproduce the experimental curves which give the differential cross section at a fixed energy by introducing a small number of non-zero phase shifts. Furthermore, the very form of the θ-dependence of the cross section often suggests the minimum number of phase shifts needed. For example, if we limit ourselves to the s-wave, formula (C-56) gives an isotropic differential cross section (Y_0^0 is a constant). If the experiments imply a variation of $\sigma(\theta)$ with θ, it means that phase shifts other than that of the s-wave are not equal to zero. Once we have thereby determined, from experimental results corresponding to different energies, the phase shifts which do effectively contribute to the cross section, we can look for theoretical models of potentials that produce these phase shifts and their energy dependence.

Comment:

The dependence of cross sections on the energy $E = \hbar^2k^2/2\mu$ of the incident particle is just as interesting as the θ-dependence of $\sigma(\theta)$. In particular, in certain cases, one observes rapid variations of the total cross section σ in the neighborhood of certain energy values. For example, if one of the phase shifts δ_l takes on the value $\pi/2$ for $E = E_0$, the corresponding contribution to σ reaches its upper limit and the cross section may show a sharp peak at $E = E_0$. This phenomenon is called "scattering resonance". We can compare it to the behavior described in Chapter I (§ D-2-c-β) of the transmission coefficient of a "square" one-dimensional potential well.

References and suggestions for further reading:

Dicke and Wittke (1.14), Chap. 16; Messiah (1.17), Chap. X; Schiff (1.18), Chaps. 5 and 9.

More advanced topics:

Coulomb scattering: Messiah (1.17), Chap. XI; Schiff (1.18), § 21; Davydov (1.20), Chap. XI, § 100.

Formal collision and S-matrix theory: Merzbacher (1.16), Chap. 19; Roman (2.3), part II, Chap. 4; Messiah (1.17), Chap. XIX; Schweber (2.16), part 3, Chap. 11.

Description of collisions in terms of wave packets: Messiah (1.17), chap. X, §§ 4, 5, 6; Goldberger and Watson (2.4), chaps. 3 and 4.

Determination of the potential from the phase shifts (the inverse problem): Wu and Ohmura (2.1), § G.

Applications: Davydov (1.20), Chap. XI; Sobel'man (11.12), Chap. 11; Mott and Massey (2.5); Martin and Spearman (16.18).

Scattering by multi-particle systems in the Born approximation and space-time correlation functions: Van Hove (2.39).

COMPLEMENTS OF CHAPTER VIII, READER'S GUIDE

The main purpose of Chapter VIII is to be a reference for other courses, for instance in nuclear physics, where various physical applications of the theory of collisions can be found.

A_{VIII} : THE FREE PARTICLE: STATIONARY STATES WITH WELL-DEFINED ANGULAR MOMENTUM	Formal examination of stationary wave functions for a free particle with well-defined angular momentum. The use of the $\mathbf{L}_+$ and $\mathbf{L}_-$ operators permits the introduction of spherical Bessel functions, and the demonstration of a certain number of their properties that were used in § C of Chapter VIII.
B_{VIII} : PHENOMENOLOGICAL DESCRIPTION OF COLLISIONS WITH ABSORPTION	This complement permits the extension of the formalism of Chapter VIII to collisions with absorption, and establishes the optical theorem. A phenomenological point of view is used, whose principle is analogous to that of Complement K_{III}. Not difficult if Chapter VIII has been well assimilated.
C_{VIII} : SOME SIMPLE APPLICATIONS OF SCATTERING THEORY	Illustration of the results of Chapter VIII by several specific examples. Section 1 is recommended for a first reading, since it presents important physical results in a simple manner (Rutherford's formula). Section 2 can be considered as a worked example. Section 3 proposes exercises without their solutions.

Complement A_{VIII}

The free particle: stationary states with well-defined angular momentum

1 The radial equation 959
2 Free spherical waves 961
2-a Recurrence relations 961
2-b Calculation of free spherical waves 962
2-c Properties 966
3 Relation between free spherical waves and plane waves . . 967

We introduced, in § C-2 of Chapter VIII, two distinct bases of stationary states of a free (spinless) particle whose Hamiltonian is written:

$$H_0 = \frac{\mathbf{P}^2}{2\mu} \tag{1}$$

The first of these bases is composed of the eigenstates common to H_0 and the three components of the momentum $\mathbf{P}$; the associated wave functions are the plane waves. The second consists of the stationary states with well-defined angular momentum, that is, the eigenstates common to H_0, $\mathbf{L}^2$ and L_z, whose principal properties we pointed out in §§ C-2-b, c and d of Chapter VIII. We intend to study here this second basis in more detail. In particular, we wish to derive a certain number of results used in Chapter VIII.

1. The radial equation

The Hamiltonian (1) commutes with the three components of the orbital angular momentum $\mathbf{L}$ of the particle:

$$[H_0, \mathbf{L}] = \mathbf{0} \tag{2}$$

Consequently, we can apply the general theory developed in § A of Chapter VII to this particular problem. We know that the free spherical waves (eigenfunctions common to H_0, $\mathbf{L}^2$ and L_z) are necessarily of the form:

$$\varphi^{(0)}_{\kappa,l,m}(\mathbf{r}) = R^{(0)}_{\kappa,l}(r)\, Y^m_l(\theta, \varphi) \tag{3}$$

The radial function $R^{(0)}_{\kappa,l}(r)$ is a solution of the equation:

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r + \frac{l(l+1)\hbar^2}{2\mu r^2}\right] R^{(0)}_{\kappa,l}(r) = E_{\kappa,l}\, R^{(0)}_{\kappa,l}(r) \tag{4}$$

where $E_{\kappa,l}$ is the eigenvalue of H_0 corresponding to $\varphi^{(0)}_{\kappa,l,m}(\mathbf{r})$. If we set:

$$R^{(0)}_{\kappa,l}(r) = \frac{1}{r}\, u^{(0)}_{\kappa,l}(r) \tag{5}$$

the function $u^{(0)}_{\kappa,l}$ is given by the equation:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{l(l+1)}{r^2} + \frac{2\mu E_{\kappa,l}}{\hbar^2}\right] u^{(0)}_{\kappa,l}(r) = 0 \tag{6}$$

to which we must add the condition:

$$u^{(0)}_{\kappa,l}(0) = 0 \tag{7}$$

It can be shown, first of all, that equations (6) and (7) enable us to find the spectrum of the Hamiltonian H_0, which we already know from the study of the plane waves [formula (C-5) of Chapter VIII]. To do this, note that the minimum value of the potential (which is, in fact, identically zero) vanishes and that consequently there cannot exist a stationary state with negative energy (*cf.* complement M_{III}). Consider, therefore, any positive value of the constant $E_{\kappa,l}$ appearing in equation (6), and set:

$$k = \frac{1}{\hbar}\sqrt{2\mu E_{\kappa,l}} \tag{8}$$

As r approaches infinity, the centrifugal term $l(l+1)/r^2$ becomes negligible compared to the constant term of equation (6), which can thus be approximated by:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + k^2\right] u^{(0)}_{\kappa,l}(r) \underset{r\to\infty}{\simeq} 0 \tag{9}$$

Consequently, all solutions of equation (6) have an asymptotic behavior (linear combination of e^{ikr} and e^{-ikr}) which is physically acceptable. Therefore, the only restriction comes from condition (7): we know (*cf.* Chap. VII, § A-3-b) that there exists, for a given value of $E_{\kappa,l}$, one and only one function (to within a constant factor) which satisfies (6) and (7). For any positive $E_{\kappa,l}$, the radial equation (6) has one and only one acceptable solution.

This means that the spectrum of H_0 includes all positive energies. Moreover, we see that the set of possible values of $E_{\kappa,l}$ does not depend on l; we shall therefore omit the index l for the energies. As for the index κ, we shall identify it with the constant defined in (8); this allows us to write:

$$E_k = \frac{\hbar^2 k^2}{2\mu} \quad ; \quad k \geq 0 \tag{10}$$

Each of these energies is infinitely degenerate. Indeed, for fixed k, there exists an acceptable solution $u^{(0)}_{\kappa,l}(r)$ of the radial equation corresponding to the energy E_k for every value (positive integral or zero) of l. Moreover, formula (3) associates $(2l+1)$ independent wave functions $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ with a given radial function $u^{(0)}_{\kappa,l}(r)$. Thus, we again find in this particular case the general result demonstrated in § A-3-b of Chapter VII: H_0, $\mathbf{L}^2$ and L_z form a C.S.C.O. in $\mathcal{E}_{\mathbf{r}}$, and the specification of the three indices k, l and m gives sufficient information for the determination of a unique function in the corresponding basis.

2. Free spherical waves

The radial functions $R^{(0)}_{k,l}(r) = u^{(0)}_{k,l}(r)/r$ can be found by solving equation (6) or equation (4) directly. The latter is easily reduced (comment of § 2-c-β below) to a differential equation known as the "spherical Bessel equation" whose solutions are well-known. Instead of using these results directly, we are going to see how the various eigenfunctions common to H_0, $\mathbf{L}^2$ and L_z can be simply deduced from those which correspond to the eigenvalue 0 of $\mathbf{L}^2$.

2-a. Recurrence relations

Let us define the operator:

$$P_+ = P_x + iP_y \tag{11}$$

in terms of the components P_x and P_y of the momentum $\mathbf{P}$. We know that $\mathbf{P}$ is a vectorial observable (*cf.* complement $\mathrm{B_{VI}}$, § 5-c), which implies the following commutation relations[1] between its components and those of the angular momentum $\mathbf{L}$:

$$\begin{aligned}
&[L_x, P_x] = 0 \\
&[L_x, P_y] = i\hbar\, P_z \\
&[L_x, P_z] = -i\hbar\, P_y
\end{aligned} \tag{12}$$

and the equations which are deduced from these by circular permutation of the indices x, y, z. Using these relations, a simple algebraic calculation gives the commutators of L_z and $\mathbf{L}^2$ with the operator P_+; we find:

$$[L_z, P_+] = \hbar\, P_+ \tag{13a}$$

$$[\mathbf{L}^2, P_+] = 2\hbar\,(P_+ L_z - P_z L_+) + 2\hbar^2\, P_+ \tag{13b}$$

Consider therefore any eigenfunction $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ common to H_0, $\mathbf{L}^2$ and L_z, the corresponding eigenvalues being E_k, $l(l+1)\hbar^2$ and $m\hbar$. By applying the operators L_+ and L_-, we can obtain the $2l$ other eigenfunctions associated with the same energy E_k and the same value of l. Since H_0 commutes with $\mathbf{L}$, we have, for example:

$$H_0 L_+\, \varphi^{(0)}_{k,l,m}(\mathbf{r}) = L_+ H_0\, \varphi^{(0)}_{k,l,m}(\mathbf{r}) = E_k L_+\, \varphi^{(0)}_{k,l,m}(\mathbf{r}) \tag{14}$$

and $L_+\, \varphi^{(0)}_{k,l,m}(\mathbf{r})$ (which is not zero if m is different from l) is an eigenfunction of H_0 with the same eigenvalue as $\varphi^{(0)}_{k,l,m}(\mathbf{r})$. Therefore:

$$L_\pm\, \varphi^{(0)}_{k,l,m}(\mathbf{r}) \propto \varphi^{(0)}_{k,l,m\pm1}(\mathbf{r}) \tag{15}$$

Let us now allow P_+ to act on $\varphi^{(0)}_{k,l,m}(\mathbf{r})$. First of all, since H_0 commutes with $\mathbf{P}$, we can repeat the preceding argument for $P_+\, \varphi^{(0)}_{k,l,m}$. Moreover, from relation (13a):

$$\begin{aligned}
L_z P_+\, \varphi^{(0)}_{k,l,m}(\mathbf{r}) &= P_+ L_z\, \varphi^{(0)}_{k,l,m} + \hbar P_+\, \varphi^{(0)}_{k,l,m} \\
&= (m+1)\hbar\, P_+\, \varphi^{(0)}_{k,l,m}(\mathbf{r})
\end{aligned} \tag{16}$$

[1] These relations can be obtained directly from the definition $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ and the canonical commutation rules.

$P_+ \varphi^{(0)}_{k,l,m}$ is therefore an eigenfunction of L_z with the eigenvalue $(m+1)\hbar$. If we use equation (13b) in the same way, we see that the presence of the term $P_z L_+$ implies that $P_+ \varphi^{(0)}_{k,l,m}$ is not, in general, an eigenfunction of $\mathbf{L}^2$; nevertheless, if $m = l$, the contribution of this term is zero:

$$\begin{aligned}\mathbf{L}^2 P_+ \varphi^{(0)}_{k,l,l} &= P_+ \mathbf{L}^2 \varphi^{(0)}_{k,l,l} + 2\hbar P_+ L_z \varphi^{(0)}_{k,l,l} + 2\hbar^2 P_+ \varphi^{(0)}_{k,l,l} \\ &= [l(l+1) + 2l + 2]\, \hbar^2 P_+ \varphi^{(0)}_{k,l,l} \\ &= (l+1)(l+2)\hbar^2 P_+ \varphi^{(0)}_{k,l,l} \end{aligned} \tag{17}$$

Consequently, $P_+ \varphi^{(0)}_{k,l,l}$ is a common eigenfunction of H_0, $\mathbf{L}^2$ and L_z with the eigenvalues E_k, $(l+1)(l+2)\hbar^2$ and $(l+1)\hbar$ respectively. Since these three observables form a C.S.C.O. (§ 1), there exists only one eigenfunction (to within a constant factor[2]) associated with this set of eigenvalues:

$$P_+ \varphi^{(0)}_{k,l,l}(\mathbf{r}) \propto \varphi^{(0)}_{k,l+1,l+1}(\mathbf{r}) \tag{18}$$

We are going to use the recurrence relations (15) and (18) to construct the $\{\varphi^{(0)}_{k,l,m}(\mathbf{r})\}$ basis from the functions $\varphi^{(0)}_{k,0,0}(\mathbf{r})$ corresponding to zero eigenvalues[3] for $\mathbf{L}^2$ and L_z.

2-b. Calculation of free spherical waves

α. Solution of the radial equation for $l = 0$

In order to determine the functions $\varphi^{(0)}_{k,0,0}(\mathbf{r})$, we return to the radial equation (6), in which we set $l = 0$; taking definition (10) into account, this equation can be written:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + k^2\right] u^{(0)}_{k,0}(r) = 0 \tag{19}$$

The solution which goes to zero at the origin [condition (7)] is of the form:

$$u^{(0)}_{k,0}(r) = a_k \sin kr \tag{20}$$

We choose the constant a_k such that the functions $\varphi^{(0)}_{k,0,0}(\mathbf{r})$ are orthonormal in the extended sense; that is:

$$\int \mathrm{d}^3 r\, \varphi^{(0)*}_{k,0,0}(\mathbf{r})\, \varphi^{(0)}_{k',0,0}(\mathbf{r}) = \delta(k - k') \tag{21}$$

It is easy to show (see below) that condition (21) is satisfied if:

$$a_k = \sqrt{\frac{2}{\pi}} \tag{22}$$

[2] Later (§ 2-b), we shall specify the coefficients that ensure the orthonormalization of the $\{\varphi^{(0)}_{k,l,m}(\mathbf{r})\}$ basis (in the extended sense, since k is a continuous index).

[3] It must not be thought that the operator $P_- = P_x - iP_y$ allows one to "step down" from an arbitrary value of l to zero. It can easily be shown, by an argument analogous to the preceding one, that:

$$P_- \varphi^{(0)}_{k,l,-l}(\mathbf{r}) \propto \varphi^{(0)}_{k,l+1,-(l+1)}(\mathbf{r})$$

.

which yields (Y_0^0 being equal to $1/\sqrt{4\pi}$):

$$\varphi^{(0)}_{k,0,0}(\mathbf{r}) = \sqrt{\frac{2k^2}{\pi}}\frac{1}{\sqrt{4\pi}}\frac{\sin kr}{kr} \tag{23}$$

Let us verify that the functions (23) satisfy the orthonormalization relation (21). To do this, it is sufficient to calculate:

$$\begin{aligned}\int \mathrm{d}^3r\,\varphi^{(0)*}_{k,0,0}(\mathbf{r})\,\varphi^{(0)}_{k',0,0}(\mathbf{r}) &= \frac{2}{\pi}kk'\frac{1}{4\pi}\int_0^\infty r^2\mathrm{d}r\frac{\sin kr}{kr}\,\frac{\sin k'r}{k'r}\int \mathrm{d}\Omega \\ &= \frac{2}{\pi}\int_0^\infty \mathrm{d}r\sin\,kr\,\sin\,k'r \end{aligned}\tag{24}$$

Replacing the sines by complex exponentials and extending the interval of integration over the range $-\infty$ to $+\infty$, we obtain:

$$\frac{2}{\pi}\int_0^\infty \mathrm{d}r\sin\,kr\,\sin\,k'r = \frac{2}{\pi}\left(-\frac{1}{4}\right)\int_{-\infty}^{+\infty}\mathrm{d}r\left[\mathrm{e}^{i(k+k')r}-\mathrm{e}^{i(k-k')r}\right] \tag{25}$$

Since k and k' are both positive, $k+k'$ is always different from zero and the contribution of the first term within the brackets is always zero. According to formula (34) of Appendix I, the second term yields finally:

$$\begin{aligned}\int \mathrm{d}^3r\,\varphi^{(0)*}_{k,0,0}(\mathbf{r})\,\varphi^{(0)}_{k',0,0}(\mathbf{r}) &= \frac{2}{\pi}\left(-\frac{1}{4}\right)(-2\pi)\,\delta(k-k') \\ &= \delta(k-k')\end{aligned}\tag{26}$$

β. Construction of the other waves by recurrence

Let us now apply the operator P_+ defined in (11) to the function $\varphi^{(0)}_{k,0,0}(\mathbf{r})$ that we have just found. According to relation (18):

$$\begin{aligned}\varphi^{(0)}_{k,1,1}(\mathbf{r}) &\propto P_+\,\varphi^{(0)}_{k,0,0}(\mathbf{r}) \\ &\propto P_+\,\frac{\sin kr}{kr}\end{aligned}\tag{27}$$

In the $\{|\mathbf{r}\rangle\}$ representation, which we have been using throughout, P_+ is the differential operator:

$$P_+ = \frac{\hbar}{i}\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right) \tag{28}$$

In formula (27), it acts on a function of r alone. Now:

$$\begin{aligned}P_+f(r) &= \frac{\hbar}{i}\left(\frac{x}{r}+i\frac{y}{r}\right)\frac{\mathrm{d}}{\mathrm{d}r}f(r) \\ &= \frac{\hbar}{i}\sin\theta\,\mathrm{e}^{i\varphi}\frac{\mathrm{d}}{\mathrm{d}r}f(r)\end{aligned}\tag{29}$$

Thus we obtain:

$$\varphi^{(0)}_{k,1,1}(\mathbf{r}) \propto \sin\theta\,\mathrm{e}^{i\varphi}\left[\frac{\cos kr}{kr}-\frac{\sin kr}{(kr)^2}\right] \tag{30}$$

We recognize the angular dependence of $Y_1^1(\theta,\varphi)$ [complement A_{VI}, formula (32)]; by applying L_-, the functions $\varphi^{(0)}_{k,1,0}(\mathbf{r})$ and $\varphi^{(0)}_{k,1,-1}(\mathbf{r})$ can be calculated.

Although $\varphi^{(0)}_{k,1,1}(\mathbf{r})$ depends on θ and φ, the application of P_+ to this function remains very simple. The canonical commutation relations indicate immediately that:

$$[P_+, X + iY] = 0 \tag{31}$$

Consequently, $\varphi^{(0)}_{k,2,2}(\mathbf{r})$ is given by:

$$\begin{aligned}\varphi^{(0)}_{k,2,2}(\mathbf{r}) &\propto P_+^2 \frac{\sin kr}{kr} \\ &\propto P_+ \frac{x+iy}{r} \frac{\mathrm{d}}{\mathrm{d}r} \frac{\sin kr}{kr} \\ &\propto (x+iy) P_+ \frac{1}{r} \frac{\mathrm{d}}{\mathrm{d}r} \frac{\sin kr}{kr} \\ &\propto (x+iy)^2 \frac{1}{r} \frac{\mathrm{d}}{\mathrm{d}r} \left[\frac{1}{r} \frac{\mathrm{d}}{\mathrm{d}r} \frac{\sin kr}{kr} \right]\end{aligned} \tag{32}$$

In general:

$$\varphi^{(0)}_{k,l,l}(\mathbf{r}) \propto (x+iy)^l \left(\frac{1}{r} \frac{\mathrm{d}}{\mathrm{d}r} \right)^l \frac{\sin kr}{kr} \tag{33}$$

The angular dependence of $\varphi^{(0)}_{k,l,l}$ is contained in the factor:

$$(x+iy)^l = r^l (\sin\theta)^l \, \mathrm{e}^{il\varphi} \tag{34}$$

which is indeed proportional to $Y_l^l(\theta,\varphi)$.

Let us define:

$$j_l(\rho) = (-1)^l \, \rho^l \left(\frac{1}{\rho} \frac{\mathrm{d}}{\mathrm{d}\rho} \right)^l \frac{\sin\rho}{\rho} \tag{35}$$

j_l, thus defined, is the *spherical Bessel function of order l*. The preceding calculation shows that $\varphi^{(0)}_{k,l,l}(\mathbf{r})$ is proportional to the product of $Y_l^l(\theta,\varphi)$ and $j_l(kr)$. We shall write (see below the problem of normalization):

$$R^{(0)}_{k,l}(r) = \sqrt{\frac{2k^2}{\pi}} j_l(kr) \tag{36}$$

The free spherical waves are then written:

$$\varphi^{(0)}_{k,l,m}(\mathbf{r}) = \sqrt{\frac{2k^2}{\pi}} j_l(kr) \, Y_l^m(\theta,\varphi) \tag{37}$$

They satisfy the orthonormalization relation:

$$\int \mathrm{d}^3 r \, \varphi^{(0)*}_{k,l,m}(\mathbf{r}) \, \varphi^{(0)}_{k',l',m'}(\mathbf{r}) = \delta(k-k') \, \delta_{ll'} \, \delta_{mm'} \tag{38}$$

and the closure relation:

$$\int_0^\infty \mathrm{d}k \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \varphi^{(0)}_{k,l,m}(\mathbf{r}) \, \varphi^{(0)*}_{k,l,m}(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \tag{39}$$

Let us now examine the normalization of the functions (37). To do so, let us begin by specifying the proportionality factors of the recurrence relations (15) and (18). For the first relation, we already know this factor from the properties of spherical harmonics (*cf.* complement A_{VI}):

$$L_{\pm}\,\varphi^{(0)}_{k,l,m}(\mathbf{r}) = \hbar\sqrt{l(l+1)-m(m\pm 1)}\,\varphi^{(0)}_{k,l,m\pm 1}(\mathbf{r}) \tag{40}$$

As for relation (18), we proceed as follows. Using the explicit expression for $Y_l^l(\theta,\varphi)$ [formulas (4) and (14) of complement A_{VI}], equations (31) and (29) as well as definition (35), and taking (37) into account, we can write this relation as:

$$P_+\,\varphi^{(0)}_{k,l,l}(\mathbf{r}) = \frac{\hbar k}{i}\sqrt{\frac{2l+2}{2l+3}}\,\varphi^{(0)}_{k,l+1,l+1}(\mathbf{r}) \tag{41}$$

In the orthonormalization relation (38), the factors $\delta_{ll'}\delta_{mm'}$ on the right-hand side arise from the angular integration and the orthonormality of the spherical harmonics. To establish relation (38), it is thus sufficient to show that the integral:

$$I_l(k,k') = \int \mathrm{d}^3r\,\varphi^{(0)*}_{k,l,l}(\mathbf{r})\,\varphi^{(0)}_{k',l,l}(\mathbf{r}) \tag{42}$$

is equal to $\delta(k-k')$. We already know from (26) that $I_0(k,k')$ has this value. Consequently, we shall demonstrate that, if:

$$I_l(k,k') = \delta(k-k') \tag{43}$$

then the same is true for $I_{l+1}(k,k')$. Relation (41) permits us to write $I_{l+1}(k,k')$ in the form:

$$\begin{aligned} I_{l+1}(k,k') &= \frac{1}{\hbar^2 kk'}\,\frac{2l+3}{2l+2}\int \mathrm{d}^3r\left[P_+\,\varphi^{(0)}_{k,l,l}(\mathbf{r})\right]^*\left[P_+\,\varphi^{(0)}_{k',l,l}(\mathbf{r})\right] \\ &= \frac{1}{\hbar^2 kk'}\,\frac{2l+3}{2l+2}\int \mathrm{d}^3r\,\varphi^{(0)*}_{k,l,l}(\mathbf{r})\,P_-\,P_+\,\varphi^{(0)}_{k',l,l}(\mathbf{r}) \end{aligned} \tag{44}$$

where $P_- = P_x - iP_y$ is the adjoint of P_+. Now:

$$P_-\,P_+ = P_x^2 + P_y^2 = \mathbf{P}^2 - P_z^2 \tag{45}$$

We know that $\varphi^{(0)}_{k',l,l}$ is an eigenfunction of $\mathbf{P}^2$. Since, in addition, P_z is Hermitian, it results that:

$$I_{l+1}(k,k') = \frac{1}{\hbar^2 kk'}\,\frac{2l+3}{2l+2}\left\{\hbar^2 k'^2 I_l(k,k') - \int \mathrm{d}^3r\left[P_z\,\varphi^{(0)}_{k,l,l}(\mathbf{r})\right]^*\left[P_z\,\varphi^{(0)}_{k',l,l}(\mathbf{r})\right]\right\} \tag{46}$$

We must now calculate $P_z\,\varphi^{(0)}_{k,l,l}(\mathbf{r})$. Using the fact that $Y_l^l(\theta,\varphi)$ is proportional to $(x+iy)^l/r^l$, we easily find:

$$\begin{aligned} P_z\,\varphi^{(0)}_{k,l,l}(\mathbf{r}) &= -\frac{\hbar k}{i}\sqrt{\frac{2k^2}{\pi}}\,\cos\theta\,Y_l^l(\theta,\varphi)\,j_{l+1}(kr) \\ &= -\frac{\hbar k}{i}\,\frac{1}{\sqrt{2l+3}}\,\varphi^{(0)}_{k,l+1,l}(\mathbf{r}) \end{aligned} \tag{47}$$

according to formula (35) of complement A_{VI}. Putting this result into (46), we finally obtain:

$$I_{l+1}(k,k') = \frac{2l+3}{2l+2}\,\frac{k'}{k}\,I_l(k,k') - \frac{1}{2l+2}\,I_{l+1}(k,k') \tag{48}$$

Hypothesis (43) thus implies:

$$I_{l+1}(k,k') = \delta(k-k') \tag{49}$$

which concludes the argument by recurrence.

2-c. Properties

α. Behavior at the origin

When ρ approaches zero, the function $j_l(\rho)$ behaves (see below) like:

$$j_l(\rho) \underset{\rho\to 0}{\sim} \frac{\rho^l}{(2l+1)!!} \tag{50}$$

Consequently, $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ is proportional to r^l in the neighborhood of the origin:

$$\varphi^{(0)}_{k,l,m}(\mathbf{r}) \underset{r\to 0}{\sim} \sqrt{\frac{2k^2}{\pi}} Y_l^m(\theta,\varphi) \frac{(kr)^l}{(2l+1)!!} \tag{51}$$

To demonstrate formula (50), starting from definition (35), it is sufficient to expand $\sin\rho/\rho$ in a power series in ρ :

$$\frac{\sin\rho}{\rho} = \sum_{p=0}^{\infty} (-1)^p \frac{\rho^{2p}}{(2p+1)!} \tag{52}$$

We then apply the operator $\left(\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^l$, which yields:

$$\begin{aligned} j_l(\rho) &= (-1)^l \rho^l \left(\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^{l-1} \sum_{p=0}^{\infty} (-1)^p \frac{2p}{(2p+1)!} \rho^{2p-1-1} \\ &= (-1)^l \rho^l \sum_{p=0}^{\infty} (-1)^p \frac{2p(2p-2)(2p-4)\ldots[2p-2(l-1)]}{(2p+1)!} \rho^{2p-2l} \end{aligned} \tag{53}$$

The first l terms of the sum ($p = 0$ to $l-1$) are zero, and the $(l+1)$th is written:

$$j_l(\rho) \underset{\rho\to 0}{\sim} (-1)^l \rho^l (-1)^l \frac{2l(2l-2)(2l-4)\ldots 2}{(2l+1)!} \tag{54}$$

which proves (50).

β. Asymptotic behavior

When their argument approaches infinity, the spherical Bessel functions are related to the trigonometric functions in the following way:

$$j_l(\rho) \underset{\rho\to\infty}{\sim} \frac{1}{\rho} \sin\left(\rho - l\frac{\pi}{2}\right) \tag{55}$$

The asymptotic behavior of the free spherical waves is therefore:

$$\varphi^{(0)}_{k,l,m}(\mathbf{r}) \underset{r\to\infty}{\sim} \sqrt{\frac{2k^2}{\pi}} Y_l^m(\theta,\varphi) \frac{\sin(kr - l\pi/2)}{kr} \tag{56}$$

If we apply the operator $\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}$ once to $\frac{\sin\rho}{\rho}$, we can write $j_l(\rho)$ in the form:

$$j_l(\rho) = (-1)^l \rho^l \left(\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^{l-1} \left[\frac{\cos\rho}{\rho^2} - \frac{\sin\rho}{\rho^3}\right] \tag{57}$$

The second term inside the brackets is negligible compared to the first term when ρ approaches infinity. Moreover, when we apply $\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}$ a second time, the dominant term still comes from the derivative of the cosine. Thus we see that:

$$j_l(\rho) \underset{\rho\to\infty}{\sim} (-1)^l \rho^l \frac{1}{\rho^l}\frac{1}{\rho}\left(\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^l \sin\rho \tag{58}$$

Since:

$$\left(\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^l \sin\rho = (-1)^l \sin\left(\rho - l\frac{\pi}{2}\right) \tag{59}$$

the result is indeed (55).

Comment:

If we set:

$$kr = \rho \tag{60}$$

[k being defined by formula (10)], the radial equation (4) becomes:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \frac{2}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho} + \left(1 - \frac{l(l+1)}{\rho^2}\right)\right] R_l(\rho) = 0 \tag{61}$$

This is the spherical Bessel equation of order l. It has two linearly independent solutions, which can be distinguished, for example, by their behavior at the origin. One of them is the spherical Bessel function $j_l(\rho)$, which satisfies (50) and (55). For the other, we can choose the "spherical Neumann function of order l", designated as $n_l(\rho)$, with the properties:

$$n_l(\rho) \underset{\rho\to 0}{\sim} \frac{(2l-1)!!}{\rho^{l+1}} \tag{62a}$$

$$n_l(\rho) \underset{\rho\to\infty}{\sim} \frac{1}{\rho}\cos\left(\rho - l\frac{\pi}{2}\right) \tag{62b}$$

3. Relation between free spherical waves and plane waves

We already know two distinct bases of eigenstates of H_0: the plane waves $v_{\mathbf{k}}^{(0)}(\mathbf{r})$ are eigenfunctions of the three components of the momentum $\mathbf{P}$; the free spherical waves $\varphi_{k,l,m}^{(0)}(\mathbf{r})$ are eigenfunctions of $\mathbf{L}^2$ and L_z. These two bases are different because $\mathbf{P}$ does not commute with $\mathbf{L}^2$ and L_z.

A given function of one of these bases can obviously be expanded in terms of the other basis. For example, we shall express a plane wave $v_{\mathbf{k}}^{(0)}(\mathbf{r})$ as a linear superposition of free spherical waves. Consider, therefore, a vector $\mathbf{k}$ in ordinary space. The plane wave $v_{\mathbf{k}}^{(0)}(\mathbf{r})$ that it characterizes is an eigenfunction of H_0 with the eigenvalue $\hbar^2\mathbf{k}^2/2\mu$. Therefore, its expansion will include only the $\varphi_{k,l,m}^{(0)}$ which correspond to this energy, that is those for which:

$$k = |\mathbf{k}| \tag{63}$$

This expansion will therefore be of the form:

$$v_{\mathbf{k}}^{(0)}(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} c_{l,m}(\mathbf{k})\, \varphi_{k,l,m}^{(0)}(\mathbf{r}) \tag{64}$$

the free indices $\mathbf{k}$ and k being related by equation (63). It is easy to show, using the properties of the spherical harmonics (*cf.* complement A_{VI}) and the spherical Bessel functions, that:

$$\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} i^l\, Y_l^{m*}(\theta_k, \varphi_k)\, j_l(kr)\, Y_l^m(\theta, \varphi) \tag{65}$$

where θ_k and φ_k are the polar angles that fix the direction of the vector $\mathbf{k}$. If $\mathbf{k}$ is directed along Oz, expansion (65) reduces to:

$$\begin{aligned} \mathrm{e}^{ikz} &= \sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\, j_l(kr)\, Y_l^0(\theta) \\ &= \sum_{l=0}^{\infty} i^l (2l+1) j_l(kr)\, P_l(\cos\theta) \end{aligned} \tag{66}$$

where P_l is the Legendre polynomial of degree l [*cf.* equation (57) of complement A_{VI}].

Let us first demonstrate relation (66). To do this, let us assume that the vector $\mathbf{k}$ chosen is collinear with Oz:

$$k_x = k_y = 0 \tag{67}$$

and points in the same direction. In this case, equation (63) becomes:

$$k_z = k \tag{68}$$

and we want to expand the function:

$$\mathrm{e}^{ikz} = \mathrm{e}^{ikr\cos\theta} \tag{69}$$

in the $\{\varphi_{k,l,m}^{(0)}(\mathbf{r})\}$ basis. Since this function is independent of the angle φ, it is a linear combination of only those basis functions for which $m = 0$:

$$\begin{aligned} \mathrm{e}^{ikr\cos\theta} &= \sum_{l=0}^{\infty} a_l\, \varphi_{k,l,0}^{(0)}(\mathbf{r}) \\ &= \sum_{l=0}^{\infty} c_l\, j_l(kr)\, Y_l^0(\theta) \end{aligned} \tag{70}$$

To calculate the numbers c_l, we can consider $\mathrm{e}^{ikr\cos\theta}$ to be a function of the variable θ, with r playing the role of a parameter. Since the spherical harmonics form an orthonormal basis for functions of θ and φ, the "coefficient" $c_l\, j_l(kr)$ can be expressed as:

$$c_l\, j_l(kr) = \int \mathrm{d}\Omega\, Y_l^{0*}(\theta)\, \mathrm{e}^{ikr\cos\theta} \tag{71}$$

Replacing Y_l^0 by its expression in terms of $Y_l^l(\theta, \varphi)$ [formula (25) of complement A_{VI}], we obtain:

$$\begin{aligned} c_l\, j_l(kr) &= \frac{1}{\sqrt{(2l)!}} \int \mathrm{d}\Omega \left[\left(\frac{L_-}{\hbar} \right)^l Y_l^l(\theta, \varphi) \right]^* \mathrm{e}^{ikr\cos\theta} \\ &= \frac{1}{\sqrt{(2l)!}} \int \mathrm{d}\Omega\, Y_l^{l*}(\theta, \varphi) \left[\left(\frac{L_+}{\hbar} \right)^l \mathrm{e}^{ikr\cos\theta} \right] \end{aligned} \tag{72}$$

since L_+ is the adjoint operator of L_-. Formula (16) of complement $\mathrm{A_{VI}}$ then yields:

$$\left(\frac{L_+}{\hbar}\right)^l \mathrm{e}^{ikr\cos\theta} = (-1)^l\,\mathrm{e}^{il\varphi}(\sin\theta)^l\frac{\mathrm{d}^l}{\mathrm{d}(\cos\theta)^l}\,\mathrm{e}^{ikr\cos\theta}$$

$$= (-1)^l\,\mathrm{e}^{il\varphi}(\sin\theta)^l(ikr)^l\,\mathrm{e}^{ikr\cos\theta} \tag{73}$$

Now $(\sin\theta)^l\,\mathrm{e}^{il\varphi}$ is just, to within a constant factor, $Y_l^l(\theta,\varphi)$ [*cf.* formulas (4) and (14) of complement $\mathrm{A_{VI}}$]. Consequently:

$$c_l\,j_l(kr) = (ikr)^l\frac{2^l\,l!}{\sqrt{(2l)!}}\sqrt{\frac{4\pi}{(2l+1)!}}\int d\Omega\left|Y_l^l(\theta,\varphi)\right|^2\mathrm{e}^{ikr\cos\theta} \tag{74}$$

It is therefore sufficient to choose a particular value of kr, for which we know the value $j_l(kr)$, in order to calculate c_l. Allow, for example, kr to approach zero: we know that $j_l(kr)$ behaves like $(kr)^l$, and so, in fact, does the right-hand side of equation (74). More precisely, using (50), we find:

$$c_l\frac{1}{(2l+1)!!} = i^l\frac{2^l\,l!}{\sqrt{(2l)!}}\sqrt{\frac{4\pi}{(2l+1)!}}\int \mathrm{d}\Omega\left|Y_l^l(\theta,\varphi)\right|^2 \tag{75}$$

that is, since Y_l^l is normalized to 1:

$$c_l = i^l\sqrt{4\pi(2l+1)} \tag{76}$$

This proves formula (66).

The general relation (65) can therefore be obtained as a consequence of the addition theorem for spherical harmonics [formula (70) of complement $\mathrm{A_{VI}}$]. Whatever the direction of $\mathbf{k}$ (defined by the polar angles θ_k and φ_k), it is always possible, through a rotation of the system of axes, to return to the case we have just considered. Consequently, expansion (66) remains valid, provided kz is replaced by $\mathbf{k}\cdot\mathbf{r}$ and $\cos\theta$ by $\cos\alpha$, where α is the angle between $\mathbf{k}$ and $\mathbf{r}$:

$$\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{l=0}^{\infty} i^l\,(2l+1)\,j_l(kr)\,P_l(\cos\alpha) \tag{77}$$

The addition theorem for spherical harmonics permits the expression of $P_l(\cos\alpha)$ in terms of the angles (θ,φ) and (θ_k,φ_k), which yields finally formula (65).

Expansions (65) and (66) show that *a state of well-defined linear momentum involves all possible orbital angular momenta.*

To obtain the expansion of a given function $\varphi^{(0)}_{k,l,m}(\mathbf{r})$ in terms of plane waves, it is sufficient to invert formula (65), using the orthonormalization relation of spherical harmonics which are functions of θ_k and φ_k. This yields:

$$\int \mathrm{d}\Omega\; Y_l^m(\theta_k,\varphi_k)\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi i^l\,j_l(kr)\,Y_l^m(\theta,\varphi) \tag{78}$$

Thus we find:

$$\varphi^{(0)}_{k,l,m}(\mathbf{r}) = \frac{(-1)^l}{4\pi}\,i^l\sqrt{\frac{2k^2}{\pi}}\int \mathrm{d}\Omega_k\;Y_l^m(\theta_k,\varphi_k)\;\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} \tag{79}$$

An eigenfunction of $\mathbf{L}^2$ and L_z is therefore a linear superposition of all plane waves with the same energy: *a state of well-defined angular momentum involves all possible directions of the linear momentum.*

References:

Messiah (1.17), App. B, § 6; Arfken (10.4), § 11.6; Butkov (10.8), Chap. 9, § 9; see the subsection "Special functions and tables" of section 10 of the bibliography.

Complement B_{VIII}

Phenomenological description of collisions with absorption

1 Principle involved 971
2 Calculation of the cross sections 972
2-a Elastic scattering cross section 972
2-b Absorption cross section 973
2-c Total cross section. Optical theorem 975

In Chapter VIII, we confined ourselves to the study of the elastic[1] scattering of particles by a potential. But we pointed out in the introduction of that chapter that collisions between particles can be inelastic and lead, under certain conditions, to numerous other reactions (creation or destruction of particles, etc...), particularly if the energy of the incident particles is high. When such reactions are possible, and one detects only elastically scattered particles, one observes that certain particles of the incident beam "disappear"; that is, they are not to be found either in the transmitted beam or amongst the elastically scattered particles. These particles are said to be "absorbed" during the interaction; in reality, they have taken part in reactions other than that of simple elastic scattering. If one is interested only in the elastic scattering, one seeks to describe the "absorption" globally, without going into detail about the other possible reactions. We are going to show here that the method of partial waves provides a convenient framework for such a phenomenological description.

1. Principle involved

We shall assume that the interactions responsible for the disappearance of the incident particles are invariant with respect to rotation about O. The scattering amplitude can therefore always be decomposed into partial waves, each of which corresponds to a fixed value of the angular momentum.

In this section, we shall see how the method of partial waves can be modified to take a possible absorption into consideration. To do this, let us return to the interpretation of partial waves that we gave in § C-3-b-α of Chapter VIII. A free incoming wave penetrates the zone of influence of the potential and gives rise to an outgoing wave. The effect of the potential is to multiply this outgoing wave by $e^{2i\delta_l}$. Since the modulus of this factor is 1 (the phase shift δ_l is real), the amplitude of the outgoing wave is equal to that of the incoming wave. Consequently (see the calculation of § 2-b below), the total flux of the incoming wave is equal to that of the outgoing wave: during the scattering, probability is conserved, that is, the total number of particles is constant. These considerations suggest that, in the cases where absorption phenomena occur, one can take them into account simply by giving the phase shift an imaginary part such that:

$$\left|e^{2i\delta_l}\right| < 1 \tag{1}$$

[1] A collision is called elastic if it changes neither the nature nor the internal state of the concerned particles; otherwise it is called inelastic.

The amplitude of the outgoing wave with angular momentum l is thus smaller than that of the incoming wave from which it arises. The fact that the outgoing probability flux is smaller than the incoming flux expresses the "disappearance" of a certain number of particles.

We are going to make this idea more explicit and deduce from it the expressions for the scattering and absorption cross sections. However, we stress the fact that this is a purely phenomenological method: the parameters with which we shall characterize the absorption (modulus of $e^{2i\delta_l}$ for each partial wave) mask an often very complicated reality. Note also that if the total probability is no longer conserved it is impossible to describe the interaction by a simple potential. A correct treatment of the set of phenomena which can then arise during the collision would demand a more elaborate formalism than the one developed in Chapter VIII.

2. Calculation of the cross sections

We return to the calculations of § C-4 of Chapter VIII, setting:

$$\eta_l = e^{2i\delta_l} \tag{2}$$

Since the possibility of producing reactions other than that of elastic scattering is always expressed by a decrease in the number of elastically scattered particles, we must have:

$$|\eta_l| \leqslant 1 \tag{3}$$

(equality corresponding to cases where only elastic scattering is possible). The asymptotic form of the wave function which describes the elastic scattering is therefore [*cf.* formula (C-51) of Chapter VIII]:

$$v_k^{(\text{scatt})}(\mathbf{r}) \underset{r\to\infty}{\sim} -\sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)} Y_l^0(\theta) \frac{e^{-ikr}\, e^{il\frac{\pi}{2}} - \eta_l\, e^{ikr}\, e^{-il\frac{\pi}{2}}}{2ikr} \tag{4}$$

2-a. Elastic scattering cross section

The argument of § C-4-a of Chapter VIII remains valid and gives the scattering amplitude $f_k(\theta)$ in the form:

$$f_k(\theta) = \frac{1}{k} \sum_{l=0}^{\infty} \sqrt{4\pi(2l+1)}\, Y_l^0(\theta) \frac{\eta_l - 1}{2i} \tag{5}$$

From this we deduce the differential elastic scattering cross section:

$$\sigma_{\text{el}}(\theta) = \frac{1}{k^2} \left| \sum_{l=0}^{\infty} \sqrt{4\pi(2l+1)}\, Y_l^0(\theta) \frac{\eta_l - 1}{2i} \right|^2 \tag{6}$$

and the total elastic scattering cross section:

$$\sigma_{\text{el}} = \frac{\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \left|1 - \eta_l\right|^2 \tag{7}$$

Comment:

According to the argument developed in § 1, the absorption of the wave (l) reaches a maximum when $|\eta_l|$ is zero, that is, when:

$$\eta_l = 0 \tag{8}$$

Formula (7) indicates however that, even in this limiting case, the contribution of the wave (l) to the elastic scattering cross section is not zero[2]. In other words, even if the interaction region is perfectly absorbing, it produces elastic scattering. This important phenomenon is a purely quantum effect. It can be compared to the behavior of a light wave which strikes an absorbing medium. Even if the absorption is total (perfectly black sphere or disc), a diffracted wave is observed (concentrated into a solid angle which becomes smaller as the surface of the disc becomes larger). Elastic scattering produced by a totally absorbing interaction is called, for this reason, *shadow scattering.*

2-b. Absorption cross section

Following the same principle as in § A-3 of Chapter VIII, we define the absorption cross section σ_{abs}: it is the ratio between the number of particles absorbed per unit time and the incident flux.

To calculate this cross section, it is sufficient, as in § B-2 of chapter VIII, to evaluate the total amount of probability $\Delta\mathcal{P}$ which "disappears" per unit time. This probability can be obtained from the current $\mathbf{J}$ associated with the wave function (4). $\Delta\mathcal{P}$ is equal to the difference between the flux of the incoming waves across a sphere (S) of very large radius R_0 and that of the outgoing waves; it is therefore equal to minus the net flux of the vector $\mathbf{J}$ leaving this sphere. Thus:

$$\Delta\mathcal{P} = -\int_{(S)} \mathbf{J}\cdot \mathrm{d}\mathbf{S} \tag{9}$$

with:

$$\mathbf{J} = \mathrm{Re}\left[v_k^{(\mathrm{scatt})*}(\mathbf{r})\frac{\hbar}{i\mu}\nabla v_k^{(\mathrm{scatt})}(\mathbf{r})\right] \tag{10}$$

Only the radial component J_r of the current contributes to the integral (9):

$$\Delta\mathcal{P} = -\int_{r=R_0} J_r\, r^2\, \mathrm{d}\Omega \tag{11}$$

with:

$$J_r = \mathrm{Re}\left[v_k^{(\mathrm{scatt})*}(\mathbf{r})\frac{\hbar}{i\mu}\frac{\partial}{\partial r}v_k^{(\mathrm{scatt})}(\mathbf{r})\right] \tag{12}$$

In formula (12), the derivative does not modify the angular dependence of the various terms which compose $v_k^{(\mathrm{scatt})}(\mathbf{r})$ [formula (4)]. Consequently, because of the

[2]This contribution is zero only if $\eta_l = 1$, that is, if the phase shift is real and equal to an integral multiple of π [this was already contained in formula (C-58) of Chapter VIII].

orthogonality of the spherical harmonics, the cross terms between a partial wave (l) in $v_k^{(\text{scatt})}(\mathbf{r})$ and a different wave (l') in $v_k^{(\text{scatt})*}(\mathbf{r})$ make a zero contribution to integral (11). We have therefore:

$$\Delta\mathcal{P} = -\sum_{l=0}^{\infty} \int_{r=R_0} J_r^{(l)}\, r^2\, \mathrm{d}\Omega \tag{13}$$

where $J_r^{(l)}$ is the radial component of the current associated with the partial wave (l). A simple calculation gives:

$$J_r^{(l)} \underset{r\to\infty}{\sim} -\frac{\hbar k}{\mu}\,\frac{\pi(2l+1)}{k^2 r^2}\left[1-|\eta_l|^2\right]\left|Y_l^0(\theta)\right|^2 \tag{14}$$

that is, finally, since $Y_l^0(\theta)$ is normalized:

$$\Delta\mathcal{P} = \frac{\hbar k}{\mu}\,\frac{\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\left[1-|\eta_l|^2\right] \tag{15}$$

The absorption cross section σ_{abs} is therefore equal to the probability $\Delta\mathcal{P}$ divided by the incident current $\hbar k/\mu$:

$$\sigma_{\text{abs}} = \frac{\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\left[1-|\eta_l|^2\right] \tag{16}$$

It is obvious that σ_{abs} is zero if all the η_l have a modulus of 1; that is, according to (2), if all the phase shifts are real. In this case, there is only elastic scattering, and the net flux of probability leaving a sphere of large radius R_0 is always zero. The total probability carried by the incoming waves is entirely transferred to the outgoing waves. On the other hand, when η_l is zero, the contribution of the wave (l) to the absorption cross section is maximum.

Comment:

The calculation of expression (15) shows that $\dfrac{\hbar k}{\mu}\,\dfrac{\pi}{k^2}(2l+1)$ is the amount of probability entering per unit time, and arising from the partial wave (l). If we divide this quantity by the incident current $\hbar k/\mu$, we obtain a surface that can be called the "incoming cross section into the partial wave (l)":

$$\sigma_l = \frac{\pi}{k^2}(2l+1) \tag{17}$$

This formula can be interpreted classically. We can consider the incident plane wave as describing a beam of particles of uniform density, having a momentum $\hbar k$ parallel to Oz. What proportion of these particles reach the scattering potential, with an angular momentum $\hbar\sqrt{l(l+1)}$? We have already mentioned the link between angular momentum and the impact parameter in classical mechanics [*cf.* formula (C-23) of Chapter VIII]:

$$\mathcal{L} = b|\mathbf{p}| = \hbar k\, b \tag{18}$$

All we must do, therefore, is to draw, in the plane passing through O and perpendicular to Oz, a circular ring centered at O, of average radius b_l such that:

$$\hbar\sqrt{l(l+1)} = \hbar k\, b_l \tag{19}$$

and of width Δb_l corresponding to $\Delta l = 1$ in formula (19) (Fig. 1). All the particles crossing this surface reach the scattering potential with an angular momentum equal to $\hbar\sqrt{l(l+1)}$, to within $\hbar$. From (19) we derive:

$$b_l = \frac{1}{k}\sqrt{l(l+1)} \simeq \frac{1}{k}\left(l + \frac{1}{2}\right) \tag{20}$$

if $l \gg 1$, and consequently:

$$\Delta b_l = \frac{1}{k} \tag{21}$$

The area of the circular ring of Figure 1 is therefore:

$$2\pi\, b_l\, \Delta b_l \simeq \frac{\pi}{k^2}(2l+1) \tag{22}$$

Thus we find again, very simply, σ_l.

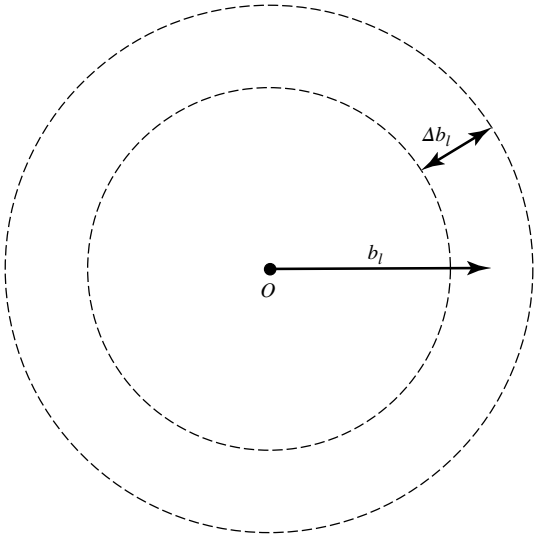


Figure 1: The incident particles must reach the potential with the impact parameter b_l to within Δb_l for their classical angular momentum to be $\hbar\sqrt{l(l+1)}$ to within $\hbar$.

2-c. Total cross section. Optical theorem

When a collision can give rise to several different reactions or scattering phenomena, the total cross section σ_{tot} is defined as the sum of the cross sections (integrated over all the directions of space) corresponding to all these processes. The total cross section is thus the number of particles which, per unit time, participate in one or another of the possible reactions, divided by the incident flux.

If, as above, we treat globally all reactions other than elastic scattering, we have simply:

$$\sigma_{\text{tot}} = \sigma_{\text{el}} + \sigma_{\text{abs}} \tag{23}$$

Formulas (7) and (16) then give:

$$\sigma_{\text{tot}} = \frac{2\pi}{k^2} \sum_{l=0}^{\infty} (2l+1)(1 - \text{Re}\,\eta_l) \tag{24}$$

Now $(1 - \text{Re}\,\eta_l)$ is the real part of $(1 - \eta_l)$, which appears in the elastic scattering amplitude [formula (5)]. Moreover, we know the value of $Y_l^0(\theta)$ for $\theta = 0$:

$$Y_l^0(0) = \sqrt{\frac{2l+1}{4\pi}} \tag{25}$$

[*cf.* complement A_{VI}, formulas (57) and (60)]. Consequently, if we calculate from (5) the imaginary part of the elastic scattering amplitude in the forward direction, we find:

$$\text{Im}\, f_k(0) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1)\, \frac{1 - \text{Re}\,\eta_l}{2} \tag{26}$$

Comparing this expression to formula (24), we see that:

$$\sigma_{\text{tot}} = \frac{4\pi}{k}\, \text{Im}\, f_k(0) \tag{27}$$

This relation between the total cross section and the imaginary part of the elastic scattering amplitude in the forward direction is valid in a very general sense; it constitutes what is called the *optical theorem.*

Comment:

The optical theorem is obviously valid in the case of purely elastic scattering ($\sigma_{\text{abs}} = 0; \sigma_{\text{tot}} = \sigma_{\text{el}}$). The fact that $f_k(0)$ – i.e. the wave scattered in the forward direction – is related to the total cross section could have been predicted from the discussion in § B-2-d of Chapter VIII. It is the interference in the forward direction between the incident plane wave and the scattered wave that accounts for the attenuation of the transmitted beam, due to the scattering of particles in all directions of space.

References and suggestions for further reading:

Optical model: Valentin (16.1), § X-3. High energy proton-proton collisions: Amaldi (16.31).

Complement C_{VIII}

Some simple applications of scattering theory

1 The Born approximation for a Yukawa potential 977
1-a Calculation of the scattering amplitude and cross section . . 977
1-b The infinite-range limit 979
2 Low energy scattering by a hard sphere 980
3 Exercises 981
3-a Scattering of the p wave by a hard sphere 981
3-b "Square spherical well": bound states and scattering resonances 982

There is no potential for which the scattering problem can be solved exactly[1] by a simple analytical calculation. Therefore, in the examples that we are going to discuss, we shall content ourselves with using the approximations that we introduced in Chapter VIII.

1. The Born approximation for a Yukawa potential

Let us consider a potential of the form:

$$V(\mathbf{r}) = V_0 \frac{\mathrm{e}^{-\alpha r}}{r} \tag{1}$$

where V_0 and α are real constants, with α positive. This potential is attractive or repulsive depending on whether V_0 is negative or positive. The larger $|V_0|$, the more intense the potential. Its range is characterized by the distance:

$$r_0 = \frac{1}{\alpha} \tag{2}$$

since, as Figure 1 shows, $V(r)$ is practically zero when r exceeds $2r_0$ or $3r_0$.

The potential (1) bears the name of Yukawa, who had the idea of associating it with nuclear forces, whose range is of the order of a fermi. To explain the origin of this potential, Yukawa was led to predict the existence of the π-meson, which was indeed later discovered. Notice that for $\alpha = 0$ this potential becomes the Coulomb potential, which can thus be considered to be a Yukawa potential of infinite range.

1-a. Calculation of the scattering amplitude and cross section

We assume that $|V_0|$ is sufficiently small for the Born approximation (§ B-4 of Chapter VIII) to be valid. According to formula (B-47) of Chapter VIII, the scattering amplitude $f_k^{(B)}(\theta, \varphi)$ is then given by:

$$f_k^{(B)}(\theta, \varphi) = -\frac{1}{4\pi} \frac{2\mu V_0}{\hbar^2} \int \mathrm{d}^3 r \, \mathrm{e}^{-i\mathbf{K}\cdot\mathbf{r}} \frac{\mathrm{e}^{-i\alpha r}}{r} \tag{3}$$

[1] Actually, we can rigorously treat the case of the Coulomb potential; however, as we pointed out in Chapter VIII (§ B-1), this necessitates a special method.

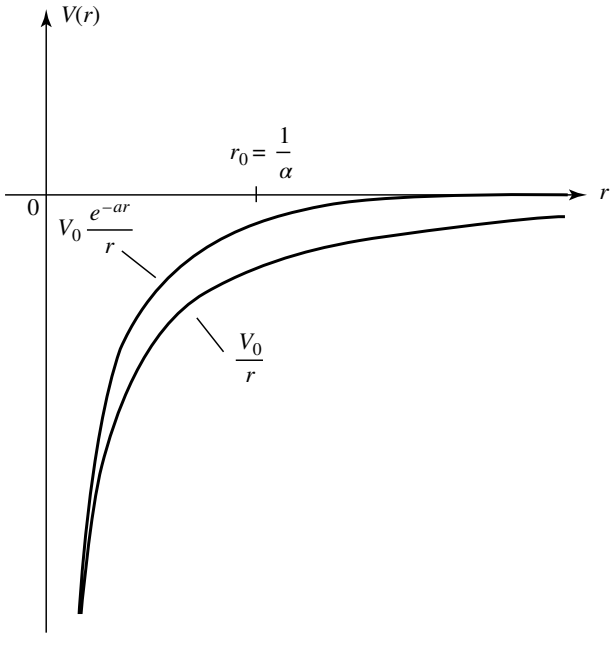


Figure 1: Yukawa potential and Coulomb potential. The presence of the term $e^{-\alpha r}$ *causes the Yukawa potential to approach zero much more rapidly when* $r \gg r_0 = 1/\alpha$ *(range of the potential).*

where $\mathbf{K}$ is the momentum transferred in the direction (θ, φ) defined by relation (B-42) of Chapter VIII.

Expression (3) involves the Fourier transform of the Yukawa potential. Since this potential depends only on the variable r, the angular integrations can easily be carried out (§ 2-e of Appendix I), putting the scattering amplitude into the form:

$$f_k^{(B)}(\theta, \varphi) = -\frac{1}{4\pi} \frac{2\mu V_0}{\hbar^2} \frac{4\pi}{|\mathbf{K}|} \int_0^\infty r dr \sin |\mathbf{K}| r \frac{e^{-\alpha r}}{r} \tag{4}$$

After a simple calculation, we then find:

$$f_k^{(B)}(\theta, \varphi) = -\frac{2\mu V_0}{\hbar^2} \frac{1}{\alpha^2 + |\mathbf{K}|^2} \tag{5}$$

Figure 6 of Chapter VIII shows that:

$$|\mathbf{K}| = 2k \sin \frac{\theta}{2} \tag{6}$$

where k is the modulus of the incident wave vector and θ is the scattering angle.

The differential scattering cross section is therefore written, in the Born approximation:

$$\sigma^{(B)}(\theta) = \frac{4\mu^2 V_0^2}{\hbar^4} \frac{1}{\left[\alpha^2 + 4k^2 \sin^2 \theta/2\right]^2} \tag{7}$$

It is independent of the azimuthal angle φ, as could have been foreseen from the fact that the problem of scattering by a central potential is symmetrical with respect to rotation about the direction of the incident beam. On the other hand, it depends, for a given energy (that is, for fixed k), on the scattering angle: in particular, the cross section in the forward direction ($\theta = 0$) is larger than the cross section in the backward direction ($\theta = \pi$). Finally, $\sigma^{(B)}(\theta)$, for fixed θ, is a decreasing function of the energy. Notice, moreover, that the sign of V_0 is of no importance in the scattering problem, at least in the Born approximation.

The total scattering cross section is easily obtained by integration:

$$\sigma^{(B)} = \int \mathrm{d}\Omega \; \sigma^{(B)}(\theta) = \frac{4\mu^2 V_0^2}{\hbar^4} \frac{4\pi}{\alpha^2(\alpha^2 + 4k^2)} \tag{8}$$

1-b. The infinite-range limit

We noted above that the Yukawa potential approaches a Coulomb potential when α tends towards zero. What happens, in this limiting case, to the formulas that we have just established?

To obtain the Coulomb interaction potential between two particles having charges of $Z_1 q$ and $Z_2 q$ (q being the charge of the electron), we write:

$$\begin{aligned} &\alpha = 0 \\ &V_0 = Z_1 Z_2 \, e^2 \end{aligned} \tag{9}$$

with:

$$e^2 = \frac{q^2}{4\pi\varepsilon_0} \tag{10}$$

Formula (7) then gives:

$$\begin{aligned} \sigma^{(C)}(\theta) &= \frac{4\mu^2}{\hbar^4} \frac{Z_1^2 Z_2^2 \, e^4}{16 \; k^4 \sin^4 \dfrac{\theta}{2}} \\ &= \frac{Z_1^2 Z_2^2 \, e^4}{16 \; E^2 \sin^4 \dfrac{\theta}{2}} \end{aligned} \tag{11}$$

(k has been replaced by its value in terms of the energy).

Expression (11) is indeed that of the Coulomb scattering cross section (*Rutherford's formula*). Of course, the way in which we have obtained it does not constitute a proof: the theory we have used is not applicable to the Coulomb potential. However, it is interesting to observe that the Born approximation for the Yukawa potential gives precisely Rutherford's formula for the limiting situation where the range of the potential approaches infinity.

Comment:

The total scattering cross section for a Coulomb potential is infinite since the corresponding integral diverges for small values of θ [expression (8) becomes infinite

when α approaches zero]. This results from the infinite range of the Coulomb potential: even if the particle passes very far from the point O, it is affected by the potential. This suggests why the scattering cross section should be infinite. However, in reality, one never observes a rigorously pure Coulomb interaction over an infinite range. The potential created by a charged particle is always modified by the presence, in its more or less immediate neighborhood, of other particles of opposite charge (screening effect).

2. Low energy scattering by a hard sphere

Let us consider a central potential such that:

$$\begin{aligned} V(r) &= 0 \quad \text{for} \quad r > r_0 \\ &= \infty \quad \text{for} \quad r < r_0 \end{aligned} \tag{12}$$

In this case, we say that we are considering a "hard sphere" of radius r_0. We assume that the energy of the incident particle is sufficiently small for kr_0 to be much smaller than 1. We can then (§ C-3-b-β of chapter VIII and exercise 3-a below) neglect all the phase shifts except that of the s wave ($l = 0$). The scattering amplitude $f_k(\theta)$ is written, under these conditions:

$$f_k(\theta) = \frac{1}{k} e^{i\delta_0(k)} \sin \delta_0(k) \tag{13}$$

(since $Y_0^0 = 1/\sqrt{4\pi}$). The differential cross section is isotropic:

$$\sigma(\theta) = |f_k(\theta)|^2 = \frac{1}{k^2} \sin^2 \delta_0(k) \tag{14}$$

so that the total cross section is simply equal to:

$$\sigma = \frac{4\pi}{k^2} \sin^2 \delta_0(k) \tag{15}$$

To calculate the phase shift $\delta_0(k)$, it is necessary to solve the radial equation corresponding to $l = 0$. This equation is written [*cf.* formula (C-35) of Chapter VIII]:

$$\left[\frac{d^2}{dr^2} + k^2 \right] u_{k,0}(r) = 0 \quad \text{for} \quad r > r_0 \tag{16}$$

which must be completed by the condition:

$$u_{k,0}(r_0) = 0 \tag{17}$$

since the potential becomes infinite for $r = r_0$. The solution $u_{k,0}(r)$ of equations (16) and (17) is unique to within a constant factor:

$$\begin{aligned} u_{k,0}(r) &= C \sin k(r - r_0) \quad \text{for} \quad r > r_0 \\ &= 0 \qquad\qquad\quad \text{for} \quad r < r_0 \end{aligned} \tag{18}$$

The phase shift δ_0 is, by definition, given by the asymptotic form of $u_{k,0}(r)$:

$$u_{k,0}(r) \underset{r \to \infty}{\sim} \sin(kr + \delta_0) \tag{19}$$

Thus, using solution (18), we find:

$$\delta_0(k) = -kr_0 \tag{20}$$

If we insert this value into expression (15) for the total cross section, we obtain:

$$\sigma = \frac{4\pi}{k^2} \sin^2 kr_0 \simeq 4\pi r_0^2 \tag{21}$$

since by hypothesis kr_0 is much smaller than 1. Therefore, σ is independent of the energy and equal to four times the apparent surface of the hard sphere seen by the particles of the incident beam. A calculation based on classical mechanics would give for the cross section the apparent surface πr_0^2: only the particles which bounce elastically off the hard sphere would be deflected. In quantum mechanics, however, one studies the evolution of the wave associated with the incident particles, and the abrupt variation of $V(r)$ at $r = r_0$ produces a phenomenon analogous to the diffraction of a light wave.

Comment:

Even when the wavelength of the incident particles becomes negligible compared to r_0 ($kr_0 \gg 1$), the quantum cross section does not approach πr_0^2. It is possible, for very large k, to sum the series which gives the total cross section in terms of phase shifts [formula (C-58) of Chapter VIII]; we then find:

$$\sigma \underset{k\to\infty}{\sim} 2\pi r_0^2 \tag{22}$$

Wave effects thus persist in the limiting case of very small wavelengths. This is due to the fact that the potential under study is discontinuous at $r = r_0$: it always varies appreciably within an interval which is smaller than the wavelength of the particles (*cf.* Chapter I, § D-2-a).

3. Exercises

3-a. Scattering of the *p* wave by a hard sphere

We wish to study the phase shift $\delta_1(k)$ produced by a hard sphere on the p wave ($l = 1$). In particular, we want to verify that it becomes negligible compared to $\delta_0(k)$ at low energy.

α. Write the radial equation for the function $u_{k,1}(r)$ for $r > r_0$. Show that its general solution is of the form:

$$u_{k,1}(r) = C\left[\frac{\sin kr}{kr} - \cos kr + a\left(\frac{\cos kr}{kr} + \sin kr\right)\right]$$

where C and a are constants.

β. Show that the definition of $\delta_1(k)$ implies that:

$$a = \tan \delta_1(k)$$

γ. Determine the constant a from the condition imposed on $u_{k,1}(r)$ at $r = r_0$.

δ. Show that, as k approaches zero, $\delta_1(k)$ behaves like[2] $(kr_0)^3$, which makes it negligible compared to $\delta_0(k)$.

3-b. "Square spherical well": bound states and scattering resonances

Consider a central potential $V(r)$ such that:

$$\begin{aligned} V(r) &= -V_0 \quad \text{for} \quad r < r_0 \\ &= 0 \qquad \text{for} \quad r > r_0 \end{aligned}$$

where V_0 is a positive constant. Set:

$$k_0 = \sqrt{\frac{2\mu V_0}{\hbar^2}}$$

We shall confine ourselves to the study of the s wave ($l = 0$).

α. *Bound states ($E < 0$)*

(*i*) Write the radial equation in the two regions $r > r_0$ and $r < r_0$, as well as the condition at the origin. Show that, if one sets:

$$\rho = \sqrt{\frac{-2\mu E}{\hbar^2}}$$

$$K = \sqrt{k_0^2 - \rho^2}$$

the function $u_0(r)$ is necessarily of the form:

$$\begin{aligned} u_0(r) &= A\,\mathrm{e}^{-\rho r} \qquad \text{for} \quad r > r_0 \\ &= B \sin Kr \quad \text{for} \quad r < r_0 \end{aligned}$$

(*ii*) Write the matching conditions at $r = r_0$. Deduce from them that the only possible values for ρ are those which satisfy the equation:

$$\tan Kr_0 = -\frac{K}{\rho}$$

(*iii*) Discuss this equation: indicate the number of s bound states as a function of the depth of the well (for fixed r_0) and show, in particular, that there are no bound states if this depth is too small.

[2] This result is true in general: for any potential of finite range r_0, the phase shift $\delta_l(k)$ behaves like $(kr_0)^{2l+1}$ at low energies.

β. Scattering resonances ($E > 0$)

(*i*) Again write the radial equation, this time setting:

$$k = \sqrt{\frac{2\mu E}{\hbar^2}}$$
$$K' = \sqrt{k_0^2 + k^2}$$

Show that $u_{k,0}(r)$ is of the form:

$$u_{k,0}(r) = A \sin(kr + \delta_0) \quad \text{for} \quad r > r_0$$
$$= B \sin K'r \quad \text{for} \quad r < r_0$$

(*ii*) Choose $A = 1$. Show, using the continuity conditions at $r = r_0$, that the constant B and the phase shift δ_0 are given by:

$$B^2 = \frac{k^2}{k^2 + k_0^2 \cos^2 K'r_0}$$
$$\delta_0 = -kr_0 + \alpha(k)$$

with:

$$\tan \alpha(k) = \frac{k}{K'} \tan K'r_0$$

(*iii*) Trace the curve representing B^2 as a function of k. This curve clearly shows resonances, for which B^2 is maximum. What are the values of k associated with these resonances? What is then the value of $\alpha(k)$? Show that, if there exists such a resonance for a small energy ($kr_0 \ll 1$), the corresponding contribution of the s wave to the total cross section is practically maximal.

γ. Relation between bound states and scattering resonances

Assume that $k_0 r_0$ is very close to $(2n+1)\frac{\pi}{2}$, where n is an integer, and set:

$$k_0 r_0 = (2n+1)\frac{\pi}{2} + \varepsilon \quad \text{with} \quad \varepsilon \ll 1$$

(*i*) Show that, if ε is positive, there exists a bound state whose binding energy $E = -\hbar^2\rho^2/2\mu$ is given by:

$$\rho \simeq \varepsilon k_0$$

(*ii*) Show that if, on the other hand, ε is negative, there exists a scattering resonance at energy $E = \hbar^2 k^2/2\mu$ such that:

$$k^2 \simeq -\frac{2k_0\varepsilon}{r_0}$$

(*iii*) Deduce from this that if the depth of the well is gradually decreased (for fixed r_0), the bound state which disappears when $k_0 r_0$ passes through an odd multiple of $\pi/2$ gives rise to a low energy scattering resonance.

References and suggestions for further reading:

Messiah (1.17), Chap. IX, § 10 and Chap. X, §§ III and IV; Valentin (16.1), Annexe II.

Chapter IX

Electron spin

A Introduction of electron spin 986
A-1 Experimental evidence 986
A-2 Quantum description: postulates of the Pauli theory 988
B Special properties of an angular momentum 1/2 990
C Non-relativistic description of a spin 1/2 particle 991
C-1 Observables and state vectors 992
C-2 Probability calculations for a physical measurement 997

Until now, we have considered the electron to be a point particle possessing three degrees of freedom associated with its three coordinates x, y and z. Consequently, the quantum theory that we have developed is based on the hypothesis that an electron state, at a given time, is characterized by a wave function $\psi(x, y, z)$ which depends only on x, y and z. Within this framework, we have studied a certain number of physical systems: amongst others, the hydrogen atom (in Chapter VII), which is particularly interesting because of the very precise experiments that can be performed on it. The results obtained in Chapter VII actually describe the emission and absorption spectra of hydrogen very accurately. They give the energy levels correctly and make it possible to explain, using the corresponding wave functions, the selection rules (which indicate which frequencies, out of all the Bohr frequencies that are *a priori* possible, appear in the spectrum). Atoms with many electrons can be treated in an analogous fashion (by using approximations, however, since the complexity of the Schrödinger equation, even for the helium atom with two electrons, makes an exact analytic solution of the problem impossible). In this case as well, agreement between theory and experiment is satisfying.

However, when atomic spectra are studied in detail, certain phenomena appear, as we shall see, which cannot be interpreted within the framework of the theory that we have developed. This result is not surprising. It is clear that it is necessary to complete the preceding theory by a certain number of *relativistic corrections*: one must take into account the modifications brought in by *relativistic kinemalics* (variation of

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

mass with velocity, etc.) and *magnetic effects* which we have neglected. We know that these corrections are small (§ C-4-a of Chapter VII): nevertheless, they do exist, and can be measured.

The *Dirac equation* gives a relativistic quantum mechanical description of the electron. Compared to the Schrödinger equation, it implies a profound modification in the quantum description of the properties of the electron; in addition to the corrections already pointed out concerning its position variables, a new characteristic of the electron appears: its *spin.* In a more general context, the structure of the Lorentz group (group of relativistic space-time transformations) reveals spin to be an intrinsic property of various particles, on the same footing, for example, as their rest mass[1].

Historically, electron spin was discovered experimentally before the introduction of the Dirac equation. Furthermore, Pauli developed a theory which allowed spin to be incorporated simply into non-relativistic quantum mechanics[2] through the addition of several supplementary postulates. Theoretical predictions for the atomic spectra are then obtained which are in excellent agreement with experimental results[3].

It is Pauli's theory, which is much simpler than Dirac's, that we are going to develop in this chapter. We shall begin, in § A, by describing a certain number of experimental results, which revealed the existence of electron spin. Then we shall specify the postulates on which Pauli's theory is based. Afterwards, we shall examine, in § B, the special properties of an angular momentum 1/2. Finally, we shall show, in § C, how one can take into account simultaneously the position variables and the spin of a particle such as the electron.

A. Introduction of electron spin

A-1. Experimental evidence

Experimental demonstrations of the existence of electron spin are numerous and appear in various important physical phenomena. For example, the magnetic properties of numerous substances, particularly of ferromagnetic metals, can only be explained if spin is taken into account. Here, however, we are going to confine ourselves to a certain number of simple phenomena observed experimentally in atomic physics: the fine structure of spectral lines, the Zeeman effect and, finally, the behavior of silver atoms in the Stern-Gerlach experiment.

A-1-a. Fine structure of spectral lines

The precise experimental study of atomic spectral lines (for the hydrogen atom, for example) reveals a *fine structure*: each line is in fact made up of several components having nearly identical frequencies[4] but which can be clearly distinguished by a device

[1] This does not mean that spin has a purely relativistic origin: it can be deduced from the structure of the non-relativistic transformation group (the Galilean group).

[2] Pauli's theory can be obtained as a limiting case of Dirac's theory when the electron's speed is small compared to that of light.

[3] We shall see, for example in Chapter XII where the general perturbation theory treated in Chapter XI is used, how relativistic corrections and the existence of spin enable us to account quantitatively for the details of the hydrogen atomic spectrum (which would be inexplicable if we limited ourselves to the theory of Chapter VII).

[4] For example, the resonance line of the hydrogen atom ($2p \leftrightarrow 1s$ transition) is actually double: the two components are separated by an interval of the order of 10^{-4} eV (that is, about 10^5 times smaller

with good resolution. This means that there exist groups of atomic levels which are very closely spaced but distinct. In particular, the calculations of § C of Chapter VII give the average energies of different groups of levels for the hydrogen atom but do not explain the splittings within each group.

A-1-b. "Anomalous" Zeeman effect

When an atom is placed in a uniform magnetic field, each of its lines (that is, each component of the fine structure) splits into a certain number of equidistant lines, the interval being proportional to the magnetic field: this is the *Zeeman effect*. The origin of the Zeeman effect can be easily understood by using the results of Chapters VI and VII (complement D_{VII}). The theoretical explanation is based on the fact that a magnetic moment $\mathbf{M}$ is associated with the orbital angular momentum $\mathbf{L}$ of an electron:

$$\mathbf{M} = \frac{\mu_B}{\hbar}\mathbf{L} \tag{A-1}$$

where μ_B is the "Bohr magneton":

$$\mu_B = \frac{q\hbar}{2m_e} \tag{A-2}$$

However, while this theory is confirmed by experiment in certain cases (the so-called "normal" Zeeman effect), it is, in other cases, incapable of accounting quantitatively for the observed phenomena (the so-called "anomalous" Zeeman effect). The most striking "anomaly" appears for atoms with odd atomic number Z (in particular, for the hydrogen atom): their levels are divided into an *even number of Zeeman sub-levels,* while, according to the theory, this number should always be odd, being equal to $(2l+1)$ with l an integer.

A-1-c. Existence of half-integral angular momenta

We are confronted with the same difficulty in connection with the Stern-Gerlach experiment, which we described in Chapter IV (§ A-1); the beam of silver atoms is split symmetrically in two. These results suggest that *half-integral values of j* (which we saw in § C-2 of Chapter VI to be *a priori* possible) do indeed exist. But this poses a serious problem, since we showed in § D-1-b of Chapter VI that the orbital angular momentum of a particle such as an electron could only be integral (more precisely, it is the quantum number l which is integral). Even in atoms with several electrons, each of these has an integral orbital angular momentum, and we shall show in Chapter X that, under these conditions, the total orbital angular momentum of the atom is necessarily integral. The existence of half-integral angular momenta thus cannot be explained without supplementary hypotheses.

> ***Comment:***
>
> It is not possible to measure directly the angular momentum of the electron using the Stern-Gerlach apparatus. Unlike silver atoms, electrons possess an electric charge q, and the force due to the interaction between their magnetic moment and the inhomogeneous magnetic field would be completely masked by the Lorentz force $q\mathbf{v} \times \mathbf{B}$.

than the average $2p \leftrightarrow 1s$ transition energy, which is equal to 10.2 eV).

A-2. Quantum description: postulates of the Pauli theory

indexSpin!quantum description

In order to resolve the preceding difficulties, Uhlenbeck and Goudsmit (1925) proposed the following hypothesis: the electron "spins" and this gives it an intrinsic angular momentum which is called the spin. To interpret the experimental results described above, one must also assume that a magnetic moment $\mathbf{M}_S$ is associated[5] with this angular momentum $\mathbf{S}$:

$$\mathbf{M}_S = 2\frac{\mu_B}{\hbar}\mathbf{S} \tag{A-3}$$

Note that the coefficient of proportionality between the angular momentum and the magnetic moment is twice as large in (A-3) as in (A-1): one says that *the spin gyromagnetic ratio is twice the orbital gyromagnetic ratio.*

Pauli later stated this hypothesis more precisely and gave a quantum description of spin which is valid in the non-relativistic limit. To the general postulates of quantum mechanics that we set forth in Chapter III must be added a certain number of postulates relating to spin.

Until now, we have studied the quantization of *orbital variables.* With the position $\mathbf{r}$ and the momentum $\mathbf{p}$ of a particle such as the electron, we associated the observables $\mathbf{R}$ and $\mathbf{P}$ acting in the state space $\mathcal{E}_\mathbf{r}$, which is isomorphic to the space $\mathcal{F}$ of wave functions. All physical quantities are functions of the fundamental variables $\mathbf{r}$ and $\mathbf{p}$, and the quantization rules enable us to associate with them observables acting in $\mathcal{E}_\mathbf{r}$. We shall call $\mathcal{E}_\mathbf{r}$ the *orbital state space.*

To these orbital variables we shall add *spin variables* which satisfy the following postulates:

(*i*) *The spin operator* $\mathbf{S}$ *is an angular momentum.* This means (§ B-2 of Chapter VI) that its three components are observables which satisfy the commutation relations:

$$[S_x,\, S_y] = i\hbar S_z \tag{A-4}$$

and the two formulas which are deduced by cyclic permutation of the indices x, y, z.

(*ii*) The spin operators act in a new space, the "*spin state space*" $\mathcal{E}_s$, where $\mathbf{S}^2$ *and* S_z *constitute a C.S.C.O.* The space $\mathcal{E}_s$ is thus spanned by the set of eigenstates $|s, m\rangle$ common to $\mathbf{S}^2$ and S_z:

$$\mathbf{S}^2\,|s, m\rangle = s(s+1)\hbar^2\,|s, m\rangle \tag{A-5a}$$

$$S_z\,|s, m\rangle = m\hbar\,|s, m\rangle \tag{A-5b}$$

According to the general theory of angular momentum (§ C of Chapter VI), we know that s must be an integral or half-integral number, and that m takes on all

[5] Actually, when one takes into account the coupling of the electron with the quantized electromagnetic field (quantum electrodynamics), one finds that the coefficient of proportionality between $\mathbf{M}_S$ and $\mathbf{S}$ is not exactly $2\mu_B/\hbar$. The difference, which is of the order of 10^{-3} in relative value, is easily observable experimentally; it is often called the "anomalous magnetic moment" of the electron.

values included between $-s$ and $+s$ differing from these two numbers by an integer (which may be zero). A given particle is characterized by *a unique value of s*: this particle is said to have a spin s. The spin state space $\mathcal{E}_s$ is therefore always of finite dimension $(2s+1)$, and all spin states are eigenvectors of $\mathbf{S}^2$ with the same eigenvalue $s(s+1)\hbar^2$.

(*iii*) *The state space* $\mathcal{E}$ of the particle being considered *is the tensor product of* $\mathcal{E}_\mathbf{r}$ *and* $\mathcal{E}_s$:

$$\mathcal{E} = \mathcal{E}_\mathbf{r} \otimes \mathcal{E}_s \tag{A-6}$$

Consequently (§ F of Chapter II), *all spin observables commute with all orbital observables.*

Except for the particular case where $s = 0$, it is therefore not sufficient to specify a ket of $\mathcal{E}_\mathbf{r}$ (that is, a square-integrable wave function) to characterize a state of the particle. In other words, the observables X, Y and Z do not constitute a C.S.C.O. in the space state $\mathcal{E}$ of the particle (no more than do P_x, P_y, P_z or any other C.S.C.O. of $\mathcal{E}_\mathbf{r}$). It is also necessary to know the spin state of the particle, that is, to add to the C.S.C.O. of $\mathcal{E}_\mathbf{r}$ a C.S.C.O. of $\mathcal{E}_s$ composed of spin observables, for example $\mathbf{S}^2$ and S_z (or $\mathbf{S}^2$ and S_x). Every particle state is a linear combination of vectors which are tensor products of a ket of $\mathcal{E}_\mathbf{r}$ and a ket of $\mathcal{E}_s$ (see § C below).

(*iv*) *The electron is a spin 1/2 particle* ($s = 1/2$) and its *intrinsic magnetic moment* is given by formula (A-3). For the electron, the space $\mathcal{E}_s$ is therefore two-dimensional.

Comments:

(*i*) The proton and the neutron, which are nuclear constituents, are also spin 1/2 particles, but their gyromagnetic ratios are different from that of the electron. At the present time, we know of the existence of particles of spin 0, 1/2, 1, 3/2, 2, etc.

(*ii*) In order to explain the existence of spin, we could imagine that a particle like the electron, instead of being a point, has a certain spatial extension. It would then be the rotation of the electron about its axis that would give rise to an intrinsic angular momentum. However, it is important to note that, in order to describe a structure that is more complex than a material point, it would be necessary to introduce more than three position variables. If, for example, the electron behaved like a solid body, six variables would be required: three coordinates to locate one of its points chosen once and for all, such as its center of gravity, and three angles to specify its orientation in space. The theory that we are considering here is radically different. It continues to treat the electron like a point (its position is fixed by three coordinates). The spin angular momentum is not derived from any position or momentum variable[6]. *Spin thus has no classical analogue.*

[6] If it were, moreover, it would necessarily be integral.

B. Special properties of an angular momentum 1/2

We shall restrict ourselves from now on to the case of the electron, which is a spin 1/2 particle. From the preceding chapters, we know how to handle its orbital variables. We are now going to study in more detail its spin degrees of freedom.

The spin state space $\mathcal{E}_s$ is two-dimensional. We shall take as a basis the orthonormal system $\{|+\rangle, |-\rangle\}$ of eigenkets common to $\mathbf{S}^2$ and S_z which satisfy the equations:

$$\begin{cases} \mathbf{S}^2 |\pm\rangle = \dfrac{3}{4}\hbar^2 |\pm\rangle & \text{(B-1a)} \\ S_z |\pm\rangle = \pm\dfrac{1}{2}\hbar |\pm\rangle & \text{(B-1b)} \end{cases}$$

$$\begin{cases} \langle + \mid - \rangle = 0 & \text{(B-2a)} \\ \langle + \mid + \rangle = \langle - \mid - \rangle = 1 & \text{(B-2b)} \end{cases}$$

$$|+\rangle\langle+| \quad + \quad |-\rangle\langle-| = \mathbb{1} \tag{B-3}$$

where $\mathbb{1}$ is the unit operator. The most general spin state is described by an arbitrary vector of $\mathcal{E}_s$:

$$|\chi\rangle = c_+ |+\rangle + c_- |-\rangle \tag{B-4}$$

where c_+ and c_- are complex numbers. According to (B-1a), all the kets of $\mathcal{E}_s$ are eigenvectors of $\mathbf{S}^2$ with the same eigenvalue $3\hbar^2/4$, which causes $\mathbf{S}^2$ to be proportional to the identity operator of $\mathcal{E}_s$:

$$\mathbf{S}^2 = \frac{3}{4}\hbar^2 \tag{B-5}$$

(in the right hand side of this equation, as is usually done, we have not written the unit operator $\mathbb{1}$ explicitly). Since $\mathbf{S}$ is, by definition, an angular momentum, it possesses all the general properties derived in § C of Chapter VI. The action of the operators:

$$S_\pm = S_x \pm iS_y \tag{B-6}$$

on the basis vectors $|+\rangle$ and $|-\rangle$ is given by the general formulas (C-50) of Chapter VI when one sets $j = s = 1/2$:

$$S_+ |+\rangle = 0 \qquad S_+ |-\rangle = \hbar |+\rangle \tag{B-7a}$$
$$S_- |+\rangle = \hbar |-\rangle \qquad S_- |-\rangle = 0 \tag{B-7b}$$

Any operator acting in $\mathcal{E}_s$ can be represented, in the $\{|+\rangle, |-\rangle\}$ basis, by a 2×2 matrix. In particular, using (B-1b) and (B-7), we find the matrices corresponding to S_x, S_y and S_z in the form:

$$(\mathbf{S}) = \frac{\hbar}{2}\,\boldsymbol{\sigma} \tag{B-8}$$

where $\boldsymbol{\sigma}$ designates the set of the three *Pauli matrices*:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{B-9}$$

The Pauli matrices possess the following properties, which can easily be verified from their explicit form (B-9) (see also Complement $\mathrm{A_{IV}}$):

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbb{1} \tag{B-10a}$$

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0 \tag{B-10b}$$

$$[\sigma_x, \sigma_y] = 2i\sigma_z \tag{B-10c}$$

$$\sigma_x\sigma_y = i\sigma_z \tag{B-10d}$$

(to the last three formulas must be added those obtained through cyclic permutation of the x, y, z indices). It also follows from (B-9) that:

$$\mathrm{Tr}\,\sigma_x = \mathrm{Tr}\,\sigma_y = \mathrm{Tr}\,\sigma_z = 0 \tag{B-11a}$$

$$\mathrm{Det}\,\sigma_x = \mathrm{Det}\,\sigma_y = \mathrm{Det}\,\sigma_z = -1 \tag{B-11b}$$

Furthermore, any 2×2 matrix can be written as a linear combination, with complex coefficients, of the three Pauli matrices and the unit matrix. This is simply due to the fact that a 2×2 matrix has only four elements. Finally, it is easy to derive (see Complement $\mathrm{A_{IV}}$) the following identity:

$$(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B}) = (\mathbf{A} \cdot \mathbf{B})\ \mathbb{1} + i\ \boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B}) \tag{B-12}$$

where $\mathbf{A}$ and $\mathbf{B}$ are two arbitrary vectors, or two vector operators whose three components commute with those of the spin $\mathbf{S}$. If $\mathbf{A}$ and $\mathbf{B}$ do not commute with each other, the identity remains valid if $\mathbf{A}$ and $\mathbf{B}$ appear in the same order on the right-hand side as on the left-hand side.

The operators associated with electron spin have all the properties that follow directly from the general theory of angular momentum. They have, in addition, some specific properties related to their particular value of s (that is, of j), which is the smallest one possible (aside from zero). These specific properties can be deduced directly from (B-8) and formulas (B-10):

$$S_x^2 = S_y^2 = S_z^2 = \frac{\hbar^2}{4} \tag{B-13a}$$

$$S_xS_y + S_yS_x = 0 \tag{B-13b}$$

$$S_xS_y = \frac{i}{2}\hbar S_z \tag{B-13c}$$

$$S_+^2 = S_-^2 = 0 \tag{B-13d}$$

where the unit operator $\mathbb{1}$ is not explicitly written in the right hand side of (B-13a), as we will do from now on for the sake of simplicity.

C. Non-relativistic description of a spin 1/2 particle

We now know how to describe separately the external (orbital) and the internal (spin) degrees of freedom of the electron. In this section, we are going to assemble these different concepts into one formalism.

C-1. Observables and state vectors

C-1-a. State space

When all its degrees of freedom are taken into account, the quantum state of an electron is characterized by a ket belonging to the space $\mathcal{E}$ which is the tensor product of $\mathcal{E}_{\mathbf{r}}$ and $\mathcal{E}_s$ (§ A-2).

We extend into $\mathcal{E}$, following the method described in § F-2-b of Chapter II, both the operators originally defined in $\mathcal{E}_{\mathbf{r}}$ and those which initially acted in $\mathcal{E}_s$ (we shall continue to use the same notation for these extended operators as for the operators from which they are derived). We thus obtain a C.S.C.O. in $\mathcal{E}$ through the juxtaposition of a C.S.C.O. of $\mathcal{E}_{\mathbf{r}}$ and one of $\mathcal{E}_s$. For example, in $\mathcal{E}_s$, we can take $\mathbf{S}^2$ and S_z (or $\mathbf{S}^2$ and any component of $\mathbf{S}$). In $\mathcal{E}_{\mathbf{r}}$, we can choose $\{X, Y, Z\}$, or $\{P_x, P_y, P_z\}$, or, if H designates the Hamiltonian associated with a central potential, $\left\{H, \mathbf{L}^2, L_z\right\}$ etc. From this we deduce various C.S.C.O. in $\mathcal{E}$:

$$\left\{X, Y, Z, \mathbf{S}^2, S_z\right\} \tag{C-1a}$$

$$\left\{P_x, P_y, P_z, \mathbf{S}^2, S_z\right\} \tag{C-1b}$$

$$\left\{H, \mathbf{L}^2, L_z, \mathbf{S}^2, S_z\right\} \tag{C-1c}$$

etc. Since all kets of $\mathcal{E}$ are eigenvectors of $\mathbf{S}^2$ with the same eigenvalue [formula (B-5)], we can omit $\mathbf{S}^2$ from the sets of observables.

We are going to use here the first of these C.S.C.O., (C-1a). We shall take as a basis of $\mathcal{E}$ the set of vectors obtained from the tensor product of the kets $|\mathbf{r}\rangle \equiv |x, y, z\rangle$ of $\mathcal{E}_{\mathbf{r}}$ and the kets $|\varepsilon\rangle$ of $\mathcal{E}_s$:

$$|\mathbf{r}, \varepsilon\rangle \equiv |x, y, z, \varepsilon\rangle = |\mathbf{r}\rangle \otimes |\varepsilon\rangle \tag{C-2}$$

where the x, y, z, components of the vector $\mathbf{r}$ can vary from $-\infty$ to $+\infty$ (continuous indices), and ε is equal to $+$ or $-$ (discrete index). By definition, $|\mathbf{r}, \varepsilon\rangle$ is an eigenvector common to X, Y, Z, $\mathbf{S}^2$ and S_z:

$$\begin{aligned}
X\,|\mathbf{r}, \varepsilon\rangle &= x\,|\mathbf{r}, \varepsilon\rangle \\
Y\,|\mathbf{r}, \varepsilon\rangle &= y\,|\mathbf{r}, \varepsilon\rangle \\
Z\,|\mathbf{r}, \varepsilon\rangle &= z\,|\mathbf{r}, \varepsilon\rangle \\
\mathbf{S}^2\,|\mathbf{r}, \varepsilon\rangle &= \frac{3}{4}\hbar^2\,|\mathbf{r}, \varepsilon\rangle \\
S_z\,|\mathbf{r}, \varepsilon\rangle &= \varepsilon\frac{\hbar}{2}\,|\mathbf{r}, \varepsilon\rangle
\end{aligned} \tag{C-3}$$

Each ket $|\mathbf{r}, \varepsilon\rangle$ is unique to within a constant factor, since X, Y, Z, $\mathbf{S}^2$ and S_z constitute a C.S.C.O. The $\{|\mathbf{r}, \varepsilon\rangle\}$ system is orthonormal (in the extended sense), since the sets $\{|\mathbf{r}\rangle\}$ and $\{|+\rangle, |-\rangle\}$ are each orthonormal in $\mathcal{E}_{\mathbf{r}}$ and $\mathcal{E}_s$ respectively:

$$\langle \mathbf{r}', \varepsilon' \mid \mathbf{r}, \varepsilon\rangle = \delta_{\varepsilon'\varepsilon}\delta(\mathbf{r}' - \mathbf{r}) \tag{C-4}$$

($\delta_{\varepsilon'\varepsilon}$ is equal to 1 or 0 depending on whether ε' and ε are the same or different). Finally, it satisfies a closure relation in $\mathcal{E}$:

$$\sum_{\varepsilon} \int \mathrm{d}^3r\,|\mathbf{r}, \varepsilon\rangle\langle\mathbf{r}, \varepsilon| = \int \mathrm{d}^3r\,|\mathbf{r}, +\rangle\langle\mathbf{r}, +| + \int \mathrm{d}^3r\,|\mathbf{r}, -\rangle\langle\mathbf{r}, -| = 1 \tag{C-5}$$

C-1-b. $\{|\mathbf{r}, \varepsilon\rangle\}$ representation

α. State vectors

Any state $|\psi\rangle$ of the space $\mathcal{E}$ can be expanded on the $\{|\mathbf{r}, \varepsilon\rangle\}$ basis. To do this, it suffices to use the closure relation (C-5):

$$|\psi\rangle = \sum_{\varepsilon} \int \mathrm{d}^3 r \, |\mathbf{r}, \varepsilon\rangle \langle \mathbf{r}, \varepsilon \mid \psi\rangle \tag{C-6}$$

The *vector* $|\psi\rangle$ can therefore be represented by the set of its coordinates in the $\{|\mathbf{r}, \varepsilon\rangle\}$ basis, that is, by the *numbers*:

$$\langle \mathbf{r}, \varepsilon \mid \psi\rangle = \psi_\varepsilon(\mathbf{r}) \tag{C-7}$$

which depend on the three continuous indices x, y, z (or, more succinctly, $\mathbf{r}$) and on the discrete index ε (+ or −). *In order to characterize the state of an electron completely, it is therefore necessary to specify two functions of the space variables x, y and z*:

$$\begin{aligned} \psi_+(\mathbf{r}) &= \langle \mathbf{r}, + \mid \psi\rangle \\ \psi_-(\mathbf{r}) &= \langle \mathbf{r}, - \mid \psi\rangle \end{aligned} \tag{C-8}$$

These two functions are often written in the form of a *two-component spinor*, which we shall write $[\psi](\mathbf{r})$:

$$[\psi](\mathbf{r}) = \begin{pmatrix} \psi_+(\mathbf{r}) \\ \psi_-(\mathbf{r}) \end{pmatrix} \tag{C-9}$$

The bra $\langle\psi|$ associated with the ket $|\psi\rangle$ is given by the adjoint of (C-6):

$$\langle\psi| = \sum_{\varepsilon} \int \mathrm{d}^3 r \, \langle\psi \mid \mathbf{r}, \varepsilon\rangle \langle \mathbf{r}, \varepsilon| \tag{C-10}$$

that is, taking (C-7) into account:

$$\langle\psi| = \sum_{\varepsilon} \int \mathrm{d}^3 r \; \psi^*_\varepsilon(\mathbf{r}) \langle \mathbf{r}, \varepsilon| \tag{C-11}$$

The bra $\langle\psi|$ is thus represented by the two functions $\psi^*_+(\mathbf{r})$ and $\psi^*_-(\mathbf{r})$, which can be written in the form of a spinor which is the adjoint of (C-9):

$$[\psi]^\dagger(\mathbf{r}) = \begin{pmatrix} \psi^*_+(\mathbf{r}) & \psi^*_-(\mathbf{r}) \end{pmatrix} \tag{C-12}$$

With this notation, the scalar product of two state vectors $|\psi\rangle$ and $|\varphi\rangle$, which, according to (C-5), is equal to:

$$\begin{aligned} \langle\psi|\varphi\rangle &= \sum_{\varepsilon} \int \mathrm{d}^3 r \, \langle\psi \mid \mathbf{r}, \varepsilon\rangle \langle \mathbf{r}, \varepsilon \mid \varphi\rangle \\ &= \int \mathrm{d}^3 r \left[\psi^*_+(\mathbf{r})\varphi_+(\mathbf{r}) + \psi^*_-(\mathbf{r})\varphi_-(\mathbf{r})\right] \end{aligned} \tag{C-13}$$

can be written in the form:

$$\langle \psi \mid \varphi \rangle = \int \mathrm{d}^3 r \; [\psi]^\dagger(\mathbf{r})\,[\varphi]\,(\mathbf{r}) \tag{C-14}$$

This formula is very similar to the one that permitted the calculation of the scalar product of two kets of $\mathcal{E}_\mathbf{r}$ from the corresponding wave functions. However, it is important to note that here the matrix multiplication of the spinors $[\psi]^\dagger(\mathbf{r})$ and $[\varphi]\,(\mathbf{r})$ must precede the spatial integration. In particular, the normalization of the vector $|\psi\rangle$ is expressed by:

$$\langle \psi | \psi \rangle \; = \int \mathrm{d}^3 r \, [\psi]^\dagger \,(\mathbf{r})\,[\psi]\,(\mathbf{r}) = \int \mathrm{d}^3 r \left[|\psi_+(\mathbf{r})|^2 + |\psi_-(\mathbf{r})|^2 \right] = 1 \tag{C-15}$$

Amongst the vectors of $\mathcal{E}$, some are the tensor products of a ket of $\mathcal{E}_\mathbf{r}$ and a ket of $\mathcal{E}_s$ (this is the case, for example, for the basis vectors). If the state vector under consideration is of this type:

$$|\psi\rangle \; = |\varphi\rangle \otimes |\chi\rangle \tag{C-16}$$

with:

$$\begin{aligned} |\varphi\rangle \; &= \int \mathrm{d}^3 r \; \varphi(\mathbf{r})\,|\mathbf{r}\rangle \; \in \mathcal{E}_\mathbf{r} \\ |\chi\rangle \; &= c_+ \,|+\rangle + c_- \,|-\rangle \; \in \mathcal{E}_s \end{aligned} \tag{C-17}$$

the spinor associated with it takes on the simple form:

$$[\psi](\mathbf{r}) = \begin{pmatrix} \varphi(\mathbf{r}) c_+ \\ \varphi(\mathbf{r}) c_- \end{pmatrix} = \varphi(\mathbf{r}) \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \tag{C-18}$$

This results from the definition of the scalar product in $\mathcal{E}$, and we have in this case:

$$\psi_+(\mathbf{r}) = \; \langle \mathbf{r}, + \mid \psi \rangle \; = \; \langle \mathbf{r} \mid \varphi \rangle \; \langle + \mid \chi \rangle \; = \varphi(\mathbf{r}) c_+ \tag{C-19a}$$

$$\psi_-(\mathbf{r}) = \; \langle \mathbf{r}, - \mid \psi \rangle \; = \; \langle \mathbf{r} \mid \varphi \rangle \; \langle - \mid \chi \rangle \; = \varphi(\mathbf{r}) c_- \tag{C-19b}$$

The square of the norm of $|\psi\rangle$ is then given by:

$$\langle \psi | \psi \rangle = \langle \varphi | \varphi \rangle \; \langle \chi | \chi \rangle = \left(\mid c_+ \mid^2 + \mid c_- \mid^2 \right) \int \mathrm{d}^3 r \mid \varphi(\mathbf{r}) \mid^2 \tag{C-20}$$

β. Operators

Let $|\psi'\rangle$ be the ket obtained from the action of the linear operator A on the ket $|\psi\rangle$ of $\mathcal{E}$. According to the results of the preceding section, $|\psi'\rangle$ and $|\psi\rangle$ can be represented by the two-component spinors $[\psi'](\mathbf{r})$ and $[\psi](\mathbf{r})$. We are now going to show that one can associate with A a 2×2 matrix $[\![A]\!]$ such that:

$$[\psi'](\mathbf{r}) = [\![A]\!][\psi](\mathbf{r}) \tag{C-21}$$

where the matrix elements remain in general differential operators with respect to the variable $\mathbf{r}$.

(i) Spin operators. These were initially defined in $\mathcal{E}_s$. Consequently, they act only on the ε index of the basis vectors $|\mathbf{r}, \varepsilon\rangle$, and their matrix form is the one stated in § B. We shall limit ourselves to one example, say that of the operator S_+. Its action on a vector $|\psi\rangle$ expanded as in (C-6) gives a vector $|\psi'\rangle$:

$$|\psi'\rangle = \hbar \int \mathrm{d}^3 r \; \psi_-(\mathbf{r}) \, |\mathbf{r}, +\rangle \tag{C-22}$$

since S_+ annihilates all the $|\mathbf{r}, +\rangle$ kets and transforms $|\mathbf{r}, -\rangle$ into $\hbar \, |\mathbf{r}, +\rangle$. The components of $|\psi'\rangle$ in the $\{|\mathbf{r}, \varepsilon\rangle\}$ basis are, according to (C-22):

$$\begin{aligned} \langle \mathbf{r}, + \, |\psi'\rangle &= \psi'_+(\mathbf{r}) = \hbar \psi_-(\mathbf{r}) \\ \langle \mathbf{r}, - \, |\psi'\rangle &= \psi'_-(\mathbf{r}) = 0 \end{aligned} \tag{C-23}$$

The spinor representing $|\psi'\rangle$ is therefore:

$$[\psi'](\mathbf{r}) = \hbar \begin{pmatrix} \psi_-(\mathbf{r}) \\ 0 \end{pmatrix} \tag{C-24}$$

This is indeed what is obtained if one performs the matrix multiplication of the spinor $[\psi](\mathbf{r})$ by:

$$[\![S_+]\!] = \frac{\hbar}{2}(\sigma_x + i\sigma_y) = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{C-25}$$

(ii) Orbital operators. Unlike the preceding operators, they always leave unchanged the ε index of the basis vector $|\mathbf{r}, \varepsilon\rangle$: their associated 2×2 matrices are always proportional to the unit matrix. On the other hand, they act on the $\mathbf{r}$-dependence of the spinors just as they act on ordinary wave functions. Consider, for example, the kets $|\psi'\rangle = X \, |\psi\rangle$ and $|\psi''\rangle = P_x \, |\psi\rangle$. Their components in the $\{|\mathbf{r}, \varepsilon\rangle\}$ basis are, respectively:

$$\psi'_\varepsilon(\mathbf{r}) = \langle \mathbf{r}, \varepsilon \mid X \mid \psi \rangle = x \; \psi_\varepsilon(\mathbf{r}) \tag{C-26a}$$

$$\psi''_\varepsilon(\mathbf{r}) = \langle \mathbf{r}, \varepsilon \mid P_x \mid \psi \rangle = \frac{\hbar}{i} \frac{\partial}{\partial x} \psi_\varepsilon(\mathbf{r}) \tag{C-26b}$$

The spinors $[\psi'](\mathbf{r})$ and $[\psi''](\mathbf{r})$ are thus obtained from $[\psi](\mathbf{r})$ by means of the 2×2 matrices:

$$[\![X]\!] = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} \tag{C-27a}$$

$$[\![P_x]\!] = \frac{\hbar}{i} \begin{pmatrix} \dfrac{\partial}{\partial x} & 0 \\ 0 & \dfrac{\partial}{\partial x} \end{pmatrix} \tag{C-27b}$$

(iii) Mixed operators. The most general operator acting in $\mathcal{E}$ is represented, in matrix notation, by a 2×2 matrix whose elements are differential operators with respect to the $\mathbf{r}$ variables. For example:

$$[\![L_z S_z]\!] = \frac{\hbar}{2} \begin{pmatrix} \dfrac{\hbar}{i} \dfrac{\partial}{\partial \varphi} & 0 \\ 0 & -\dfrac{\hbar}{i} \dfrac{\partial}{\partial \varphi} \end{pmatrix} \tag{C-28}$$

or:

$$[\![\mathbf{S}\cdot\mathbf{P}]\!] = \frac{\hbar}{2}(\sigma_x P_x + \sigma_y P_y + \sigma_z P_z) = \frac{\hbar^2}{2i}\begin{pmatrix} \dfrac{\partial}{\partial z} & \dfrac{\partial}{\partial x} - i\dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial x} + i\dfrac{\partial}{\partial y} & -\dfrac{\partial}{\partial z} \end{pmatrix} \tag{C-29}$$

Comments:

(*i*) The spinor representation $\{|\mathbf{r},\varepsilon\rangle\}$ is analogous to the $\{|\mathbf{r}\rangle\}$ representation $\mathcal{E}_{\mathbf{r}}$. The matrix element $\langle\psi|A|\varphi\rangle$ of any operator A of $\mathcal{E}$ is given by the formula:

$$\langle\psi|A|\varphi\rangle = \int \mathrm{d}^3 r [\psi]^\dagger(\mathbf{r})[\![A]\!][\varphi](\mathbf{r}) \tag{C-30}$$

where $[\![A]\!]$ designates the 2×2 matrix that represents the operator A (one first carries out the matrix multiplications and then integrates over all space). This representation will only be used when it simplifies the reasoning and the calculations: as in $\mathcal{E}_{\mathbf{r}}$, the vectors and operators themselves will be used as much as possible.

(*ii*) Obviously, there also exists a $\{|\mathbf{p},\varepsilon\rangle\}$ representation, whose basis vectors are the eigenvectors common to the C.S.C.O. $\{P_x, P_y, P_z, \mathbf{S}^2, S_z\}$. The definition of the scalar product in $\mathcal{E}$ yields:

$$\langle\mathbf{r},\varepsilon \mid \mathbf{p},\varepsilon'\rangle = \langle\mathbf{r}\mid\mathbf{p}\rangle\langle\varepsilon\mid\varepsilon'\rangle = \frac{1}{(2\pi\hbar)^{3/2}}\mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\delta_{\varepsilon\varepsilon'} \tag{C-31}$$

In the $\{|\mathbf{p},\varepsilon\rangle\}$ representation, one associates with each vector $|\psi\rangle$ of $\mathcal{E}$ a two-component spinor:

$$[\overline{\psi}](\mathbf{p}) = \begin{pmatrix}\overline{\psi}_+(\mathbf{p}) \\ \overline{\psi}_-(\mathbf{p})\end{pmatrix} \tag{C-32}$$

with:

$$\begin{aligned}\overline{\psi}_+(\mathbf{p}) &= \langle\mathbf{p},+\mid\psi\rangle \\ \overline{\psi}_-(\mathbf{p}) &= \langle\mathbf{p},-\mid\psi\rangle\end{aligned} \tag{C-33}$$

According to (C-31), $\overline{\psi}_+(\mathbf{p})$ and $\overline{\psi}_-(\mathbf{p})$ are the Fourier transforms of $\psi_+(\mathbf{r})$ and $\psi_-(\mathbf{r})$:

$$\begin{aligned}\overline{\psi}_\varepsilon(\mathbf{p}) = \langle\mathbf{p},\varepsilon\mid\psi\rangle &= \sum_{\varepsilon'}\int \mathrm{d}^3 r\,\langle\mathbf{p},\varepsilon\mid\mathbf{r},\varepsilon'\rangle\langle\mathbf{r},\varepsilon'\mid\psi\rangle \\ &= \frac{1}{(2\pi\hbar)^{3/2}}\int \mathrm{d}^3 r\ \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}\psi_\varepsilon(\mathbf{r})\end{aligned} \tag{C-34}$$

The operators are still represented by 2×2 matrices, and those corresponding to the spin operators remain the same as in the $\{|\mathbf{r},\varepsilon\rangle\}$ representation.

C-2. Probability calculations for a physical measurement

Using the formalism we have just described, we can apply the postulates of Chapter III to obtain predictions concerning the various measurements that one can imagine carrying out on an electron. We are going to give several examples.

First of all, consider the probabilistic interpretation of the components $\psi_+(\mathbf{r})$ and $\psi_-(\mathbf{r})$ of the state vector $|\psi\rangle$ which we assume to be normalized [formula (C-15)]. Imagine that we are simultaneously measuring the position of the electron and the component of its spin along Oz. Since X, Y, Z and S_z constitute a C.S.C.O., there exists only one state vector that corresponds to a given result: x, y, z and $\pm\hbar/2$. The probability $\mathrm{d}^3\mathcal{P}(\mathbf{r},+)$ of the electron being found in the infinitesimal volume d^3r around the point $\mathbf{r}(x,y,z)$ with its spin "up" (component along Oz equal to $+\hbar/2$) is equal to:

$$\mathrm{d}^3\mathcal{P}(\mathbf{r},+) = |\langle \mathbf{r},+|\psi\rangle|^2\,\mathrm{d}^3r = |\psi_+(\mathbf{r})|^2\,\mathrm{d}^3r \tag{C-35}$$

In the same way:

$$\mathrm{d}^3\mathcal{P}(\mathbf{r},-) = |\langle \mathbf{r},-|\psi\rangle|^2\,\mathrm{d}^3r = |\psi_-(\mathbf{r})|^2\,\mathrm{d}^3r \tag{C-36}$$

is the probability of the electron being found in the same volume as before but with its spin "down" (component along Oz equal to $-\hbar/2$).

If it is the component of the spin along Ox that is being measured at the same time as the position, all we need to do is use formulas (A-20) of Chapter IV. The X, Y, Z and S_x operators also form a C.S.C.O.: to the measurement result $\{x,\ y,\ z,\ \pm\hbar/2\}$ corresponds a single state vector:

$$|\mathbf{r}\rangle\,|\pm\rangle_x = \frac{1}{\sqrt{2}}\left[|\mathbf{r},+\rangle \pm |\mathbf{r},-\rangle\right] \tag{C-37}$$

The probability of the electron being found in the volume d^3r around the point $\mathbf{r}$ with its spin in the positive direction of the Ox axis is then:

$$\mathrm{d}^3r \times \left|\frac{1}{\sqrt{2}}\left[\langle \mathbf{r},+\,|\,\psi\rangle + \langle \mathbf{r},-\,|\,\psi\rangle\right]\right|^2 = \frac{1}{2}\,|\psi_+(\mathbf{r}) + \psi_-(\mathbf{r})|^2\,\mathrm{d}^3r \tag{C-38}$$

Obviously, one can measure the momentum of the electron instead of its position. One then uses the components of $|\psi\rangle$ relative to the vectors $|\mathbf{p},\ \varepsilon\rangle$ [*cf.* comment *(ii)* of § 1], that is, the Fourier transforms $\overline{\psi}_\pm(\mathbf{p})$ of $\psi_\pm(\mathbf{r})$. The probability $\mathrm{d}^3\mathcal{P}(\mathbf{p},\ \pm)$ of the momentum being $\mathbf{p}$ to within d^3p and of the spin component along Oz being $\pm\hbar/2$ is given by:

$$\mathrm{d}^3\mathcal{P}(\mathbf{p},\ \pm) = |\ \langle \mathbf{p},\ \pm|\psi\rangle\ |^2\ \mathrm{d}^3p = |\overline{\psi}_\pm(\mathbf{p})|^2\ \mathrm{d}^3p \tag{C-39}$$

The various measurements that we have envisaged until now are all "complete" in the sense that they each relate to a C.S.C.O. For "incomplete measurements", several orthogonal states correspond to the same result, and it is necessary to sum the squares of the moduli of the corresponding probability amplitudes.

For example, if one does not seek to measure its spin, the probability $\mathrm{d}^3\mathcal{P}(\mathbf{r})$ of finding the electron in the volume d^3r in the neighborhood of the point $\mathbf{r}$ is equal to:

$$\mathrm{d}^3\mathcal{P}(\mathbf{r}) = \left[|\psi_+(\mathbf{r})|^2 + |\psi_-(\mathbf{r})|^2\right]\mathrm{d}^3r \tag{C-40}$$

This is because two orthogonal state vectors, $|\mathbf{r}, +\rangle$ and $|\mathbf{r}, -\rangle$, are associated with the result $\{x, y, z\}$, their corresponding probability amplitudes being $\psi_+(\mathbf{r})$ and $\psi_-(\mathbf{r})$.

Finally, let us calculate the probability $\mathcal{P}_+$ that the spin component along Oz is $+\hbar/2$ (one is not seeking to measure the orbital variables). There exist an infinite number of orthogonal states, for example all the $|\mathbf{r}, +\rangle$ with arbitrary $\mathbf{r}$, which correspond to the result of the measurement. One must therefore sum over all possible values of $\mathbf{r}$ the squares of the moduli of the amplitudes $\langle \mathbf{r}, + \mid \psi \rangle = \psi_+(\mathbf{r})$, which gives:

$$\mathcal{P}_+ = \int \mathrm{d}^3 r \; |\psi_+(\mathbf{r})|^2 \tag{C-41}$$

Of course, if we are considering the component of the spin along Ox instead of along Oz, we integrate the result (C-38) over all space. These ideas generalize those of § B-2 of Chapter IV, where we considered only the spin observables since the orbital variables could be treated classically.

References and suggestions for further reading:

History of the discovery of spin and references to original articles: Jammer (4.8), § 3-4.

Evidence of spin in atomic physics: Eisberg and Resnick (1.3), Chap. 8; Born (11.4), Chap. VI; Kuhn (11.1), Chap. III, §§ A.5, A.6 and F; see references of Chapter IV relating to the Stern-Gerlach experiment.

The spin magnetic moment of the electron: Cagnac and Pebay-Peyroula (11.2). Chap. XII; Crane (11.16).

The Dirac equation: Schiff (1.18), Chap. 13; Messiah (1.17). Chap. XX: Bjorkcn and Drell (2.6), Chaps. 1 to 4

The Lorentz group: Omnes (16.13), Chap. 4; Bacry (10.31). Chaps. 7 and 8.

Spin 1 particles: Messiah (1.17), § XIII.21.

Several complements concerning the properties of spins 1/2 can be found at the end of Chapter IV; this is why Chapter IX has only two complements.

A_{IX} : ROTATION OPERATORS FOR A SPIN 1/2 PARTICLE	This complement is a continuation of Complement B_{VI}. It studies in detail the relationship between the spin $\frac{1}{2}$ angular momentum and the geometric rotations of this spin. Moderately difficult. Can be omitted upon a first reading.
B_{IX} : EXERCISES	Exercice 4 is worked out in detail. It studies the polarization of a beam of spin $\frac{1}{2}$ particles caused by their reflection from a magnetized ferromagnetic material. This method is actually used in certain experiments.

Complement A_{IX}

Rotation operators for a spin 1/2 particle

1 Rotation operators in state space 1001
1-a Total angular momentum 1001
1-b Decomposition of rotation operators into tensor products . . 1001
2 Rotation of spin states 1002
2-a Explicit calculation of the rotation operators in $\mathcal{E}_s$ 1002
2-b Operator associated with a rotation through an angle of 2π . 1003
2-c Relationship between the vectorial nature of $\mathbf{S}$ and the behavior of a spin state upon rotation 1004
3 Rotation of two-component spinors 1005

We are going to apply the ideas about rotation introduced in Complement B_{VI} to the case of a spin 1/2 particle. First, we shall study the form that rotation operators take on in this case. We shall then examine the behavior, under rotation, of the ket representing the particle's state and of the two-component spinor associated with it.

1. Rotation operators in state space

1-a. Total angular momentum

A spin 1/2 particle possesses an orbital angular momentum $\mathbf{L}$ and a spin angular momentum $\mathbf{S}$. It is natural to define its total angular momentum as the sum of these two angular momenta:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{1}$$

This definition is clearly consistent with the general considerations discussed in Complement B_{VI}. It insures that not only $\mathbf{R}$ and $\mathbf{P}$, but also $\mathbf{S}$, be vectorial observables. Note that, to test this, it is sufficient to calculate the commutators between the components of these observables and those of $\mathbf{J}$; *cf.* § 5-c of Complement B_{VI}.

1-b. Decomposition of rotation operators into tensor products

In the state space of the particle under study, the rotation operator $R_{\mathbf{u}}(\alpha)$ is associated with the geometrical rotation $\mathcal{R}_{\mathbf{u}}(\alpha)$ through an angle α about the unit vector $\mathbf{u}$ (*cf.* Complement B_{VI}, § 4):

$$R_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\alpha \mathbf{J}\cdot\mathbf{u}} \tag{2}$$

where $\mathbf{J}$ is the total angular momentum (1).

Since $\mathbf{L}$ acts only in $\mathcal{E}_{\mathbf{r}}$, and $\mathbf{S}$ only in $\mathcal{E}_s$ (which implies, in particular, that all components of $\mathbf{L}$ commute with all components of $\mathbf{S}$), we can write $R_{\mathbf{u}}(\alpha)$ in the form of a tensor product:

$$R_{\mathbf{u}}(\alpha) = {}^{(\mathbf{r})}R_{\mathbf{u}}(\alpha) \otimes {}^{(s)}R_{\mathbf{u}}(\alpha) \tag{3}$$

where:

$$ {}^{(\mathbf{r})}R_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\alpha \mathbf{L}\cdot\mathbf{u}} \tag{4} $$

and:

$$ {}^{(s)}R_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\alpha \mathbf{S}\cdot\mathbf{u}} \tag{5} $$

are the rotation operators associated with $\mathcal{R}_{\mathbf{u}}(\alpha)$ in $\mathcal{E}_{\mathbf{r}}$ and $\mathcal{E}_s$ respectively.

Consequently, if one performs the rotation $\mathcal{R}_{\mathbf{u}}(\alpha)$ on a spin 1/2 particle whose state is represented by a ket which is a tensor product:

$$ |\psi\rangle = |\varphi\rangle \otimes |\chi\rangle \tag{6} $$

with:

$$ \begin{aligned} |\varphi\rangle &\in \mathcal{E}_{\mathbf{r}} \\ |\chi\rangle &\in \mathcal{E}_s \end{aligned} \tag{7} $$

its state after rotation will be:

$$ |\psi'\rangle = R_{\mathbf{u}}(\alpha)\,|\psi\rangle = \left[{}^{(\mathbf{r})}R_{\mathbf{u}}(\alpha)\,|\varphi\rangle\right] \otimes \left[{}^{(s)}R_{\mathbf{u}}(\alpha)\,|\chi\rangle\right] \tag{8} $$

The spin state of the particle is therefore also affected by the rotation. This is what we are going to study in more detail in § 2.

2. Rotation of spin states

We have already studied (§ 3 of Complement B_{VI}) the rotation operators ${}^{(\mathbf{r})}R$ in the space $\mathcal{E}_{\mathbf{r}}$. Here we are interested in the operators ${}^{(s)}R$ which act in the spin state space $\mathcal{E}_s$.

2-a. Explicit calculation of the rotation operators in $\mathcal{E}_s$

As in Chapter IX, we set:

$$ \mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma} \tag{9} $$

We want to calculate the operator:

$$ {}^{(s)}R_{\mathbf{u}}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\alpha \mathbf{S}\cdot\mathbf{u}} = \mathrm{e}^{-i\frac{\alpha}{2}\boldsymbol{\sigma}\cdot\mathbf{u}} \tag{10} $$

To do this, let us use the definition of the exponential of an operator:

$$ {}^{(s)}R_{\mathbf{u}}(\alpha) = 1 - \frac{i\alpha}{2}\boldsymbol{\sigma}\cdot\mathbf{u} + \frac{1}{2!}\left(-i\frac{\alpha}{2}\right)^2(\boldsymbol{\sigma}\cdot\mathbf{u})^2 + \ldots + \frac{1}{n!}\left(-i\frac{\alpha}{2}\right)^n(\boldsymbol{\sigma}\cdot\mathbf{u})^n + \ldots \tag{11} $$

Now, applying identity (B-12) of Chapter IX, we immediately see that:

$$ (\boldsymbol{\sigma}\cdot\mathbf{u})^2 = \mathbf{u}^2 = 1 \tag{12} $$

which leads to:

$$(\boldsymbol{\sigma}\cdot\mathbf{u})^n = \begin{cases} 1 & \text{if } n \text{ is even} \\ \boldsymbol{\sigma}\cdot\mathbf{u} & \text{if } n \text{ is odd} \end{cases} \tag{13}$$

Consequently, if we group together the even and odd terms respectively, expansion (11) can be written:

$$\begin{aligned} {}^{(s)}R_{\mathbf{u}}(\alpha) = &\left[1 - \frac{1}{2!}\left(\frac{\alpha}{2}\right)^2 + \ldots + \frac{(-1)^p}{(2p)!}\left(\frac{\alpha}{2}\right)^{2p} + \ldots\right] \\ &- i\boldsymbol{\sigma}\cdot\mathbf{u}\left[\frac{\alpha}{2} - \frac{1}{3!}\left(\frac{\alpha}{2}\right)^3 + \ldots + \frac{(-1)^p}{(2p+1)!}\left(\frac{\alpha}{2}\right)^{2p+1} + \ldots\right] \end{aligned} \tag{14}$$

that is, finally:

$$\boxed{{}^{(s)}R_{\mathbf{u}}(\alpha) = \cos\frac{\alpha}{2} - i\boldsymbol{\sigma}\cdot\mathbf{u}\sin\frac{\alpha}{2}} \tag{15}$$

It will be very easy to calculate the action of the operator ${}^{(s)}R$, in this form, on any spin state.

Using this formula, we can write the rotation matrix $R_{\mathbf{u}}{}^{(1/2)}(\alpha)$ explicitly in the $\{|+\rangle, |-\rangle\}$ basis, since we already know [formulas (B-9) of Chapter IX] the matrices which represent the σ_x, σ_y and σ_z operators. We find:

$$R_{\mathbf{u}}^{(1/2)}(\alpha) = \begin{pmatrix} \cos\frac{\alpha}{2} - iu_z\sin\frac{\alpha}{2} & (-iu_x - u_y)\sin\frac{\alpha}{2} \\ (-iu_x + u_y)\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} + iu_z\sin\frac{\alpha}{2} \end{pmatrix} \tag{16}$$

where u_x, u_y and u_z are the cartesian components of the vector $\mathbf{u}$.

2-b. Operator associated with a rotation through an angle of 2π

If we take 2π for the angle of rotation α, the geometrical rotation $\mathcal{R}_{\mathbf{u}}(2\pi)$ coincides, whatever the vector $\mathbf{u}$ may be, with the identity rotation. However, if we set $\alpha = 2\pi$ in formula (15), we see that:

$${}^{(s)}R_{\mathbf{u}}(2\pi) = -1 \tag{17}$$

whereas:

$${}^{(s)}R_{\mathbf{u}}(0) = 1 \tag{18}$$

The operator associated with a rotation through an angle of 2π is not the identity operator, but minus this operator. The group law is therefore conserved only locally in the correspondance between geometrical rotations and rotation operators in $\mathcal{E}_s$ [see discussion in Complement B_{VI}, comment (*iii*) of § 3-c-γ]. This is due to the half-integral value of the spin angular momentum of the particle which we are considering.

The fact that the spin state changes sign during a rotation through an angle of 2π is not disturbing, since two state vectors differing only by a global phase factor have the

same physical properties. It is more important to study the way in which an observable A transforms during such a rotation. It is easy to show that:

$$A' = {}^{(s)}R_{\mathbf{u}}(2\pi)\, A\, {}^{(s)}R^{\dagger}_{\mathbf{u}}(2\pi) = A \tag{19}$$

This result is quite satisfying since a rotation through 2π cannot modify the measuring device associated with A. Consequently, the spectrum of A' must remain the same as that of A.

Comment:

We showed in Complement B_{VI} [comment *(iii)* of § 3-c-γ] that:

$${}^{(\mathbf{r})}R_{\mathbf{u}}(2\pi) = 1 \tag{20}$$

Consequently, in the global state space $\mathcal{E} = \mathcal{E}_{\mathbf{r}} \otimes \mathcal{E}_s$, as in $\mathcal{E}_s$, we have:

$$R_{\mathbf{u}}(2\pi) = {}^{(\mathbf{r})}R_{\mathbf{u}}(2\pi) \otimes {}^{(s)}R_{\mathbf{u}}(2\pi) = -1 \tag{21}$$

2-c. Relationship between the vectorial nature of S and the behavior of a spin state upon rotation

Consider an arbitrary spin state $|\chi\rangle$. We showed in Chapter IV (§ B-1-c) that there must exist angles θ and φ such that $|\chi\rangle$ can be written (except for a global phase factor which has no physical meaning):

$$|\chi\rangle = \mathrm{e}^{-i\varphi/2}\cos\frac{\theta}{2}\,|+\rangle + \mathrm{e}^{i\varphi/2}\sin\frac{\theta}{2}\,|-\rangle \tag{22}$$

$|\chi\rangle$ then appears as the eigenvector associated with the eigenvalue $+\hbar/2$ of the component $\mathbf{S}\cdot\mathbf{v}$ of the spin $\mathbf{S}$ along the unit vector $\mathbf{v}$ defined by the polar angles θ and φ. Now let us perform an arbitrary rotation on the state $|\chi\rangle$. Let us call $\mathbf{v}'$ the result of the transformation of $\mathbf{v}$ by the rotation being considered. Since $\mathbf{S}$ is a vectorial observable, the state $|\chi'\rangle$ after the rotation must be an eigenvector, with the eigenvalue $+\hbar/2$, of the component $\mathbf{S}\cdot\mathbf{v}'$ of $\mathbf{S}$ along the unit vector $\mathbf{v}'$ (*cf.* Complement B_{VI}, § 5):

$$|\chi\rangle = |+\rangle_v \Longrightarrow \left|\chi'\right\rangle = R\,|\chi\rangle \quad \propto |+\rangle_{v'} \tag{23}$$

with:

$$\mathbf{v}' = \mathcal{R}\mathbf{v} \tag{24}$$

We shall be satisfied with verifying this for a specific case (*cf.* Fig. 1). Choose for $\mathbf{v}$ the unit vector $\mathbf{e}_z$ of the Oz axis, and for $\mathbf{v}'$ an arbitrary unit vector, with polar angles θ and φ. $\mathbf{v}'$ is obtained from $\mathbf{v} = \mathbf{e}_z$ by a rotation through an angle θ about the unit vector $\mathbf{u}$, which is fixed by the polar angles:

$$\begin{aligned}\theta_u &= \frac{\pi}{2}\\ \varphi_u &= \varphi + \frac{\pi}{2}\end{aligned} \tag{25}$$

Thus we must show that:

$${}^{(s)}R_{\mathbf{u}}(\theta)\,|+\rangle \propto |+\rangle_{v'} \tag{26}$$

The cartesian components of the vector $\mathbf{u}$ are:

$$\begin{aligned}u_x &= -\sin\varphi\\ u_y &= \cos\varphi\\ u_z &= 0\end{aligned} \tag{27}$$

so the operator ${}^{(s)}R_{\mathbf{u}}(\theta)$ can be written, using formula (15):

$$\begin{aligned} {}^{(s)}R_{\mathbf{u}}(\theta) &= \cos\frac{\theta}{2} - i\boldsymbol{\sigma}\cdot\mathbf{u}\sin\frac{\theta}{2} \\ &= \cos\frac{\theta}{2} - i\left(-\sigma_x\sin\varphi + \sigma_y\cos\varphi\right)\sin\frac{\theta}{2} \\ &= \cos\frac{\theta}{2} - \frac{1}{2}\left(\sigma_+ \mathrm{e}^{-i\varphi} - \sigma_- \mathrm{e}^{i\varphi}\right)\sin\frac{\theta}{2} \end{aligned} \tag{28}$$

with:

$$\sigma_\pm = \sigma_x \pm i\sigma_y \tag{29}$$

Now we know [*cf.* formulas (B-7) of Chapter IX] that:

$$\begin{aligned} &\sigma_+ |+\rangle = 0 \\ &\sigma_- |+\rangle = 2\,|-\rangle \end{aligned} \tag{30}$$

The result of the transformation of the ket $|+\rangle$ by the operator ${}^{(s)}R_{\mathbf{u}}(\theta)$ is therefore:

$${}^{(s)}R_{\mathbf{u}}(\theta)\,|+\rangle = \cos\frac{\theta}{2}\,|+\rangle + \mathrm{e}^{i\varphi}\sin\frac{\theta}{2}\,|-\rangle \tag{31}$$

We recognize, to within a phase factor, the ket $|+\rangle_{v'}$ [*cf.* formula (22)]:

$${}^{(s)}R_{\mathbf{u}}(\theta)\,|+\rangle = \mathrm{e}^{i\varphi/2}\,|+\rangle_{v'} \tag{32}$$

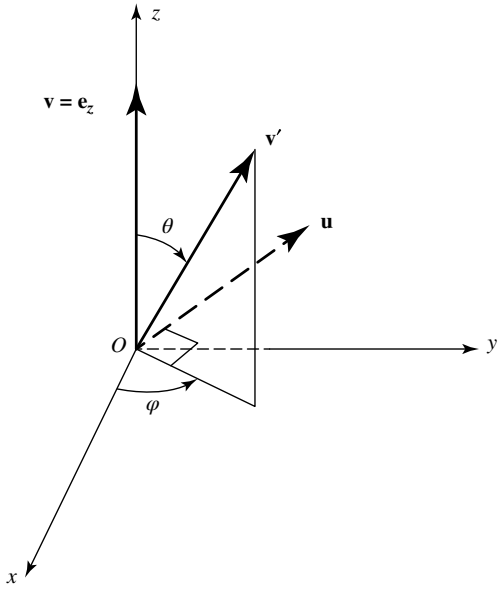


Figure 1: A rotation through an angle θ about $\mathbf{u}$ brings the vector $\mathbf{v} = \mathbf{e}_z$ onto the unit vector $\mathbf{v}'$, with polar angles θ and φ.

3. Rotation of two-component spinors

We are now prepared to study the global behavior of a spin 1/2 particle under rotation. That is, we shall now take into account both its external and internal degrees of freedom.

Consider a spin 1/2 particle whose state is represented by the ket $|\psi\rangle$ of the state space $\mathcal{E} = \mathcal{E}_{\mathbf{r}} \otimes \mathcal{E}_s$. The ket $|\psi\rangle$ can be represented by the spinor $[\psi](\mathbf{r})$, having the components:

$$\psi_\varepsilon(\mathbf{r}) = \langle \mathbf{r}, \varepsilon \mid \psi \rangle \tag{33}$$

If we perform an arbitrary geometrical rotation $\mathcal{R}$ on this particle, its state then becomes:

$$|\psi'\rangle = R\,|\psi\rangle \tag{34}$$

where:

$$R = {}^{(\mathbf{r})}R \otimes {}^{(s)}R \tag{35}$$

is the operator associated, in $\mathcal{E}$, with the geometrical rotation $\mathcal{R}$. How is the spinor, $[\psi'](\mathbf{r})$, which corresponds to the state $|\psi'\rangle$, obtained from $[\psi](\mathbf{r})$?

In order to answer this question, let us write the components $\psi'_\varepsilon(\mathbf{r})$ of $[\psi']$:

$$\psi'_\varepsilon(\mathbf{r}) = \langle \mathbf{r}, \varepsilon \mid \psi' \rangle = \langle \mathbf{r}, \varepsilon \mid R \mid \psi \rangle \tag{36}$$

We can find the components of $\psi(\mathbf{r})$ by inserting the closure relation relative to the $\{|\mathbf{r}', \varepsilon'\rangle\}$ basis between R and $|\psi\rangle$:

$$\psi'_\varepsilon(\mathbf{r}) = \sum_{\varepsilon'} \int \mathrm{d}^3 r' \, \langle \mathbf{r}, \varepsilon \mid R \mid \mathbf{r}', \varepsilon' \rangle \, \langle \mathbf{r}', \varepsilon' \mid \psi \rangle \tag{37}$$

Now, since the vectors of the $\{|\mathbf{r}, \varepsilon\rangle\}$ basis are tensor products, the matrix elements of the operator R in this basis can be decomposed in the following manner:

$$\langle \mathbf{r}, \varepsilon \mid R \mid \mathbf{r}', \varepsilon' \rangle = \left\langle \mathbf{r} \mid {}^{(\mathbf{r})}R \mid \mathbf{r}' \right\rangle \left\langle \varepsilon \mid {}^{(s)}R \mid \varepsilon' \right\rangle \tag{38}$$

We already know [*cf.* Complement B_{VI}, formula (26)] that:

$$\left\langle \mathbf{r} \mid {}^{(\mathbf{r})}R \mid \mathbf{r}' \right\rangle = \left\langle \mathcal{R}^{-1}\mathbf{r} \mid \mathbf{r}' \right\rangle = \delta\left[\mathbf{r}' - (\mathcal{R}^{-1}\mathbf{r})\right] \tag{39}$$

Consequently, if we set:

$$\left\langle \varepsilon \mid {}^{(s)}R \mid \varepsilon' \right\rangle = R^{(1/2)}_{\varepsilon\varepsilon'} \tag{40}$$

formula (37) can finally be written:

$$\boxed{\psi'_\varepsilon(\mathbf{r}) = \sum_{\varepsilon'} R^{(1/2)}_{\varepsilon\varepsilon'} \psi_{\varepsilon'}(\mathcal{R}^{-1}\mathbf{r})} \tag{41}$$

that is, explicitly:

$$\begin{pmatrix} \psi'_+(\mathbf{r}) \\ \psi'_-(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} R^{(1/2)}_{++} & R^{(1/2)}_{+-} \\ R^{(1/2)}_{-+} & R^{(1/2)}_{--} \end{pmatrix} \begin{pmatrix} \psi_+(\mathcal{R}^{-1}\mathbf{r}) \\ \psi_-(\mathcal{R}^{-1}\mathbf{r}) \end{pmatrix} \tag{42}$$

Thus we obtain the following result: each component of the new spinor $[\psi']$ at the point $\mathbf{r}$ is a linear combination of the two components of the original spinor $[\psi]$ evaluated at the point $\mathcal{R}^{-1}\mathbf{r}$ (that is, at the point that the rotation maps into $\mathbf{r}$)[1]. The coefficients of these linear combinations are the elements of the 2×2 matrix which represents ${}^{(s)}R$ in the $\{|+\rangle, |-\rangle\}$ basis of $\mathcal{E}_s$ [*cf.* formula (16)].

[1] Note the close analogy between this behavior and that of a vector field under rotation.

References and suggestions for further reading:

Feynman III (1.2), Chap. 6; Chap. 18, § 18-4 and added note 1; Messiah (1.17), App. C; Edmonds (2.21), Chap. 4.

Rotation groups and SU(2): Bacry (10.31), Chap. 6; Wigner (2.23), Chap. 15; Meijer and Bauer (2.18), Chap. 5.

Experiments dealing with rotations of a spin 1/2: article by Werner et al. (11.18).

Complement B_{IX}

Exercises

1. Consider a spin 1/2 particle. Call its spin $\mathbf{S}$, its orbital angular momentum $\mathbf{L}$ and its state vector $|\psi\rangle$. The two functions $\psi_+(\mathbf{r})$ and $\psi_-(\mathbf{r})$ are defined by:

$$\psi_\pm(\mathbf{r}) = \langle \mathbf{r}, \pm \mid \psi \rangle$$

Assume that:

$$\psi_+(\mathbf{r}) = R(r)\left[Y_0^0(\theta, \varphi) + \frac{1}{\sqrt{3}}Y_1^0(\theta, \varphi)\right]$$
$$\psi_-(\mathbf{r}) = \frac{R(r)}{\sqrt{3}}\left[Y_1^1(\theta, \varphi) - Y_1^0(\theta, \varphi)\right]$$

where r, θ, φ are the coordinates of the particle and $R(r)$ is a given function of r.

a. What condition must $R(r)$ satisfy for $|\psi\rangle$ to be normalized?

b. S_z is measured with the particle in the state $|\psi\rangle$. What results can be found, and with what probabilities? Same question for L_z, then for S_x.

c. A measurement of $\mathbf{L}^2$, with the particle in the state $|\psi\rangle$, yielded zero. What state describes the particle just after this measurement? Same question if the measurement of $\mathbf{L}^2$ had given $2\hbar^2$.

2. Consider a spin 1/2 particle. $\mathbf{P}$ and $\mathbf{S}$ designate the observables associated with its momentum and its spin. We choose as the basis of the state space the orthonormal basis $|p_x, p_y, p_z, \pm\rangle$ of eigenvectors common to P_x, P_y, P_z and S_z (whose eigenvalues are, respectively, p_x, p_y, p_z and $\pm\hbar/2$).

We intend to solve the eigenvalue equation of the operator A which is defined by:

$$A = \mathbf{S} \cdot \mathbf{P}$$

a. Is A Hermitian?

b. Show that there exists a basis of eigenvectors of A which are also eigenvectors of P_x, P_y, P_z. In the subspace spanned by the kets $|p_x, p_y, p_z, \pm\rangle$, where p_x, p_y, p_z are fixed, what is the matrix representing A?

c. What are the eigenvalues of A, and what is their degree of degeneracy? Find a system of eigenvectors common to A and P_x, P_y, P_z.

3. The Pauli Hamiltonian

The Hamiltonian of an electron ot mass m, charge q, spin $\hbar\boldsymbol{\sigma}/2$ (where σ_x, σ_y, σ_z are the Pauli matrices), placed in an electromagnetic field described by the vector potential $\mathbf{A}(\mathbf{r}, t)$ and the scalar potential $U(\mathbf{r}, t)$, is written:

$$H = \frac{1}{2m}[\mathbf{P} - q\mathbf{A}(\mathbf{R}, t)]^2 + qU(\mathbf{R}, t) - \frac{q\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}(\mathbf{R}, t)$$

The last term represents the interaction between the spin magnetic moment $\frac{q\hbar}{2m}\boldsymbol{\sigma}$ and the magnetic field $\mathbf{B}(\mathbf{R}, t) = \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{R}, t)$.

Show, using the properties of the Pauli matrices, that this Hamiltonian can also be written in the following form ("the Pauli Hamiltonian"):

$$H = \frac{1}{2m}\left\{\boldsymbol{\sigma}.[\mathbf{P} - q\mathbf{A}(\mathbf{R}, t)]\right\}^2 + qU(\mathbf{R}, t)$$

4. We intend to study the reflection of a monoenergetic neutron beam which is perpendicularly incident on a block of a ferromagnetic material. We call Ox the direction of propagation of the incident beam and yOz the surface of the ferromagnetic material, which fills the entire $x > 0$ region (see Figure 1). Let each incident neutron have an energy E and a mass m. The spin of the neutrons is $s = 1/2$ and their magnetic moment is written $\mathbf{M} = \gamma\mathbf{S}$ (γ is the gyromagnetic ratio and $\mathbf{S}$ is the spin operator).

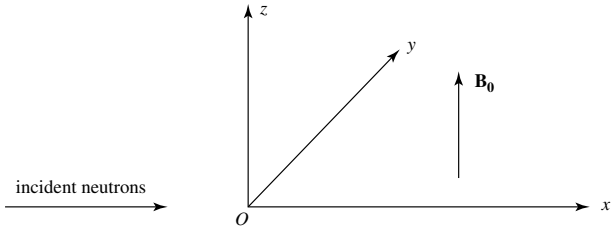


Figure 1

The potential energy of the neutrons is the sum of two terms:

- the first one corresponds to the interaction with the nucleons of the substance. Phenomenologically, it is represented by a potential $V(x)$, defined by $V(x) = 0$ for $x \leq 0$, $V(x) = V_0 > 0$ for $x > 0$.

- the second term corresponds to the interaction of the magnetic moment of each neutron with the internal magnetic field $\mathbf{B}_0$ of the material ($\mathbf{B}_0$ is assumed to be uniform and parallel to Oz). Thus we have $W = 0$ for $x \leq 0$, $W = \omega_0 S_z$ for $x > 0$ (with $\omega_0 = -\gamma B_0$). Throughout this exercise we shall confine ourselves to the case:

$$0 < \frac{\hbar\omega_0}{2} < V_0$$

a. Determine the stationary states of the particle that correspond to a positive incident momentum and a spin which is either parallel or antiparallel to Oz.

b. We assume in this question that $V_0 - \hbar\omega_0/2 < E < V_0 + \hbar\omega_0/2$. The incident neutron beam is unpolarized. Calculate the degree of polarization of the reflected beam. Can you imagine an application of this effect?

c. Now consider the general case where E has an arbitrary positive value. The spin of the incident neutrons points in the Ox direction. What is the direction of the spin of the reflected particles (there are three cases, depending on the relative values of E and $V_0 \pm \hbar\omega_0/2$)?

Solution of exercise 4

a. The Hamiltonian H of the particle is:

$$H = \frac{\mathbf{P}^2}{2m} + V(X) + W \tag{1}$$

$V(X)$, which acts only on the orbital variables, commutes with S_z. Since W is proportional to S_z, it also commutes with this operator. Furthermore, $V(X)$ commutes with P_y and P_z, as well as with W (obviously, since W acts only on the spin variables). We can therefore find a basis of eigenvectors common to H, S_z, P_y, P_z, which can be written:

$$\left|\varphi^{\pm}_{E,p_y,p_z}\right\rangle = \left|\varphi^{\pm}_{E}\right\rangle \otimes |p_y\rangle \otimes |p_z\rangle \otimes |\pm\rangle \tag{2}$$

with:

$$\begin{aligned}
\left|\varphi^{\pm}_{E}\right\rangle &\in \mathcal{E}_x \\
|p_y\rangle &\in \mathcal{E}_y\,;\ P_y\,|p_y\rangle = p_y\,|p_y\rangle \\
|p_z\rangle &\in \mathcal{E}_z\,;\ P_z\,|p_z\rangle = p_z\,|p_z\rangle \\
|\pm\rangle &\in \mathcal{E}_s\,;\ S_z\,|\pm\rangle = \pm\frac{\hbar}{2}\,|\pm\rangle
\end{aligned} \tag{3}$$

where the ket $\left|\varphi^{\pm}_{E}\right\rangle$ is a solution of the eigenvalue equation:

$$\left[\frac{P_x^2}{2m} + V(X) + \frac{1}{2m}\left(p_y^2 + p_z^2\right) \pm \frac{\hbar\omega_0}{2}\right]\left|\varphi^{\pm}_{E}\right\rangle = E\left|\varphi^{\pm}_{E}\right\rangle \tag{4}$$

We assume in the statement of the problem that the neutron beam is normally incident, so we can set $p_y = p_z = 0$. Let $\varphi^{\pm}_{E}(x) = \langle x \mid \varphi^{\pm}_{E}\rangle$ be the wave function associated with $\left|\varphi^{\pm}_{E}\right\rangle$; it satisfies the equation:

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \pm \frac{\hbar\omega_0}{2}\right]\varphi^{\pm}_{E}(x) = E\varphi^{\pm}_{E}(x) \tag{5}$$

Thus the problem is reduced to that of a classical one-dimensional "square well": reflection from a "potential step" (*cf.* Complement $\mathrm{H_I}$).

In the $x < 0$ region, $V(x)$ is zero and the total energy E (which is positive) is greater than the potential energy. We know in this case that the wave function is a superposition of imaginary oscillatory exponentials:

$$\varphi^{\pm}_{E}(x) = A_{\pm}\mathrm{e}^{ikx} + B_{\pm}\mathrm{e}^{-ikx} \qquad \text{if} \quad x < 0 \tag{6}$$

with:

$$k = \sqrt{\frac{2m}{\hbar^2}E} \tag{7}$$

$A_\pm$ gives the amplitude of the wave associated with an incident particle having a spin either parallel or antiparallel to Oz. $B_\pm$ gives the amplitude of the wave associated with a reflected particle for the same two spin directions.

In the $x > 0$ region, $V(x)$ is equal to V_0 and, depending on the relative values of E and $V_0 \pm \hbar\omega_0/2$, the wave functions can behave like oscillatory or damped exponentials. We shall consider three cases:

(*i*) If $E > V_0 + \hbar\omega_0/2$, we set:

$$k'_\pm = \sqrt{\frac{2m}{\hbar^2}\left(E - V_0 \mp \frac{\hbar\omega_0}{2}\right)} \tag{8}$$

and the transmitted wave behaves like an oscillatory exponential:

$$\varphi_E^\pm(x) = C_\pm e^{ik'_\pm x} \qquad \text{if} \quad x > 0 \tag{9}$$

Moreover, the continuity conditions for the wave function and its derivative imply [*cf.* Complement H_I, relations (13) and (14)]:

$$\frac{B_\pm}{A_\pm} = \frac{k - k'_\pm}{k + k'_\pm} \qquad \frac{C_\pm}{A_\pm} = \frac{2k}{k + k'_\pm} \tag{10}$$

(*ii*) If, on the other hand, $E < V_0 - \hbar\omega_0/2$, we must introduce the quantities $\rho_\pm$:

$$\rho_\pm = \sqrt{\frac{2m}{\hbar^2}\left(V_0 \pm \frac{\hbar\omega_0}{2} - E\right)} \tag{11}$$

and the wave in the $x > 0$ region is a real, damped exponential (evanescent wave):

$$\varphi_E^\pm(x) = D_\pm e^{-\rho_\pm x} \qquad \text{if} \quad x > 0 \tag{12}$$

with, in this case [*cf.* Complement H_I, equations (22) and (23)]:

$$\frac{B_\pm}{A_\pm} = \frac{k - i\rho_\pm}{k + i\rho_\pm}\ ; \qquad \frac{D_\pm}{A_\pm} = \frac{2k}{k + i\rho_\pm} \tag{13}$$

(*iii*) Finally, in the intermediate case $V_0 - \hbar\omega_0/2 < E < V_0 + \hbar\omega_0/2$, we have:

$$\varphi_E^+(x) = D_+ e^{-\rho_+ x} \qquad \text{if} \quad x > 0 \tag{14a}$$

$$\varphi_E^-(x) = C_- e^{-ik'_- x} \qquad \text{if} \quad x > 0 \tag{14b}$$

[definitions (8) and (11) of k'_- and ρ_+ are still valid]. Depending on the spin orientation, the wave is either a damped or an oscillatory exponential. We then have:

$$\frac{B_+}{A_+} = \frac{k - i\rho_+}{k + i\rho_+}\ ; \qquad \frac{D_+}{A_+} = \frac{2k}{k + i\rho_+} \tag{15a}$$

$$\frac{B_-}{A_-} = \frac{k - k'_-}{k + k'_-}\ ; \qquad \frac{C_-}{A_-} = \frac{2k}{k + k'_-} \tag{15b}$$

b. When $V_0 - \hbar\omega_0/2 < E < V_0 + \hbar\omega_0/2$, we are in the situation of case (*iii*) above. If the projection onto Oz of the incident neutron spin is equal to $\hbar/2$, the corresponding reflection coefficient is:

$$R_+ = \left|\frac{B_+}{A_+}\right|^2 = \left|\frac{k - i\rho_+}{k + i\rho_+}\right|^2 = 1 \tag{16}$$

On the other hand, if the projection of the spin onto Oz is equal to $-\hbar/2$, the reflection coefficient is no longer 1, since it is given by:

$$R_- = \left|\frac{B_-}{A_-}\right|^2 = \left(\frac{k - k'_-}{k + k'_-}\right)^2 < 1 \tag{17}$$

Thus we see how the reflected beam can be polarized since, depending on the direction of its spin, the neutron has a different probability of being reflected. An unpolarized incident beam can be considered to be formed of neutrons whose spins have a probability $1/2$ of being in the state $|+\rangle$ and a probability $1/2$ of being in the state $|-\rangle$. Taking (16) and (17) into account, we see that the probability that a particle of the reflected beam will have its spin in the state $|+\rangle$ is $1/(1+R_-)$, while for the state $|-\rangle$ it is $R_-/(1+R_-)$. Therefore, the degree of polarization of the reflected beam is:

$$T = \frac{1 - R_-}{1 + R_-} = \frac{2kk'_-}{k^2 + k'^2_-} \tag{18}$$

In practice, reflection from a saturated ferromagnetic substance is actually used in the laboratory to obtain beams of polarized neutrons. To increase the degree of polarization obtained, the beam is made to fall obliquely on the surface of the ferromagnetic mirror; thus, the theoretical results obtained here are not directly applicable. However, the principle of the experiment is the same. The ferromagnetic substance chosen is often cobalt. When cobalt is magnetized to saturation, one can obtain high degrees of polarization T ($T \gtrsim 80\%$). Note, furthermore, that the same neutron beam reflection device can serve as an "analyzer" as well as a "polarizer" for spin directions. This possibility has been exploited in precision measurements of the magnetic moment of the neutron.

c. Consider a neutron whose momentum, of magnitude $p = \hbar k$, is parallel to Ox. Assume that the projection $\langle S_x \rangle$ of its spin is equal to $\hbar/2$. Its state is [*cf.* Chap. IV, relation (A-20)]:

$$|\psi\rangle = |p\rangle \otimes \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle] \tag{19}$$

with:

$$\langle \mathbf{r}|p\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \mathrm{e}^{ipx/\hbar} \tag{20}$$

How can we construct a stationary state of the particle in which the incident wave has the form (19)? We simply have to consider the state:

$$|\psi_S\rangle = \frac{1}{\sqrt{2}}\left[\left|\varphi^+_{E,0,0}\right\rangle + \left|\varphi^-_{E,0,0}\right\rangle\right] \tag{21}$$

which is a linear combination of two eigenkets of H defined in (2), associated with the same eigenvalue $E = p^2/2m$. The part of the ket $|\psi_S\rangle$ which describes the reflected wave is then:

$$|-p\rangle \otimes \frac{1}{\sqrt{2}}[B_+ |+\rangle + B_- |-\rangle] \tag{22}$$

where B_+ and B_- are given, depending on the case, by (10), (13) or (15) (A_+ and A_- being replaced by 1). Let us calculate, for a state such as (22), the average value $\langle \mathbf{S} \rangle$. Since this state is a tensor product, the spin variables and the orbital variables are not correlated. Therefore, $\langle \mathbf{S} \rangle$ can easily be obtained from the spin state vector $B_+ |+\rangle + B_- |-\rangle$, which gives:

$$\langle S_x \rangle = \frac{\hbar}{2} \frac{B_+^* B_- + B_-^* B_+}{|B_+|^2 + |B_-|^2} \tag{23a}$$

$$\langle S_y \rangle = \frac{\hbar}{2} \frac{i(B_-^* B_+ - B_+^* B_-)}{|B_+|^2 + |B_-|^2} \tag{23b}$$

$$\langle S_z \rangle = \frac{\hbar}{2} \frac{|B_+|^2 - |B_-|^2}{|B_+|^2 + |B_-|^2} \tag{23c}$$

Three cases can then be distinguished:

(*i*) If $E > V_0 + \hbar\omega_0/2$, we see from (10) that B_+ and B_- are real. Formulas (23) then show that $\langle S_x \rangle$ and $\langle S_z \rangle$ are not zero but that $\langle S_y \rangle = 0$. Upon reflection of the neutron, the spin has thus undergone a rotation about Oy. Physically, it is the difference between the degrees of reflection of neutrons whose spin is parallel to Oz and those whose spin is antiparallel to Oz which explains why the $\langle S_z \rangle$ component becomes positive.

(*ii*) If $E < V_0 - \hbar\omega_0/2$, equations (13) show that B_+ and B_- are not real: they are two complex numbers having different phases but the same modulus. According to (23), we have, in this case, $\langle S_z \rangle = 0$ but $\langle S_x \rangle \neq 0$ and $\langle S_y \rangle \neq 0$. Upon reflection of the neutron, the spin thus undergoes a rotation about Oz. The physical origin of this rotation is the following: because of the existence of the evanescent wave, the neutron spends a certain time in the $x > 0$ region; the Larmor precession about $\mathbf{B}_0$ that it undergoes during this time accounts for the rotation of its spin.

(*iii*) If $V_0 - \hbar\omega_0/2 < E < V_0 + \hbar\omega_0/2$, B_+ is a complex number while B_- is a real number, and their moduli are different. None of the spin components, $\langle S_x \rangle$, $\langle S_y \rangle$ or $\langle S_z \rangle$, is then zero. This rotation of the spin upon reflection of the neutron is explained by a combination of the effects pointed out in (*i*) and (*ii*).

Chapter X

Addition of angular momenta

A Introduction 1015
A-1 Total angular momentum in classical mechanics 1015
A-2 The importance of total angular momentum in quantum mechanics 1016
B Addition of two spin 1/2's. Elementary method 1019
B-1 Statement of the problem 1019
B-2 The eigenvalues of S_z and their degrees of degeneracy 1021
B-3 Diagonalization of S^2 1022
B-4 Results: triplet and singlet 1024
C Addition of two arbitrary angular momenta. General method 1025
C-1 Review of the general theory of angular momentum 1025
C-2 Statement of the problem 1026
C-3 Eigenvalues of $\mathbf{J}^2$ and J_z 1029
C-4 Common eigenvectors of $\mathbf{J}^2$ and J_z 1033

A. Introduction

A-1. Total angular momentum in classical mechanics

Consider a system of N classical particles. The total angular momentum $\mathcal{L}$ of this system with respect to a fixed point O is the vector sum of the individual angular momenta of the N particles with respect to this point O :

$$\boldsymbol{\mathcal{L}} = \sum_{i=1}^{N} \boldsymbol{\mathcal{L}}_i \tag{A-1}$$

with:

$$\boldsymbol{\mathcal{L}}_i = \mathbf{r}_i \times \mathbf{p}_i \tag{A-2}$$

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

The time derivative of $\boldsymbol{\mathcal{L}}$ is equal to the moment with respect to O of the external forces. Consequently, when the external forces are zero (an isolated system) or all directed towards the same center, the total angular momentum of the system (with respect to any point in the first case and with respect to the center of force in the second one) is a constant of the motion. This is not the case for each of the individual angular momenta $\boldsymbol{\mathcal{L}}_i$ if there are internal forces, that is, if the various particles of the system interact.

We shall illustrate this point with an example. Consider a system composed of two particles, (1) and (2), subject to the same central force field (which can be created by a third particle assumed to be heavy enough to remain motionless at the origin). If these two particles exert no force on each other, their angular momenta $\boldsymbol{\mathcal{L}}_1$ and $\boldsymbol{\mathcal{L}}_2$ with respect to the center of force O are both constants of the motion. The only force then acting on particle (1), for example, is directed towards O; its moment with respect to this point is therefore zero, as is $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{L}}_1$. On the other hand, if particle (1) is also subject to a force due to the presence of particle (2), the moment with respect to O of this force is not generally zero, and, consequently, $\boldsymbol{\mathcal{L}}_1$ is no longer a constant of the motion. However, if the interaction between the two particles obeys the principle of action and reaction, the moment of the force exerted by (1) on (2) with respect to O exactly compensates that of the force exerted by (2) on (1): the total angular momentum $\boldsymbol{\mathcal{L}}$ is conserved over time.

Therefore, in a system of interacting particles, *only the total angular momentum is a constant of the motion*: forces inside the system induce a transfer of angular momentum from one particle to the other. Thus we see why it is useful to study the properties of the total angular momentum.

A-2. The importance of total angular momentum in quantum mechanics

Let us treat the preceding example quantum mechanically. In the case of two non-interacting particles, the Hamiltonian of the system is given simply, in the $\{|\mathbf{r}_1, \mathbf{r}_2\rangle\}$ representation:

$$H_0 = H_1 + H_2 \tag{A-3}$$

with:

$$\begin{aligned} H_1 &= -\frac{\hbar^2}{2\mu_1}\Delta_1 + V(r_1) \\ H_2 &= -\frac{\hbar^2}{2\mu_2}\Delta_2 + V(r_2) \end{aligned} \tag{A-4}$$

[μ_1 and μ_2 are the masses of the two particles, $V(r)$ is the central potential to which they are subject; Δ_1 and Δ_2 denote the Laplacian operators relative to the coordinates of particles (1) and (2) respectively]. We know from Chapter VII (§ A-2-a) that the three components of the operator $\mathbf{L}_1$ associated with the angular momentum $\mathcal{L}_1$ of particle (1) commute with H_1:

$$[\mathbf{L}_1, H_1] = 0 \tag{A-5}$$

Also, all observables relating to one of the particles commute with all those corresponding to the other one ; in particular:

$$[\mathbf{L}_1, H_2] = 0 \tag{A-6}$$

From (A-5) and (A-6), we see that the three components of $\mathbf{L}_1$ are constants of the motion. An analogous argument is obviously valid for $\mathbf{L}_2$.

Now assume that the two particles interact, and that the corresponding potential energy $v(|\mathbf{r}_1 - \mathbf{r}_2|)$ depends only on the distance between them $|\mathbf{r}_1 - \mathbf{r}_2|$[1]:

$$|\mathbf{r}_1 - \mathbf{r}_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{A-7}$$

In this case, the Hamiltonian of the system is:

$$H = H_1 + H_2 + v(|\mathbf{r}_1 - \mathbf{r}_2|) \tag{A-8}$$

where H_1 and H_2 are given by (A-4). According to (A-5) and (A-6), the commutator of $\mathbf{L}_1$ with H reduces to:

$$[\mathbf{L}_1,\, H] = [\mathbf{L}_1,\, v(|\mathbf{r}_1 - \mathbf{r}_2|)] \tag{A-9}$$

that is, for example for the component L_{1z}:

$$[L_{1z},\, H] = [L_{1z},\, v(|\mathbf{r}_1 - \mathbf{r}_2|)] = \frac{\hbar}{i}\left(x_1 \frac{\partial v}{\partial y_1} - y_1 \frac{\partial v}{\partial x_1}\right) \tag{A-10}$$

Expression (A-10) is generally not zero: $\mathbf{L}_1$ is no longer a constant of the motion. On the other hand, if we define the *total angular momentum operator* $\mathbf{L}$ by an expression similar to (A-1):

$$\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2 \tag{A-11}$$

we obtain an operator whose three components are constants of the motion. For example, we find:

$$[L_z,\, H] = [L_{1z} + L_{2z},\, H] \tag{A-12}$$

According to (A-10), this commutator is equal to:

$$\begin{aligned}[L_z,\, H] &= [L_{1z} + L_{2z},\, H] \\ &= \frac{\hbar}{i}\left(x_1 \frac{\partial v}{\partial y_1} - y_1 \frac{\partial v}{\partial x_1} + x_2 \frac{\partial v}{\partial y_2} - y_2 \frac{\partial v}{\partial x_2}\right)\end{aligned} \tag{A-13}$$

But, since v depends only on $|\mathbf{r}_1 - \mathbf{r}_2|$ given by (A-7), we have:

$$\frac{\partial v}{\partial x_1} = v' \frac{\partial |\mathbf{r}_1 - \mathbf{r}_2|}{\partial x_1} = v' \frac{x_1 - x_2}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{A-14a}$$

$$\frac{\partial v}{\partial x_2} = v' \frac{\partial |\mathbf{r}_1 - \mathbf{r}_2|}{\partial x_2} = v' \frac{x_2 - x_1}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{A-14b}$$

and analogous expressions for $\partial v/\partial y_1$, $\partial v/\partial y_2$, $\partial v/\partial z_1$ and $\partial v/\partial z_2$ (v' is the derivative of v, considered as a function of a single variable). Substituting these values into (A-13):

$$\begin{aligned}[L_z,\, H] &= \frac{\hbar}{i}\frac{v'}{|\mathbf{r}_1 - \mathbf{r}_2|}\{x_1(y_1 - y_2) - y_1(x_1 - x_2) + x_2(y_2 - y_1) - y_2(x_2 - x_1)\} \\ &= 0\end{aligned} \tag{A-15}$$

[1]The corresponding classical forces then necessarily obey the principle of action and reaction.

We therefore arrive at the same conclusion as in classical mechanics.

Until now we have implicitly assumed that the particles being studied had no spin. Now let us examine another important example: that of a single particle with spin. First, we assume that this particle is subject only to a central potential $V(r)$. Its Hamiltonian is then the one studied in § A of Chapter VII. We know that the three components of the orbital angular momentum $\mathbf{L}$ commute with this Hamiltonian. In addition, since the spin operators commute with the orbital observables, the three components of the spin $\mathbf{S}$ are also constants of the motion. But we shall see in Chapter XII that relativistic corrections introduce into the Hamiltonian a *spin-orbit coupling* term of the form:

$$H_{S0} = \xi(r)\mathbf{L} \cdot \mathbf{S} \tag{A-16}$$

where $\xi(r)$ is a known function of the single variable r (the physical meaning of this coupling will be explained in Chapter XII). When this term is taken into account, $\mathbf{L}$ and $\mathbf{S}$ no longer commute with the total Hamiltonian. For example[2]:

$$\begin{aligned}[L_z,\, H_{S0}] &= \xi(r)[L_z,\, L_xS_x + L_yS_y + L_zS_z] \\ &= \xi(r)(i\hbar L_yS_x - i\hbar L_xS_y)\end{aligned} \tag{A-17}$$

and, similarly:

$$\begin{aligned}[S_z,\, H_{S0}] &= \xi(r)[S_z,\, L_xS_x + L_yS_y + L_zS_z] \\ &= \xi(r)(i\hbar L_xS_y - i\hbar L_yS_x)\end{aligned} \tag{A-18}$$

However, if we set:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{A-19}$$

the three components of $\mathbf{J}$ *are constants of the motion.* To see this, we can simply add equations (A-17) and (A-18):

$$[J_z,\, H_{S0}] = [L_z + S_z,\, H_{S0}] = 0 \tag{A-20}$$

(an analogous proof could be given for the other components of $\mathbf{J}$). The operator $\mathbf{J}$ defined by (A-19) is said to be the total angular momentum of a particle with spin.

In the two cases just described, we have two partial angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$, which commute. We know a basis of the state space composed of eigen-vectors common to $\mathbf{J}_1^2,\ J_{1z},\ \mathbf{J}_2^2,\ J_{2z}$. However, $\mathbf{J}_1$ and $\mathbf{J}_2$ are not constants of the motion, while the components of the total angular momentum:

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \tag{A-21}$$

commute with the Hamiltonian of the system. We shall therefore try to construct, using the preceding basis, *a new basis formed by eigenvectors of* $\mathbf{J}^2$ *and* J_z. The problem thus posed in general terms is that of the *addition (or composition) of two angular momenta* $\mathbf{J}_1$ *and* $\mathbf{J}_2$.

The importance of this new basis, formed of eigenvectors of $\mathbf{J}^2$ and J_z, is easy to understand. To determine the stationary states of the system, that is, the eigenstates of

[2] To establish (A-17) and (A-18), one uses the fact that $\mathbf{L}$, which acts only on the angular variables θ and φ, commutes with $\xi(r)$, which depends only on r.

H, it is simpler to diagonalize the matrix which represents H in this new basis. Since H commutes with $\mathbf{J}^2$ and J_z, this matrix can be broken down into as many blocks as there are eigensubspaces associated with the various sets of eigenvalues of $\mathbf{J}^2$ and J_z (*cf.* Chap. II, § D-3-a). Its structure is much simpler than that of the matrix which represents H in the basis of eigenvectors common to $\mathbf{J}_1^2$, J_{1z}, $\mathbf{J}_2^2$, J_{2z}, since neither J_{1z} nor J_{2z} generally commutes with H.

We shall leave aside for now the problem of the diagonalization of H (whether exact or approximate) in the basis of eigenstates of $\mathbf{J}^2$ and J_z. Rather, we shall concentrate on the construction of this new basis from the one formed by the eigenstates of $\mathbf{J}_1^2$, J_{1z}, $\mathbf{J}_2^2$, J_{2z}. A certain number of physical applications (many-electron atoms, fine and hyperfine line structure, etc.) will be considered after we have studied perturbation theory (complements of Chapter XI and Chapter XII).

We shall begin (§ B) with an elementary treatment of a simple case, in which the two partial angular momenta we wish to add are spin 1/2's. This will allow us to familiarize ourselves with various aspects of the problem, before we treat, in § C, the addition of two arbitrary angular momenta.

B. Addition of two spin 1/2's. Elementary method

B-1. Statement of the problem

We shall consider a system of two spin 1/2 particles (electrons or silver atoms in the ground state, for example), and we shall be concerned only with their spin degrees of freedom. Let $\mathbf{S}_1$, and $\mathbf{S}_2$ be the spin operators of the two particles.

B-1-a. State space

We have already defined the state space of such a system. Recall that it is a four-dimensional space, obtained by taking the tensor product of the individual spin spaces of the two particles. We know an orthonormal basis of this space, which we shall denote by $\{|\varepsilon_1, \varepsilon_2\rangle\}$, that is, explicitly:

$$\{|\varepsilon_1, \varepsilon_2\rangle\} = \{|+, +\rangle, |+, -\rangle, |-, +\rangle, |-, -\rangle\} \tag{B-1}$$

These vectors are eigenstates of the four observables $\mathbf{S}_1^2$, S_{1z}, $\mathbf{S}_2^2$, S_{2z} (which are actually the extensions, into the tensor product space, of operators, defined in each of the spin spaces):

$$\mathbf{S}_1^2|\varepsilon_1, \varepsilon_2\rangle = \mathbf{S}_2^2|\varepsilon_1, \varepsilon_2\rangle = \frac{3}{4}\hbar^2|\varepsilon_1, \varepsilon_2\rangle \tag{B-2a}$$

$$S_{1z}|\varepsilon_1, \varepsilon_2\rangle = \varepsilon_1\frac{\hbar}{2}|\varepsilon_1, \varepsilon_2\rangle \tag{B-2b}$$

$$S_{2z}|\varepsilon_1, \varepsilon_2\rangle = \varepsilon_2\frac{\hbar}{2}|\varepsilon_1, \varepsilon_2\rangle \tag{B-2c}$$

$\mathbf{S}_1^2$, $\mathbf{S}_2^2$, S_{1z} and S_{2z} constitute a C.S.C.O. (the first two observables are actually multiples of the identity operator, and the set of operators remains complete even if they are omitted).

B-1-b. Total spin S. Commutation relations

We define the total spin $\mathbf{S}$ of the system by:

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \tag{B-3}$$

It is simple, knowing that $\mathbf{S}_1$ and $\mathbf{S}_2$ are angular momenta, to show that $\mathbf{S}$ is as well. We can calculate, for example, the commutator of S_x and S_y:

$$\begin{aligned}[S_x,\, S_y] &= [S_{1x} + S_{2x},\, S_{1y} + S_{2y}] \\ &= [S_{1x},\, S_{1y}] + [S_{2x},\, S_{2y}] \\ &= i\hbar S_{1z} + i\hbar S_{2z} \\ &= i\hbar S_z \end{aligned} \tag{B-4}$$

The operator $\mathbf{S}^2$ can be obtained by taking the (scalar) square of equation (B-3):

$$\mathbf{S}^2 = (\mathbf{S}_1 + \mathbf{S}_2)^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 \tag{B-5}$$

since $\mathbf{S}_1$ and $\mathbf{S}_2$ commute. The scalar product $\mathbf{S}_1 \cdot \mathbf{S}_2$ can be expressed in terms of the operators $S_{1\pm}$, S_{1z} and $S_{2\pm}$, S_{2z}; it is easy to show that:

$$\begin{aligned}\mathbf{S}_1 \cdot \mathbf{S}_2 &= S_{1x}S_{2x} + S_{1y}S_{2y} + S_{1z}S_{2z} \\ &= \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}) + S_{1z}S_{2z} \end{aligned} \tag{B-6}$$

Note that, since $\mathbf{S}_1$ and $\mathbf{S}_2$ each commute with $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$, so do the three components of $\mathbf{S}$. In particular, $\mathbf{S}^2$ and S_z commute with $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$:

$$[S_z,\, \mathbf{S}_1^2] = [S_z,\, \mathbf{S}_2^2] = 0 \tag{B-7a}$$

$$[\mathbf{S}^2,\, \mathbf{S}_1^2] = [\mathbf{S}^2,\, \mathbf{S}_2^2] = 0 \tag{B-7b}$$

In addition, S_z obviously commutes with S_{1z} and S_{2z}:

$$[S_z,\, S_{1z}] = [S_z,\, S_{2z}] = 0 \tag{B-8}$$

However, $\mathbf{S}^2$ *commutes with neither* S_{1z} *nor* S_{2z} since, according to (B-5):

$$\begin{aligned}[\mathbf{S}^2,\, S_{1z}] &= [\mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2,\, S_{1z}] \\ &= 2[\mathbf{S}_1 \cdot \mathbf{S}_2,\, S_{1z}] \\ &= 2[S_{1x}S_{2x} + S_{1y}S_{2y},\, S_{1z}] \\ &= 2i\hbar(-S_{1y}S_{2x} + S_{1x}S_{2y}) \end{aligned} \tag{B-9}$$

[this calculation is analogous to the one performed in (A-17) and (A-18)]. The commutator of $\mathbf{S}^2$ with S_{2z} is, of course, equal and opposite to the preceding one, so that $S_z = S_{1z} + S_{2z}$ commutes with $\mathbf{S}^2$.

B-1-c. The basis change to be performed

The basis (B-1), as we have seen, is composed of eigenvectors common to the C.S.C.O.:

$$\{\mathbf{S}_1^2,\, \mathbf{S}_2^2,\, S_{1z},\, S_{2z}\} \tag{B-10}$$

Also, we have just shown that the four observables:

$$\mathbf{S}_1^2,\, \mathbf{S}_2^2,\, \mathbf{S}^2,\, S_z \tag{B-11}$$

commute. We shall see in what follows that they also form a C.S.C.O.

Adding the two spins $\mathbf{S}_1$ and $\mathbf{S}_2$ amounts to constructing the orthonormal system of eigenvectors common to the set (B-11). This system will be different from (B-1), since $\mathbf{S}^2$ does not commute with S_{1z} and S_{2z}. We shall write the vectors of this new basis $|S,\, M\rangle$, with the eigenvalues of $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$ (which remain the same) implicit. The vectors $|S,\, M\rangle$ therefore satisfy the equations:

$$\mathbf{S}_1^2|S,\, M\rangle = \mathbf{S}_2^2|S,\, M\rangle = \frac{3}{4}\hbar^2|S,\, M\rangle \tag{B-12a}$$

$$\mathbf{S}^2|S,\, M\rangle = S(S+1)\hbar^2|S,\, M\rangle \tag{B-12b}$$

$$S_z|S,\, M\rangle = M\hbar|S,\, M\rangle \tag{B-12c}$$

We know that $\mathbf{S}$ is an angular momentum. Consequently, S must be a positive integer or half-integer, and M varies by one-unit jumps between $-S$ and $+S$. The problem is therefore to find what values S and M can actually have, and to express the basis vectors $\{|S,\, M\rangle\}$ in terms of those of the known basis.

In this section, we shall confine ourselves to solving this problem by the elementary method involving the calculation and diagonalization of the 4×4 matrices representing $\mathbf{S}^2$ and S_z in the $\{|\varepsilon_1,\, \varepsilon_2\rangle\}$ basis. In § C, we shall use another, more elegant, method, and generalize it to the case of two arbitrary angular momenta.

B-2. The eigenvalues of S_z and their degrees of degeneracy

The observables $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$ are easy to deal with: all vectors of the state space are their eigenvectors, with the same eigenvalue $3\hbar^2/4$. Consequently, equations (B-12a) are automatically satisfied for all kets $|S,\, M\rangle$.

We have already noted [formulas (B-7) and (B-8)] that S_z commutes with the four observables of the C.S.C.O. (B-10). We should therefore expect *the basis vectors* $\{|\varepsilon_1,\, \varepsilon_2\rangle\}$ *to be automatically eigenvectors of* S_z. We can indeed show, using (B-2b) and (B-2c), that:

$$S_z|\varepsilon_1,\, \varepsilon_2\rangle = (S_{1z} + S_{2z})|\varepsilon_1,\, \varepsilon_2\rangle = \frac{1}{2}(\varepsilon_1 + \varepsilon_2)\hbar|\varepsilon_1,\, \varepsilon_2\rangle \tag{B-13}$$

$|\varepsilon_1,\, \varepsilon_2\rangle$ is therefore an eigenstate of S_z with the eigenvalue:

$$M = \frac{1}{2}(\varepsilon_1 + \varepsilon_2) \tag{B-14}$$

Since ε_1 and ε_2 can each be equal to ± 1, we see that M *can take on the values* $+1$, 0 *and* -1.

The values $M = 1$ and $M = -1$ are not degenerate. Only one eigenvector corresponds to each of them: $|+, +\rangle$ for the first one and $|-, -\rangle$ for the second one. *On the other hand, $M = 0$ is two-fold degenerate*: two orthogonal eigenvectors are associated with it, $|+, -\rangle$ and $|-, +\rangle$. Any linear combination of these two vectors is an eigenstate of S_z, with the eigenvalue 0.

These results appear clearly in the matrix which represents S_z in the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis. Choosing the basis vectors in the order indicated in (B-1), that matrix can be written:

$$(S_z) = \hbar \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{B-15}$$

B-3. Diagonalization of $\mathbf{S}^2$

All that remains to be done is to find and then diagonalize the matrix which represents $\mathbf{S}^2$ in the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis. We know in advance that it is not diagonal, since $\mathbf{S}^2$ does not commute with S_{1z} and S_{2z}.

B-3-a. Calculation of the matrix representing $\mathbf{S}^2$

We are going to apply $\mathbf{S}^2$ to each of the basis vectors. To do this, we shall use formulas (B-5) and (B-6):

$$\mathbf{S}^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + 2S_{1z}S_{2z} + S_{1+}S_{2-} + S_{1-}S_{2+} \tag{B-16}$$

The four vectors $|\varepsilon_1, \varepsilon_2\rangle$ are eigenvectors of $\mathbf{S}_1^2$, $\mathbf{S}_2^2$, S_{1z} and S_{2z} [formulas (B-2)], and the action of the operators $S_{1\pm}$ and $S_{2\pm}$ can be derived from formulas (B-7) of Chapter IX. We therefore find:

$$\begin{aligned} \mathbf{S}^2|+, +\rangle &= \left(\frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2\right)|+, +\rangle + \frac{1}{2}\hbar^2|+, +\rangle \\ &= 2\hbar^2|+, +\rangle \end{aligned} \tag{B-17a}$$

$$\begin{aligned} \mathbf{S}^2|+, -\rangle &= \left(\frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2\right)|+, -\rangle - \frac{1}{2}\hbar^2|+, -\rangle + \hbar^2|-, +\rangle \\ &= \hbar^2[|+, -\rangle + |-, +\rangle] \end{aligned} \tag{B-17b}$$

$$\begin{aligned} \mathbf{S}^2|-, +\rangle &= \left(\frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2\right)|-, +\rangle - \frac{1}{2}\hbar^2|-, +\rangle + \hbar^2|+, -\rangle \\ &= \hbar^2[|-, +\rangle + |+, -\rangle] \end{aligned} \tag{B-17c}$$

$$\begin{aligned} \mathbf{S}^2|-, -\rangle &= \left(\frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2\right)|-, -\rangle + \frac{1}{2}\hbar^2|-, -\rangle \\ &= 2\hbar^2|-, -\rangle \end{aligned} \tag{B-17d}$$

The matrix representing $\mathbf{S}^2$ in the basis of the four vectors $|\varepsilon_1, \varepsilon_2\rangle$, arranged in

the order given in (B-1), is therefore:

$$(\mathbf{S}^2) = \hbar^2 \left(\begin{array}{c|cc|c} 2 & 0 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right) \tag{B-18}$$

Comment:

The zeros appearing in this matrix were to be expected. $\mathbf{S}^2$ commutes with S_z, and therefore has non-zero matrix elements only between eigenvectors of S_z associated with the same eigenvalue. According to the results of § 2, the only non-diagonal elements of $\mathbf{S}^2$ which could be different from zero are those which relate $|+,\,-\rangle$ to $|-,\,+\rangle$.

B-3-b. Eigenvalues and eigenvectors of $\mathbf{S}^2$

Matrix (B-18) can be broken down into three submatrices (as shown by the dotted lines). Two of them are one-dimensional: *the vectors* $|+,\,+\rangle$ *and* $|-,\,-\rangle$ *are eigenvectors of* $\mathbf{S}^2$, as is also shown by relations (B-17a) and (B-17d). The associated eigenvalues are both equal to $2\hbar^2$.

We must now diagonalize the 2×2 submatrix:

$$(\mathbf{S}^2)_0 = \hbar^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{B-19}$$

which represents $\mathbf{S}^2$ inside the two-dimensional subspace spanned by $|+,\,-\rangle$ and $|-,\,+\rangle$, that is, the eigensubspace of S_z corresponding to $M = 0$. The eigenvalues $\lambda\hbar^2$ of matrix (B-19) can be obtained by solving the characteristic equation:

$$(1-\lambda)^2 - 1 = 0 \tag{B-20}$$

The roots of this equation are $\lambda = 0$ and $\lambda = 2$. This yields the last two eigenvalues of $\mathbf{S}^2$: 0 and $2\hbar^2$. An elementary calculation yields the corresponding eigenvectors:

$$\frac{1}{\sqrt{2}}[|+,\,-\rangle + |-,\,+\rangle] \quad \text{for the eigenvalue } 2\hbar^2 \tag{B-21a}$$

$$\frac{1}{\sqrt{2}}[|+,\,-\rangle - |-,\,+\rangle] \quad \text{for the eigenvalue } 0 \tag{B-21b}$$

(of course, they are defined only to within a global phase factor; the coefficients $1/\sqrt{2}$ insure their normalization).

The operator $\mathbf{S}^2$ therefore possesses two distinct eigenvalues: 0 and $2\hbar^2$. The first one is non-degenerate and corresponds to vector (B-21b). The second one is three-fold degenerate, and the vectors $|+,\,+\rangle$, $|-,\,-\rangle$ and (B-21a) form an orthonormal basis in the associated eigensubspace.

B-4. Results: triplet and singlet

Thus we have obtained the eigenvalues of $\mathbf{S}^2$ and S_z, as well as a system of eigenvectors common to these two observables. We shall summarize these results by expressing them in the notation of equations (B-12).

The quantum number S of (B-12b) can take on two values: 0 and 1. The first one is associated with a single vector, (B-21b), which is also an eigenvector of S_z with the eigenvalue 0, since it is a linear combination of $|+,\,-\rangle$ and $|-,\,+\rangle$; we shall therefore denote this vector by $|0,\,0\rangle$:

$$|0,\,0\rangle = \frac{1}{\sqrt{2}}[|+,\,-\rangle - |-,\,+\rangle] \tag{B-22}$$

Three vectors which differ by their values of M are associated with the value $S = 1$:

$$\begin{cases} |1,\,1\rangle &= |+,\,+\rangle \\ |1,\,0\rangle &= \dfrac{1}{\sqrt{2}}[|+,\,-\rangle + |-,\,+\rangle] \\ |1,\,-1\rangle &= |-,\,-\rangle \end{cases} \tag{B-23}$$

It can easily be shown that the four vectors $|S,\,M\rangle$ given in (B-22) and (B-23) form an orthonormal basis. Specification of S and M suffices to define uniquely a vector of this basis. From this, it can be shown that $\mathbf{S}^2$ and S_z constitute a C.S.C.O. (which could include $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$, although it is not necessary here).

Therefore, *when two spin* 1/2's ($s_1 = s_2 = 1/2$) *are added, the number* S which characterizes the eigenvalues $S(S+1)\hbar^2$ of the observable $\mathbf{S}^2$ *can be equal either to* 1 *or to* 0. *With each of these two values of* S *is associated a family of* $(2S+1)$ *orthogonal vectors* (three for $S = 1$, one for $S = 0$) *corresponding to the* $(2S+1)$ *values of* M *which are compatible with* S.

Comments:

(*i*) The family (B-23) of the three vectors $|1,\,M\rangle$ ($M = 1,\,0,\,-1$) constitutes what is called a *triplet* ; the vector $|0,\,0\rangle$ is called a *singlet* state.

(*ii*) The triplet states are *symmetric* with respect to exchange of the two spins, whereas the singlet state is *antisymmetric*. This means that if each vector $|\varepsilon_1,\,\varepsilon_2\rangle$ is replaced by the vector $|\varepsilon_2,\,\varepsilon_1\rangle$, expressions (B-23) remain invariant, while (B-22) changes sign. We shall see in Chapter XIV the importance of this property when the two particles whose spins are added are identical. Furthermore, it enables us to find the right linear combination of $|+,\,-\rangle$ and $|-,\,+\rangle$ which must be associated with $|+,\,+\rangle$ and $|-,\,-\rangle$ (clearly symmetric) in order to complete the triplet. The singlet state, on the other hand, is the antisymmetric linear combination of $|+,\,-\rangle$ and $|-,\,+\rangle$, which is orthogonal to the preceding one.

C. Addition of two arbitrary angular momenta. General method

C-1. Review of the general theory of angular momentum

Consider an arbitrary system, whose state space is $\mathcal{E}$, and an angular momentum $\mathbf{J}$ relative to this system ($\mathbf{J}$ can be either a partial angular momentum or the total angular momentum of the system). We showed in Chapter VI (§ C-3) that it is always possible to construct a standard basis $\{|k,\, j,\, m\rangle\}$ composed of eigenvectors common to $\mathbf{J}^2$ and J_z:

$$\mathbf{J}^2|k,\, j,\, m\rangle = j(j+1)\hbar^2|k,\, j,\, m\rangle \tag{C-1a}$$

$$J_z|k,\, j,\, m\rangle = m\hbar|k,\, j,\, m\rangle \tag{C-1b}$$

such that the action of the operators J_+ and J_- obeys the relations:

$$J_\pm|k,\, j,\, m\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}|k,\, j,\, m\pm 1\rangle \tag{C-2}$$

We denote by $\mathcal{E}(k,\, j)$ the vector space spanned by the set of vectors of the standard basis which correspond to fixed values of k and j. There are $(2j+1)$ of these vectors, and, according to (C-1) and (C-2), they can be transformed into each other by $\mathbf{J}^2$, J_z, J_+ and J_-. The state space can be considered to be a direct sum of orthogonal subspaces $\mathcal{E}(k,\, j)$ which possess the following properties:

(i) $\mathcal{E}(k,\, j)$ is $(2j+1)$-dimensional.

(ii) $\mathcal{E}(k,\, j)$ is globally invariant under the action of $\mathbf{J}^2$, J_z, $J_\pm$, and, more generally, of any function $F(\mathbf{J})$. In other words, these operators have non-zero matrix elements only inside each of the subspaces $\mathcal{E}(k,\, j)$.

(iii) Inside a subspace $\mathcal{E}(k,\, j)$ the matrix elements of any function $F(\mathbf{J})$ of the angular momentum $\mathbf{J}$ are independent of k.

Comment:

As we pointed out in § C-3-a of Chapter VI, we can give the index k a concrete physical meaning by choosing for the standard basis the system of eigenvectors common to $\mathbf{J}^2$, J_z and one or several observables which commute with the three components of $\mathbf{J}$ and form a C.S.C.O. with $\mathbf{J}^2$ and J_z. If, for example:

$$[A,\, \mathbf{J}] = 0 \tag{C-3}$$

and if the set $\{A,\, \mathbf{J}^2,\, J_z\}$ is a C.S.C.O., we require the vectors $|k,\, j,\, m\rangle$ to be eigenvectors of A:

$$A|k,\, j,\, m\rangle = a_{k,j}|k,\, j,\, m\rangle \tag{C-4}$$

Relations (C-1), (C-2) and (C-4) determine the standard basis $\{|k,\, j,\, m\rangle\}$ in this case. Each of the $\mathcal{E}(k,\, j)$ is an eigensubspace of A, and the index k distinguishes between the various eigenvalues $a_{k,j}$ associated with each value of j.

C-2. Statement of the problem

C-2-a. State space

Consider a physical system formed by the union of two subsystems (for example, a two-particle system). We shall use indices 1 and 2 to label quantities relating to the two subsystems.

We shall assume that we know, in the state space $\mathcal{E}_1$ of subsystem (1), a standard basis $\{|k_1, j_1, m_1\rangle\}$ composed of common eigenvectors of $\mathbf{J}_1^2$ and J_{1z}, where $\mathbf{J}_1$ is the angular momentum operator of subsystem (1):

$$\mathbf{J}_1^2|k_1, j_1, m_1\rangle = j_1(j_1+1)\hbar^2|k_1, j_1, m_1\rangle \tag{C-5a}$$

$$J_{1z}|k_1, j_1, m_1\rangle = m_1\hbar|k_1, j_1, m_1\rangle \tag{C-5b}$$

$$J_{1\pm}|k_1, j_1, m_1\rangle = \hbar\sqrt{j_1(j_1+1) - m_1(m_1\pm 1)}|k_1, j_1, m_1\pm 1\rangle \tag{C-5c}$$

Similarly, we assume that the state space $\mathcal{E}_2$ of subsystem (2) is spanned by a standard basis $\{|k_2, j_2, m_2\rangle\}$:

$$\mathbf{J}_2^2|k_2, j_2, m_2\rangle = j_2(j_2+1)\hbar^2|k_2, j_2, m_2\rangle \tag{C-6a}$$

$$J_{2z}|k_2, j_2, m_2\rangle = m_2\hbar|k_2, j_2, m_2\rangle \tag{C-6b}$$

$$J_{2\pm}|k_2, j_2, m_2\rangle = \hbar\sqrt{j_2(j_2+1) - m_2(m_2\pm 1)}|k_2, j_2, m_2\pm 1\rangle \tag{C-6c}$$

The state space of the global system is the tensor product of $\mathcal{E}_1$ and $\mathcal{E}_2$:

$$\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2 \tag{C-7}$$

We know a basis of the global system, formed by taking the tensor product of the bases chosen in $\mathcal{E}_1$ and $\mathcal{E}_2$. We shall denote by $|k_1, k_2; j_1, j_2; m_1, m_2\rangle$ the vectors of this basis:

$$|k_1, k_2; j_1, j_2; m_1, m_2\rangle = |k_1, j_1, m_1\rangle \otimes |k_2, j_2, m_2\rangle \tag{C-8}$$

The spaces $\mathcal{E}_1$ and $\mathcal{E}_2$ can be considered to be the direct sums of the sub-spaces $\mathcal{E}_1(k_1, j_1)$ and $\mathcal{E}_2(k_2, j_2)$, which possess the properties recalled in § C-1:

$$\mathcal{E}_1 = \sum_{\oplus} \mathcal{E}_1(k_1, j_1) \tag{C-9a}$$

$$\mathcal{E}_2 = \sum_{\oplus} \mathcal{E}_2(k_2, j_2) \tag{C-9b}$$

Consequently, $\mathcal{E}$ is the direct sum of the subspaces $\mathcal{E}(k_1, k_2; j_1, j_2)$ obtained by taking the tensor product of a space $\mathcal{E}_1(k_1, j_1)$ and a space $\mathcal{E}_2(k_2, j_2)$:

$$\mathcal{E} = \sum_{\oplus} \mathcal{E}(k_1, k_2; j_1, j_2) \tag{C-10}$$

with:

$$\mathcal{E}(k_1, k_2; j_1, j_2) = \mathcal{E}_1(k_1, j_1) \otimes \mathcal{E}_2(k_2, j_2) \tag{C-11}$$

The dimension of the subspace $\mathcal{E}(k_1, k_2; j_1, j_2)$ is $(2j_1+1)(2j_2+1)$. This subspace is globally invariant under the action of any function of $\mathbf{J}_1$ and $\mathbf{J}_2$ ($\mathbf{J}_1$ and $\mathbf{J}_2$ here denote the extensions into $\mathcal{E}$ of the angular momentum operators originally defined in $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively).

C-2-b. Total angular momentum. Commutation relations

The total angular momentum of the system under consideration is defined by:

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \tag{C-12}$$

where $\mathbf{J}_1$ and $\mathbf{J}_2$, extensions of operators acting in the different spaces $\mathcal{E}_1$ and $\mathcal{E}_2$, commute. Of course, the components of $\mathbf{J}_1$, on the one hand, and of $\mathbf{J}_2$, on the other, satisfy the commutation relations that characterize angular momenta. It is easy to verify that the components of $\mathbf{J}$ also satisfy such relations [the calculation is the same as in (B-4)].

Since $\mathbf{J}_1$ and $\mathbf{J}_2$ each commute with $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$, so does $\mathbf{J}$. In particular, $\mathbf{J}^2$ and J_z commute with $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$:

$$[J_z,\, \mathbf{J}_1^2] = [J_z,\, \mathbf{J}_2^2] = 0 \tag{C-13a}$$

$$[\mathbf{J}^2,\, \mathbf{J}_1^2] = [\mathbf{J}^2,\, \mathbf{J}_2^2] = 0 \tag{C-13b}$$

Furthermore, J_{1z} and J_{2z} obviously commute with J_z :

$$[J_{1z},\, J_z] = [J_{2z},\, J_z] = 0 \tag{C-14}$$

but not with $\mathbf{J}^2$ since this last operator can be written in terms of $\mathbf{J}_1$ and $\mathbf{J}_2$ in the form:

$$\mathbf{J}^2 = \mathbf{J}_1^2 + \mathbf{J}_2^2 + 2\mathbf{J}_1 \cdot \mathbf{J}_2 \tag{C-15}$$

and, as in (B-9), J_{1z} and J_{2z} do not commute with $\mathbf{J}_1 \cdot \mathbf{J}_2$. We can also transform the expression for $\mathbf{J}^2$ into:

$$\mathbf{J}^2 = \mathbf{J}_1^2 + \mathbf{J}_2^2 + 2J_{1z}J_{2z} + J_{1+}J_{2-} + J_{1-}J_{2+} \tag{C-16}$$

C-2-c. The basis change to be performed

A vector $|k_1,\, k_2;\, j_1,\, j_2;\, m_1,\, m_2\rangle$ of basis (C-8) is a simultaneous eigenstate of the observables:

$$\mathbf{J}_1^2,\, \mathbf{J}_2^2,\, J_{1z},\, J_{2z} \tag{C-17}$$

with the respective eigenvalues $j_1(j_1+1)\hbar^2$, $j_2(j_2+1)\hbar^2$, $m_1\hbar$, $m_2\hbar$. *Basis* (C-8) *is well adapted to the study of the individual angular momenta* $\mathbf{J}_1$ *and* $\mathbf{J}_2$ of the two subsystems.

According to (C-13), the observables:

$$\mathbf{J}_1^2,\, \mathbf{J}_2^2,\, J^2,\, J_z \tag{C-18}$$

also commute. We are going to construct an orthonormal system of common eigenvectors of these observables: *this new basis will be well adapted to the study of the total angular momentum* of the system. Note that this basis will be different from the preceding one, since $\mathbf{J}^2$ does not commute with J_{1z} and J_{2z} (§ C-2-b above).

Comment:

To give a physical meaning to the two indices k_1 and k_2, let us assume (comment of § C-1) that we know, in $\mathcal{E}_1$, a C.S.C.O., $\{A_1,\, \mathbf{J}_1^2,\, J_{1z}\}$ where A_1 commutes with the three components of $\mathbf{J}_1$,

and, in $\mathcal{E}_2$, a C.S.C.O., $\{A_2, \mathbf{J}_2^2, J_{2z}\}$ where A_2 commutes with the three components of $\mathbf{J}_2$. We can choose for a standard basis $\{|k_1, j_1, m_1\rangle\}$ the orthonormal system of eigenvectors common to A_1, $\mathbf{J}_1^2$ and J_{1z}, and for $\{|k_2, j_2, m_2\rangle\}$ the orthonormal system of eigenvectors common to A_2, $\mathbf{J}_2^2$ and J_{2z}. The set:

$$\{A_1, A_2; \mathbf{J}_1^2, \mathbf{J}_2^2; J_{1z}, J_{2z}\} \tag{C-19}$$

then constitutes a C.S.C.O. in $\mathcal{E}$, whose eigenvectors are the kets (C-8). Since the observable A_1 commutes separately with the components of $\mathbf{J}_1$, and with those of $\mathbf{J}_2$, it also commutes with $\mathbf{J}$ and, in particular, with $\mathbf{J}^2$ and J_z. The same is, of course, true of A_2. Consequently, the observables:

$$A_1, A_2, \mathbf{J}_1^2, \mathbf{J}_2^2, \mathbf{J}^2, J_z \tag{C-20}$$

commute. We shall see that they in fact form a C.S.C.O.: the new basis we are trying to find is the orthonormal system of eigenvectors of this C.S.C.O.

The subspace $\mathcal{E}(k_1, k_2; j_1, j_2)$ of $\mathcal{E}$ defined in (C-11) is globally invariant under the action of any operator which is a function of $\mathbf{J}_1$ and $\mathbf{J}_2$, and, therefore, under the action of any function of the total angular momentum $\mathbf{J}$. It follows that the observables $\mathbf{J}^2$ and J_z, which we want to diagonalize, have non-zero matrix elements only between vectors belonging to the same subspace $(k_1, k_2; j_1, j_2)$. The matrices (which are, in general, infinite) representing $\mathbf{J}^2$ and J_z in the basis (C-8) are "block diagonal", that is, they can be broken down into a series of submatrices, each of which corresponds to a particular subspace $\mathcal{E}(k_1, k_2; j_1, j_2)$. *The problem therefore reduces to a change of basis inside each of the subspaces* $\mathcal{E}(k_1, k_2; j_1, j_2)$, which are of finite dimension $(2j_1 + 1)(2j_2 + 1)$.

Moreover, the matrix elements in the basis (C-8) of any function of $\mathbf{J}_1$ and $\mathbf{J}_2$ are independent of k_1 and k_2. This is therefore true of those of $\mathbf{J}^2$ and J_z. Consequently, *the problem of the diagonalization of* $\mathbf{J}^2$ *and* J_z *is the same inside all the subspaces* $\mathcal{E}(k_1, k_2; j_1, j_2)$ *that correspond to the same values* j_1 *and* j_2. It is for this reason that one usually speaks of *adding angular momenta* j_1 *and* j_2 without specifying the other quantum numbers. To simplify the notation, we shall henceforth omit the indices k_1 and k_2. We shall denote by $\mathcal{E}(j_1, j_2)$ the subspace $\mathcal{E}(k_1, k_2; j_1, j_2)$ and by $|j_1, j_2; m_1, m_2\rangle$, the vectors of basis (C-8) belonging to this subspace:

$$\mathcal{E}(j_1, j_2) \equiv \mathcal{E}(k_1, k_2; j_1, j_2) \tag{C-21a}$$

$$|j_1, j_2; m_1, m_2\rangle \equiv |k_1, k_2; j_1, j_2; m_1, m_2\rangle \tag{C-21b}$$

Since $\mathbf{J}$ is an angular momentum and $\mathcal{E}(j_1, j_2)$ is globally invariant under the action of any function of $\mathbf{J}$, the results of Chapter VI recalled above (§ C-1) are applicable. Consequently, $\mathcal{E}(j_1, j_2)$ is a direct sum of orthogonal subspaces $\mathcal{E}(k, J)$, each of which is globally invariant under the action of $\mathbf{J}^2$, J_z, J_+ and J_-:

$$\mathcal{E}(j_1, j_2) = \sum_{\oplus} \mathcal{E}(k, J) \tag{C-22}$$

Thus, finally, we are left with the following double problem:

(*i*) Given j_1 and j_2, *what are the values of* J which appear in (C-22), and *how many distinct subspaces* $\mathcal{E}(k, J)$ are associated with each of them?

(*ii*) *How can the eigenvectors of* $\mathbf{J}^2$ *and* J_z belonging to $\mathcal{E}(j_1, j_2)$ *be expanded on the* $\{|j_1, j_2; m_1, m_2\rangle\}$ *basis*?

§ C-3 supplies the answer to the first question, and § C-4 to the second.

Comments:

(*i*) We have introduced $\mathbf{J}_1$ and $\mathbf{J}_2$ as the angular momenta of two distinct subspaces. In fact, we know (§ A-2) that we may be adding the orbital and spin angular momenta of the same particle. All the discussions and results of this section are applicable to this case, with $\mathcal{E}_1$ and $\mathcal{E}_2$ simply being replaced by $\mathcal{E}_\mathbf{r}$ and $\mathcal{E}_s$.

(*ii*) In order to add several angular momenta, one first adds the first two, then the angular momentum so obtained to the third one, and so on until the last one has been added.

C-3. Eigenvalues of $\mathbf{J}^2$ and J_z

C-3-a. Special case of two spin 1/2's

First of all, let us again take up the simple problem treated in § B. The spaces $\mathcal{E}_1$ and $\mathcal{E}_2$ each contain, in this case, a single invariant subspace, and the tensor product space $\mathcal{E}$, a single subspace $\mathcal{E}(j_1, j_2)$, for which $j_1 = j_2 = 1/2$.

The results recalled in § C-1 make it very simple to find the values of the quantum number S associated with the total spin. The space $\mathcal{E} = \mathcal{E}(1/2, 1/2)$ must be a direct sum of $(2S+1)$-dimensional subspaces $\mathcal{E}(k, S)$. Each of these subspaces contains one and only one eigenvector of S_z corresponding to each of the values of M such that $|M| \leq S$. Now, we know (*cf.* § B-2) that the only values taken on by M are 1, -1 and 0, the first two being non-degenerate and the third, two-fold degenerate. From this, the following conclusions can be deduced directly:

(*i*) Values of S greater than 1 are excluded. For example, for $S = 2$ to be possible there would have to exist at least one eigenvector of S_z of eigenvalue $2\hbar$.

(*ii*) $S = 1$ occurs (since $M = 1$ does) only once: $M = 1$ is not degenerate.

(*iii*) This is also true for $S = 0$. The subspace characterized by $S = 1$ includes only one vector for which $M = 0$, and this value of M is doubly degenerate in the space $\mathcal{E}(1/2, 1/2)$.

The four-dimensional space $\mathcal{E}(1/2, 1/2)$ can therefore be broken down into a subspace associated with $S = 1$ (which is three-dimensional) and a subspace associated with $S = 0$ (which is one-dimensional).

Using a completely analogous argument, we shall determine the possible values of J in the general case in which j_1 and j_2 are arbitrary.

C-3-b. The eigenvalues of J_z and their degrees of degeneracy

In accordance with the conclusions of § C-2-c, we shall consider a well-defined subspace $\mathcal{E}(j_1, j_2)$, of dimension $(2j_1 + 1)(2j_2 + 1)$. We shall assume that j_1 and j_2 are labeled such that:

$$j_1 \geq j_2 \tag{C-23}$$

The vectors $|j_1, j_2; m_1, m_2\rangle$ are already eigenstates of J_z:

$$\begin{aligned} J_z|j_1, j_2; m_1, m_2\rangle &= (J_{1z} + J_{2z})|j_1, j_2; m_1, m_2\rangle \\ &= (m_1 + m_2)\hbar|j_1, j_2; m_1, m_2\rangle \end{aligned} \tag{C-24}$$

and the corresponding eigenvalues $M\hbar$ are such that:

$$M = m_1 + m_2 \tag{C-25}$$

Consequently, M takes on the following values:

$$j_1 + j_2\ ,\quad j_1 + j_2 - 1\ ,\quad j_1 + j_2 - 2\ \ ,...,-(j_1 + j_2) \tag{C-26}$$

To find the degree of degeneracy $g_{j_1,j_2}(M)$ of these values, we can use the following geometrical procedure. In a two-dimensional diagram, we associate with each vector $|j_1, j_2; m_1, m_2\rangle$ the point whose abscissa is m_1 and whose ordinate is m_2. All these points are situated inside, or on the sides of, the rectangle whose corners are at (j_1, j_2), $(j_1, -j_2)$, $(-j_1, -j_2)$ and $(-j_1, j_2)$. Figure 1 represents the 15 points associated with the basis vectors in the case in which $j_1 = 2$ and $j_2 = 1$ (the values of m_1 and m_2 are shown beside each point). All points situated on the same dashed line (of slope -1) correspond to the same value of $M = m_1 + m_2$. The number of such points is therefore equal to the degeneracy $g_{j_1,j_2}(M)$ of this value of M.

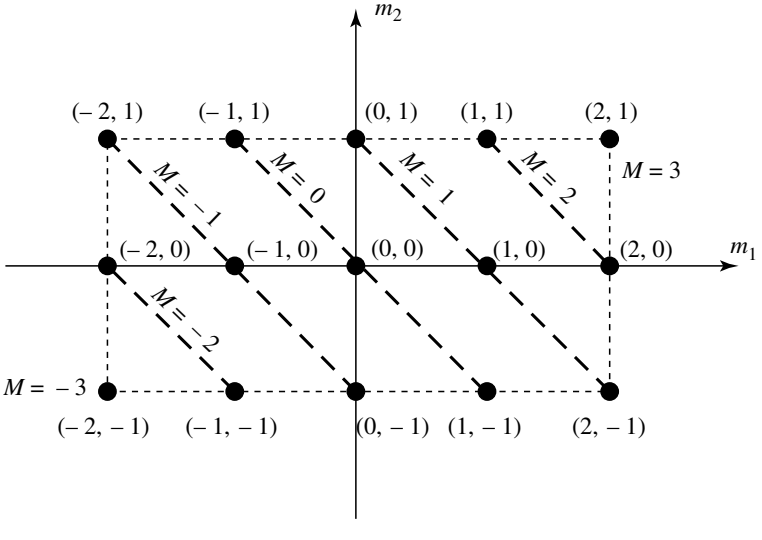


Figure 1: Pairs of possible values $(m_1,\ m_2)$ for the kets $|j_1, j_2; m_1, m_2\rangle$. We have chosen the case in which $j_1 = 2$ and $j_2 = 1$. The points associated with a given value of $M = m_1 + m_2$ are situated on a straight line of slope -1 (dashed lines).

Now consider the various values of M, in decreasing order, tracing the line defined by each of them (Fig. 1). $M = j_1 + j_2$ is not degenerate, since the line it characterizes

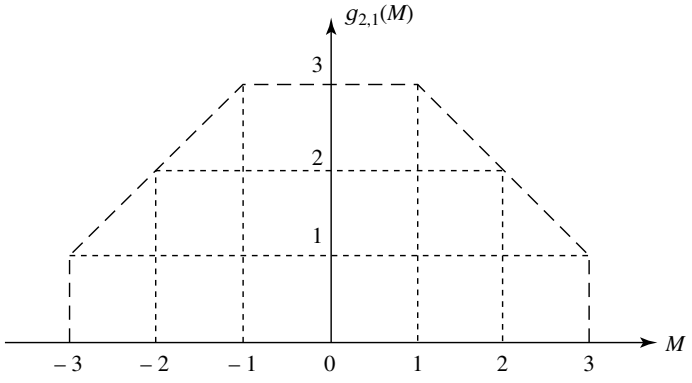


Figure 2: Value of the degree of degeneracy $g_{j_1,j_2}(M)$ as a fonction of M. As in Figure 1, we have shown the case in which $j_1 = 2$ and $j_2 = 1$. The degree of degeneracy $g_{j_1,j_2}(M)$ is simply obtained by counting the number of points on the corresponding dashed line of Figure 1.

---

passes only through the upper right-hand corner, whose coordinates are (j_1, j_2):

$$g_{j_1,j_2}(j_1 + j_2) = 1 \tag{C-27}$$

$M = j_1 + j_2 - 1$ is doubly degenerate, since the corresponding line contains the points $(j_1, j_2 - 1)$ and $(j_1 - 1, j_2)$:

$$g_{j_1,j_2}(j_1 + j_2 - 1) = 2 \tag{C-28}$$

The degree of degeneracy thus increases by one when M decreases by one, until we reach the lower right-hand corner of the rectangle $(m_1 = j_1, m_2 = -j_2)$, that is, the value $M = j_1 - j_2$. The number of points on the line is then at a maximum and is equal to:

$$g_{j_1,j_2}(j_1 - j_2) = 2j_2 + 1 \tag{C-29}$$

When M falls below $j_1 - j_2$, $g_{j_1,j_2}(M)$ first remains constant and equal to its maximum value as long as the line associated with M cuts across the entire width of the rectangle, that is, until it passes through the upper left-hand corner of the rectangle $(m_1 = -j_1, m_2 = j_2)$:

$$g_{j_1,j_2}(M) = 2j_2 + 1 \quad \text{for} \quad -(j_1 - j_2) \leq M \leq j_1 - j_2 \tag{C-30}$$

Finally, for M less than $-(j_1 - j_2)$, the corresponding line no longer intersects with the upper horizontal side of the rectangle, and $g_{j_1,j_2}(M)$ steadily decreases by one each time M decreases by one, again reaching 1 when $M = -(j_1 + j_2)$ (lower left-hand corner of the rectangle). Consequently:

$$g_{j_1,j_2}(-M) = g_{j_1,j_2}(M) \tag{C-31}$$

These results are summarized, for $j_1 = 2$ and $j_2 = 1$, in Figure 2, which gives $g_{2,1}(M)$ as a function of M.

C-3-c. The eigenvalues of $\mathbf{J}^2$

Note, first of all, that the values (C-26) of M are all integral if j_1 and j_2 are both integral or both half-integral, and all half-integral if one of them is integral and the other half-integral. Consequently, the corresponding values of J will also be all integral in the first case and all half-integral in the second.

Since the maximum value attained by M is $j_1 + j_2$, none of the values of J greater than $j_1 + j_2$ is found in $\mathcal{E}(j_1, j_2)$ and therefore none appears in the direct sum (C-22). With $J = j_1 + j_2$ is associated one invariant subspace (since $M = j_1 + j_2$ exists) and only one (since $M = j_1 + j_2$ is not degenerate). In this subspace $\mathcal{E}(J = j_1 + j_2)$, there is one and only one vector which corresponds to $M = j_1 + j_2 - 1$; now this value of M is two-fold degenerate in $\mathcal{E}(j_1, j_2)$; therefore, $J = j_1 + j_2 - 1$ also occurs, and to it corresponds a single invariant subspace $\mathcal{E}(J = j_1 + j_2 - 1)$.

More generally, we shall denote by $p_{j_1,j_2}(J)$ the number of subspaces $\mathcal{E}(k, J)$ of $\mathcal{E}(j_1, j_2)$ associated with a given value of J, that is, the number of different values of k for this value of J (j_1 and j_2 having been fixed at the beginning). $p_{j_1,j_2}(J)$ and $g_{j_1,j_2}(M)$ are very simply related. Consider a particular value of M. To it corresponds one and only one vector in each subspace $\mathcal{E}(k, J)$ such that $J \geq |M|$. Its degree of degeneracy $g_{j_1,j_2}(M)$ in $\mathcal{E}(j_1, j_2)$ can therefore be written:

$$\begin{aligned} g_{j_1,j_2}(M) &= p_{j_1,j_2}(J = |M|) + p_{j_1,j_2}(J = |M| + 1) \\ &\quad + p_{j_1,j_2}(J = |M| + 2) + \ldots \end{aligned} \tag{C-32}$$

Inverting, we obtain $p_{j_1,j_2}(J)$ in terms of $g_{j_1,j_2}(M)$:

$$\begin{aligned} p_{j_1,j_2}(J) &= g_{j_1,j_2}(M = J) - g_{j_1,j_2}(M = J + 1) \\ &= g_{j_1,j_2}(M = -J) - g_{j_1,j_2}(M = -J - 1) \end{aligned} \tag{C-33}$$

The results of § C-3-b then enable us to determine simply the values of the quantum number J which actually occur in $\mathcal{E}(j_1, j_2)$ and the number of invariant subspaces $\mathcal{E}(k, J)$ which are associated with them. First of all, we obviously have:

$$p_{j_1,j_2}(J) = 0 \quad \text{for} \quad J > j_1 + j_2 \tag{C-34}$$

since $g_{j_1,j_2}(M)$ is zero for $|M| > j_1 + j_2$. Furthermore, according to (C-27) and (C-28):

$$p_{j_1,j_2}(J = j_1 + j_2) = g_{j_1,j_2}(M = j_1 + j_2) = 1 \tag{C-35a}$$

$$p_{j_1,j_2}(J = j_1 + j_2 - 1) = g_{j_1,j_2}(M = j_1 + j_2 - 1) - g_{j_1,j_2}(M = j_1 + j_2) = 1 \tag{C-35b}$$

Thus, by iteration, we find all the values of $p_{j_1,j_2}(J)$:

$$p_{j_1,j_2}(J = j_1 + j_2 - 2) = 1, \ \ldots, \tag{C-36a}$$

$$\ldots, \ p_{j_1,j_2}(J = j_1 - j_2) = 1 \tag{C-36b}$$

and, finally, according to (C-30):

$$p_{j_1,j_2}(J) = 0 \quad \text{for} \quad J < j_1 - j_2 \tag{C-37}$$

Therefore, for fixed j_1 and j_2 that is, inside a given space $\mathcal{E}(j_1, j_2)$, the eigenvalues of $\mathbf{J}^2$ are such that[3]:

$$J = j_1 + j_2 \quad , \quad j_1 + j_2 - 1 \quad , \quad j_1 + j_2 - 2 \ \ , \ldots, \ |j_1 - j_2| \tag{C-38}$$

[3]Thus far, we have assumed $j_1 \geq j_2$, but it is simple to extend the discussion to the opposite case $j_1 < j_2$: all we need to do is invert indices 1 and 2.

With each of these values is associated *a single* invariant subspace $\mathcal{E}(J)$, so that the index k which appears in (C-22) is actually unnecessary. This means, in particular, that if we fix a value of J belonging to the set (C-38) and a value of M which is compatible with it, there corresponds to them one and only one vector in $\mathcal{E}(j_1, j_2)$: the specification of J suffices for the determination of the subspace $\mathcal{E}(J)$, in which the specification of M then defines one and only one vector. In other words, $\mathbf{J}^2$ and J_z form a C.S.C.O. in $\mathcal{E}(j_1, j_2)$.

Comment:

It can be shown that the number of pairs (J, M) found in $\mathcal{E}(j_1, j_2)$ is indeed equal to the dimension $(2j_1 + 1)(2j_2 + 1)$ of this space. This number (if, for example, $j_1 \geq j_2$) is equal to:

$$\sum_{J=j_1-j_2}^{j_1+j_2} (2J + 1) \tag{C-39}$$

If we set:

$$J = j_1 - j_2 + i \tag{C-40}$$

it is easy to calculate the sum (C-39):

$$\begin{aligned} \sum_{J=j_1-j_2}^{j_1+j_2} (2J + 1) &= \sum_{i=0}^{2j_2} [2(j_1 - j_2 + i) + 1] \\ &= [2(j_1 - j_2) + 1](2j_2 + 1) + 2\frac{2j_2(2j_2 + 1)}{2} \\ &= (2j_2 + 1)(2j_1 + 1) \end{aligned} \tag{C-41}$$

C-4. Common eigenvectors of $\mathbf{J}^2$ and J_z

We shall denote by $|J, M\rangle$ the common eigenvectors of $\mathbf{J}^2$ and J_z belonging to the space $\mathcal{E}(j_1, j_2)$. To be completely rigorous, we should have to recall the values of j_1 and j_2 in this notation, but we shall not write them explicitly, since they are the same as in the vectors (C-21b) of which the $|J, M\rangle$ are linear combinations. Of course, the indices J and M refer to the eigenvalues of $\mathbf{J}^2$ and J_z:

$$\mathbf{J}^2|J, M\rangle = J(J + 1)\hbar^2|J, M\rangle \tag{C-42a}$$

$$J_z|J, M\rangle = M\hbar|J, M\rangle \tag{C-42b}$$

and the vectors $|J, M\rangle$, like all those of the space $\mathcal{E}(j_1, j_2)$, are eigenvectors of $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$ with eigenvalues $j_1(j_1 + 1)\hbar^2$ and $j_2(j_2 + 1)\hbar^2$ respectively.

C-4-a. Special case of two spin 1/2's

First of all, we shall show how use of the general results concerning angular momenta leads us to the expression for the vectors $|S, M\rangle$ established in § B-3. It will not be necessary to diagonalize the matrix which represents $\mathbf{S}^2$. By generalizing this method, we shall then construct (§ 4-b) the vectors $|J, M\rangle$ for the case of arbitrary j_1 and j_2.

α. The subspace $\mathcal{E}(S = 1)$

The ket $|+, +\rangle$ is, in the state space $\mathcal{E} = \mathcal{E}(1/2, 1/2)$, the only eigenvector of S_z associated with $M = 1$. Since $\mathbf{S}^2$ and S_z commute, and the value $M = 1$ is not degenerate, $|+, +\rangle$ must also be an eigenvector of $\mathbf{S}^2$ (§ D-3-a of Chapter II). According to the reasoning of § C-3-a, the corresponding value of S must be 1. Therefore, we can choose the phase of the vector $|S = 1, M = 1\rangle$ such that:

$$|1, 1\rangle = |+, +\rangle \tag{C-43}$$

It is then easy to find the other states of the triplet, since we know from the general theory of angular momentum that:

$$\begin{aligned} S_-|1, 1\rangle &= \hbar\sqrt{1(1+1) - 1(1-1)}|1, 0\rangle \\ &= \hbar\sqrt{2}|1, 0\rangle \end{aligned} \tag{C-44}$$

Consequently:

$$|1, 0\rangle = \frac{1}{\hbar\sqrt{2}} S_-|+, +\rangle \tag{C-45}$$

To calculate $|1, 0\rangle$ explicitly in the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis, it suffices to recall that definition (B-3) of the total spin $\mathbf{S}$ implies:

$$S_- = S_{1-} + S_{2-} \tag{C-46}$$

We then obtain:

$$\begin{aligned} |1, 0\rangle &= \frac{1}{\hbar\sqrt{2}}(S_{1-} + S_{2-})|+, +\rangle \\ &= \frac{1}{\hbar\sqrt{2}}[\hbar|-, +\rangle + \hbar|+, -\rangle] \\ &= \frac{1}{\sqrt{2}}[|-, +\rangle + |+, -\rangle] \end{aligned} \tag{C-47}$$

Finally, we can again apply S_- to $|1, 0\rangle$, that is, $(S_{1-} + S_{2-})$ to expression (C-47). This yields:

$$\begin{aligned} |1, -1\rangle &= \frac{1}{\hbar\sqrt{2}} S_-|1, 0\rangle \\ &= \frac{1}{\hbar\sqrt{2}}(S_{1-} + S_{2-})\frac{1}{\sqrt{2}}[|-, +\rangle + |+, -\rangle] \\ &= \frac{1}{2\hbar}[\hbar|-, -\rangle + \hbar|-, -\rangle] \\ &= |-, -\rangle \end{aligned} \tag{C-48}$$

Of course, this last result could have been obtained directly, using an argument analogous to the one applied above to $|+, +\rangle$. However, the preceding calculation has a slight advantage: it enables us, in accordance with the general conventions set forth in § C-3-a of Chapter VI, to fix the phase factors which could appear in $|1, 0\rangle$ and $|1, -1\rangle$ with respect to the one chosen for $|1, 1\rangle$ in (C-43).

β. The state $|S = 0,\ M = 0\rangle$

The only vector $|S = 0,\ M = 0\rangle$ of the subspace $\mathcal{E}(S = 0)$ is determined, to within a constant factor, by the condition that it must be orthogonal to the three vectors $|1,\ M\rangle$ that we have just constructed.

Since it is orthogonal to $|1,\ 1\rangle = |+,\ +\rangle$ and $|1,\ -1\rangle = |-,\ -\rangle$, $|0,\ 0\rangle$ must be a linear combination of $|+,\ -\rangle$ and $|-,\ +\rangle$:

$$|0,\ 0\rangle = \alpha|+,\ -\rangle + \beta|-,\ +\rangle \tag{C-49}$$

which will be normalized if:

$$\langle 0,\ 0|0,\ 0\rangle = |\alpha|^2 + |\beta|^2 = 1 \tag{C-50}$$

We now insist that its scalar product with $|1,\ 0\rangle$ [*cf.* (C-47)] be zero:

$$\frac{1}{\sqrt{2}}(\alpha + \beta) = 0 \tag{C-51}$$

The coefficients α and β are therefore equal in absolute value and of opposite sign. With (C-50) taken into account, this fixes them to within a phase factor:

$$\alpha = -\beta = \frac{1}{\sqrt{2}}\mathrm{e}^{i\chi} \tag{C-52}$$

where χ is any real number. We shall choose $\chi = 0$, which yields:

$$|0,\ 0\rangle = \frac{1}{\sqrt{2}}[|+,\ -\rangle - |-,\ +\rangle] \tag{C-53}$$

Thus we have calculated the four vectors $|S,\ M\rangle$ without explicitly having had to write the matrix that represents $\mathbf{S}^2$ in the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis.

C-4-b. General case (arbitrary j_1 and j_2)

We showed in § C-3-c that the decomposition of $\mathcal{E}(j_1,\ j_2)$ into a direct sum of invariant subspaces $\mathcal{E}(J)$ is:

$$\mathcal{E}(j_1,\ j_2) = \mathcal{E}(j_1 + j_2) \oplus \mathcal{E}(j_1 + j_2 - 1) \oplus \ldots \oplus \mathcal{E}\left(|j_1 - j_2|\right) \tag{C-54}$$

We shall now see how to determine the vectors $|J,\ M\rangle$ that span these subspaces.

α. The subspace $\mathcal{E}(J = j_1 + j_2)$

The ket $|j_1,\ j_2;\ m_1 = j_1,\ m_2 = j_2\rangle$ is, in $\mathcal{E}(j_1,\ j_2)$, the only eigenvector of J_z associated with $M = j_1 + j_2$. Since $\mathbf{J}^2$ and J_z commute, and the value $M = j_1 + j_2$ is not degenerate, $|j_1,\ j_2;\ m_1 = j_1,\ m_2 = j_2\rangle$ must also be an eigenvector of $\mathbf{J}^2$. According to (C-54), the corresponding value of J can only be $j_1 + j_2$. We can choose the phase of the vector:

$$|J = j_1 + j_2\ ,\ M = j_1 + j_2\rangle$$

such that:

$$|j_1 + j_2\ ,\ j_1 + j_2\rangle = |j_1, j_2;\ j_1, j_2\rangle \tag{C-55}$$

Repeated application of the operator J_- on this expression enables us to complete the family of vectors $|J,\ M\rangle$ for which $J = j_1 + j_2$. According to the general formulas (C-50) of Chapter VI:

$$J_- \,|j_1 + j_2\ ,\ j_1 + j_2\rangle = \hbar\sqrt{2(j_1 + j_2)}\,|j_1 + j_2\ ,\ j_1 + j_2 - 1\rangle \tag{C-56}$$

We can therefore calculate the vector corresponding to $J = j_1 + j_2$ and $M = j_1 + j_2 - 1$ by applying $J_- = J_{1-} + J_{2-}$ to the vector $|j_1,\ j_2;\ j_1,\ j_2\rangle$:

$$\begin{aligned}
|j_1 + j_2\ ,\ j_1 + j_2 - 1\rangle &= \frac{1}{\hbar\sqrt{2(j_1 + j_2)}} J_- \,|j_1 + j_2\ ,\ j_1 + j_2\rangle \\
&= \frac{1}{\hbar\sqrt{2(j_1 + j_2)}} (J_{1-} + J_{2-})\,|j_1, j_2;\ j_1, j_2\rangle \\
&= \frac{1}{\hbar\sqrt{2(j_1 + j_2)}} \big[\hbar\sqrt{2j_1}\,|j_1, j_2;\ j_1 - 1,\ j_2\rangle \\
&\quad + \hbar\sqrt{2j_2}\,|j_1, j_2\ ;\ j_1, j_2 - 1\rangle\,\big]
\end{aligned} \tag{C-57}$$

that is:

$$\begin{aligned}
|j_1 + j_2\ ,\ j_1 + j_2 - 1\rangle &= \sqrt{\frac{j_1}{j_1 + j_2}}\,|j_1, j_2\ ;\ j_1 - 1, j_2\rangle \\
&\quad + \sqrt{\frac{j_2}{j_1 + j_2}}\,|j_1,\ j_2\ ;\ j_1, j_2 - 1\rangle
\end{aligned} \tag{C-58}$$

Note that we obtain in this way a linear combination of the two basis vectors that correspond to $M = j_1 + j_2 - 1$, and that this combination is directly normalized.

We then repeat the procedure: we construct $|j_1 + j_2,\ j_1 + j_2 - 2\rangle$ by letting J_- act on both sides of (C-58) (for the right-hand side, we take this operator in the form $J_{1-} + J_{2-}$), and so on, through $|j_1 + j_2, -(j_1 + j_2)\rangle$, which is found to be equal to $|j_1,\ j_2;\ -j_1, -j_2\rangle$.

We therefore know how to calculate the first $[2(j_1 + j_2) + 1]$ vectors of the $\{|J,\ M\rangle\}$ basis, which correspond to $J = j_1 + j_2$ and $M = j_1 + j_2,\ j_1 + j_2 - 1,\ ...,\ -(j_1 + j_2)$ and span the subspace $\mathcal{E}(J = j_1 + j_2)$ of $\mathcal{E}(j_1,\ j_2)$.

β. The other subspaces $\mathcal{E}(J)$

Now consider the space $\mathscr{S}(j_1 + j_2)$, the supplement of $\mathcal{E}(j_1 + j_2)$ in $\mathcal{E}(j_1,\ j_2)$. According to (C-54), $\mathscr{S}(j_1 + j_2)$ can be broken down into:

$$\mathscr{S}(j_1 + j_2) = \mathcal{E}(j_1 + j_2 - 1) \oplus \mathcal{E}(j_1 + j_2 - 2) \oplus ... \oplus \mathcal{E}(|j_1 - j_2|) \tag{C-59}$$

We can therefore apply to it the same reasoning as was used in § α.

In $\mathscr{S}(j_1+j_2)$, the degree of degeneracy $g'_{j_1,j_2}(M)$ of a given value of M is smaller by one than $g_{j_1,j_2}(M)$, since $\mathcal{E}(j_1+j_2)$ possesses one and only one vector associated with this value of M :

$$g'_{j_1,j_2}(M) = g_{j_1,j_2}(M) - 1 \tag{C-60}$$

This means, in particular, that $M = j_1+j_2$ no longer exists in $\mathscr{S}(j_1+j_2)$, and that the new maximum value $M = j_1+j_2-1$ is not degenerate. From this we see, as in § α, that the corresponding vector must be proportional to $|J = j_1+j_2-1,\ M = j_1+j_2-1\rangle$. It is easy to find its expansion on the $\{|j_1,\ j_2;\ m_1,\ m_2\rangle\}$ basis, since, because of the value of M, it is surely of the form:

$$\begin{aligned}|j_1+j_2-1\ ,\ j_1+j_2-1\rangle = \alpha\,|j_1,\ j_2\ ;\ j_1,\ j_2-1\rangle \\ + \beta\,|j_1,\ j_2\ ;\ j_1-1, j_2\rangle\end{aligned} \tag{C-61}$$

with:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{C-62}$$

to insure its normalization. It must also be orthogonal to $|j_1+j_2,\ j_1+j_2-1\rangle$, which belongs to $\mathcal{E}(j_1+j_2)$ and whose expression is given by (C-58). The coefficients α and β must therefore satisfy:

$$\alpha\sqrt{\frac{j_2}{j_1+j_2}} + \beta\sqrt{\frac{j_1}{j_1+j_2}} = 0 \tag{C-63}$$

Relations (C-62) and (C-63) determine α and β to within a phase factor. We shall choose α and β to be real and, for example, α positive. With these conventions:

$$\begin{aligned}|j_1+j_2-1\ ,\ j_1+j_2-1\rangle = \sqrt{\frac{j_1}{j_1+j_2}}\,|j_1,\ j_2\ ;\ j_1, j_2-1\rangle \\ - \sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,\ j_2\ ;\ j_1-1, j_2\rangle\end{aligned} \tag{C-64}$$

This vector is the first of a new family, characterized by $J = j_1+j_2-1$. As in § α, we can derive the others by applying J_- as many times as necessary. Thus we obtain $[2(j_1+j_2-1)+1]$ vectors $|J,\ M\rangle$ corresponding to

$$J = j_1+j_2-1 \quad \text{and} \quad M = j_1+j_2-1,\ j_1+j_2-2,\ ...,\ -(j_1+j_2-1),$$

and spanning the subspace $\mathcal{E}(J = j_1+j_2-1)$.

Now consider the space $\mathscr{S}(j_1+j_2, j_1+j_2-1)$, the supplement of the direct sum $\mathcal{E}(j_1+j_2)\oplus\mathcal{E}(j_1+j_2-1)$ in $\mathcal{E}(j_1,\ j_2)$[4]:

$$\mathscr{S}(j_1+j_2\ , j_1+j_2-1) = \mathcal{E}(j_1+j_2-2)\oplus ...\oplus\mathcal{E}(|j_1-j_2|) \tag{C-65}$$

In the space $\mathscr{S}(j_1+j_2\ ,\ j_1+j_2-1)$, the degeneracy of each value of M is again decreased by one with respect to what it was in $\mathscr{S}(j_1+j_2)$. In particular, the maximum

[4]Of course, $\mathscr{S}(j_1+j_2, j_1+j_2-1)$ exists only if j_1+j_2-2 is not less than $|j_1-j_2|$

value is now $M = j_1 + j_2 - 2$, and it is not degenerate. The corresponding vector of $\mathscr{S}(j_1 + j_2\ , j_1 + j_2 - 1)$ must therefore be $|J = j_1 + j_2 - 2\ , M = j_1 + j_2 - 2\rangle$. To express it in the $\{|j_1, j_2\ ;\ m_1, m_2\rangle\}$ basis, it is sufficient to note that it is a linear combination of the three vectors $|j_1, j_2\ ;\ j_1, j_2 - 2\rangle$, $|j_1, j_2\ ;\ j_1 - 1,\ j_2 - 1\rangle$, $|j_1, j_2\ ;\ j_1 - 2, j_2\rangle$. The coefficients of this combination are fixed to within a phase factor by the triple condition that it be normalized and orthogonal to $|j_1 + j_2\ ,\ j_1 + j_2 - 2\rangle$ and $|j_1 + j_2 - 1\ ,\ j_1 + j_2 - 2\rangle$ (which are already known). Finally, the use of J_- enables us to find the other vectors of this third family, thus defining $\mathcal{E}(j_1 + j_2 - 2)$.

The procedure can be repeated without difficulty until we have exhausted all values of M greater than or equal to $|j_1 - j_2|$ [and, consequently, according to (C-31), also all those less than or equal to $-|j_1 - j_2|$]. We then know all the desired $|J,\ M\rangle$ vectors. This method will be illustrated by two examples in Complement $\mathrm{A_X}$.

C-4-c. Clebsch-Gordan coefficients

In each space $\mathcal{E}(j_1,\ j_2)$, the eigenvectors of $\mathbf{J}^2$ and J_z are linear combinations of vectors of the initial $\{|j_1,\ j_2\ ;\ m_1, m_2\rangle\}$ basis:

$$|J,\ M\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1, j_2\ ;\ m_1, m_2\rangle\, \langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle \tag{C-66}$$

The coefficients $\langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle$ of these expansions are called *Clebsch-Gordan coefficients.*

Comment:

To be completely rigorous, we should write the vectors $|j_1, j_2\ ;\ m_1, m_2\rangle$ and $|J,\ M\rangle$ as $|k_1, k_2;\ j_1, j_2;\ m_1, m_2\rangle$ and $|k_1, k_2;\ j_1, j_2;\ J, M\rangle$ respectively [the values of k_1 and k_2, like those of j_1 and j_2, would then be the same on both sides of relations (C-66)]. However, we shall not write k_1 and k_2 in the symbols which represent the Clebsch-Gordan coefficients, since we know that these coefficients are independent of k_1 and k_2 (§ C-2-c).

It is not possible to give a general expression for the Clebsch-Gordan coefficients, but the method presented in § C-4-b enables us to calculate them by iteration for any values of j_1 and j_2. For practical applications, there are *numerical tables* of Clebsch-Gordan coefficients.

Actually, to determine the Clebsch-Gordan coefficients uniquely, a certain number of *phase conventions* must be chosen. [We mentioned this fact when we wrote expressions (C-55) and (C-64)]. Clebsch-Gordan coefficients are always chosen to be real. The choice then bears on the signs of some of them (obviously, the relative signs of the coefficients appearing in the expansion of the same vector $|J,\ M\rangle$ are fixed; only the global sign of the expansion can be chosen arbitrarily).

The results of § C-4-b imply that $\langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle$ is different from zero only if:

$$M = m_1 + m_2 \tag{C-67a}$$

$$|j_1 - j_2| \leq J \leq j_1 + j_2 \tag{C-67b}$$

where J is of the same type (integral or half-integral) as $j_1 + j_2$ and $|j_1 - j_2|$. Condition (C-67b) is often called the "triangle rule": one must be able to form a triangle with three line segments of lengths j_1, j_2 and J.

Since the vectors $|J,\ M\rangle$ also form an orthonormal basis of the space $\mathcal{E}(j_1,\ j_2)$, the expressions which are the inverse of (C-66) can be written:

$$|j_1, j_2\ ;\ m_1, m_2\rangle = \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^{J} |J, M\rangle \langle J, M|j_1,\ j_2\ ;\ m_1, m_2\rangle \tag{C-68}$$

As the Clebsch-Gordan coefficients have all been chosen to be real, the scalar products appearing in (C-68) are such that:

$$\langle J, M|j_1, j_2\ ;\ m_1, m_2\rangle = \langle j_1, j_2\ ;\ m_1, m_2|J, M\rangle \tag{C-69}$$

The Clebsch-Gordan coefficients therefore enable us to express the vectors of the old basis $\{|j_1, j_2\ ;\ m_1, m_2\rangle\}$, in terms of those of the new basis $\{\langle J, M\rangle\}$.

The Clebsch-Gordan coefficients possess interesting properties, some of which will be studied in Complement $\mathrm{B_X}$.

References and suggestions for further reading:

Messiah (1.17), Chap. XIII, § V; Rose (2.19), Chap. III; Edmonds (2.21), Chaps. 3 and 6.

Relation with group theory: Meijer and Bauer (2.18), Chap. 5 § 5 and App. III of that Chapter; Bacry (10.31),Chap. 6; Wigner (2.23), Chaps. 14 and 15.

Vectorial spherical harmonics: Edmonds (2.21), § 5-10; Jackson (7.5), Chap. 16; Berestetskii et al. (2.8), §§ 6 and 7 ; Akhiezer and Berestetskii (2.14), § 4.

A_X: EXAMPLES OF ADDITION OF ANGULAR MOMENTA	Illustrates the results of Chapter X by the simplest cases not treated in detail in this chapter: two angular momenta equal to 1, and an integral angular momentum l with a spin 1/2. Easy, recommended as an exercise illustrating methods of addition of angular momenta.
B_X: CLEBSCH-GORDAN COEFFICIENTS **C_X: ADDITION OF SPHERICAL HARMONICS**	Technical complements intended to demonstrate certain useful mathematical results; can be used as references. B_X: study of Clebsch-Gordan coefficients, which frequently appear in physical problems involving angular momentum and rotational invariance. C_X: proof of an expression concerning the product of spherical harmonics; useful for certain subsequent complements and exercises.
D_X: VECTOR OPERATORS: THE WIGNER-ECKART THEOREM **E_X: ELECTRIC MULTIPOLE MOMENTS**	Introduction of physical concepts (vector observables, multipole moments) which play important roles in numerous fields. D_X: study of vector operators; proof of the Wigner-Eckart theorem, which establishes proportionality rules between the matrix elements of these operators. Rather theoretical, but recommended for its numerous applications. Can be helpful in an atomic physics course (the vector model, calculation of Landé factors, etc.). E_X: definition and properties of electric multipole moments of a classical or quantum mechanical system; study of their selection rules (these multipole moments are frequently used in atomic and nuclear physics). Moderately difficult.
F_X: EVOLUTION OF TWO ANGULAR MOMENTA $\mathbf{J}_1$ AND $\mathbf{J}_2$ COUPLED BY AN INTERACTION $a\mathbf{J}_1 \cdot \mathbf{J}_2$	Can be considered to be a worked exercise, treating a problem fundamental to the vector model of the atom: the time evolution of two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ coupled by an interaction $W = a\mathbf{J}_1 \cdot \mathbf{J}_2$. This dynamical point of view completes, as it were, the results of Chapter X concerning the eigenstates of W. Fairly simple.
G_X : EXERCISES	Exercises 7 to 10 are more difficult than the others. Exercices 7, 8, 9 are extensions of Complements D_X and F_X (concept of a standard component and that of an irreducible tensor operator, the Wigner-Eckart theorem). Exercise 10 takes up the problem of the various ways of coupling three angular momenta.

Complement A_X

Examples of addition of angular momenta

1 Addition of $\mathbf{j}_1 = 1$ and $\mathbf{j}_2 = 1$ 1043

1-a The subspace $\mathcal{E}(J = 2)$ 1043

1-b The subspace $\mathcal{E}(J = 1)$ 1044

1-c The vector $|J = 0, M = 0\rangle$ 1045

2 Addition of an integral orbital angular momentum l and a spin 1/2 1046

2-a The subspace $\mathcal{E}(J = l + 1/2)$ 1046

2-b The subspace $\mathcal{E}(J = l - 1/2)$ 1048

To illustrate the general method of addition of angular momenta described in Chapter X, we shall apply it here to two examples.

1. Addition of $\mathbf{j}_1 = 1$ and $\mathbf{j}_2 = 1$

First consider the case in which $j_1 = j_2 = 1$. This is the case, for example, for a two-particle system in which both orbital angular momenta are equal to 1. Since each of the two particles is then in a p state, this is said to be a "p^2 configuration".

The space $\mathcal{E}(1, 1)$ with which we are concerned has $3 \times 3 = 9$ dimensions. We assume the basis composed of common eigenstates of $\mathbf{J}_1^2$, $\mathbf{J}_2^2$, J_{1z} and J_{2z} to be known:

$$\{|1, 1; m_1, m_2\rangle\}, \quad \text{with} \quad m_1, m_2 = 1, 0, -1 \tag{1}$$

and we want to determine the $\{|J, M\rangle\}$ basis of common eigenvectors of $\mathbf{J}_1^2$, $\mathbf{J}_2^2$, $\mathbf{J}^2$ and J_z, where $\mathbf{J}$ is the total angular momentum.

According to § C-3 of Chapter X, the possible values of the quantum number J are:

$$J = 2, 1, 0. \tag{2}$$

We must therefore construct three families of vectors $|J, M\rangle$, containing, respectively, five, three and one vectors of the new basis.

1-a. The subspace $\mathcal{E}(J = 2)$

The ket $|J = 2,\ M = 2\rangle$ can be written simply:

$$|2, 2\rangle = |1, 1\ ;\ 1, 1\rangle \tag{3}$$

Applying J_- to it, we find the vector $|J = 2,\ M = 1\rangle$:

$$\begin{aligned}|2,1\rangle &= \frac{1}{2\hbar} J_- |2,\ 2\rangle \\ &= \frac{1}{2\hbar}(J_{1-} + J_{2-})|1,1\ ;\ 1,1\rangle \\ &= \frac{1}{2\hbar}\left[\hbar\sqrt{2}\,|1,1\ ;\ 0,1\rangle + \hbar\sqrt{2}\,|1,1\ ;\ 1,0\rangle\right] \\ &= \frac{1}{\sqrt{2}}\left[|1,1\ ;\ 1,0\rangle + |1,1\ ;\ 0,1\rangle\right] \end{aligned} \tag{4}$$

We use J_- again to calculate $|J = 2,\ M = 0\rangle$. After a simple calculation, we find:

$$|2,0\rangle = \frac{1}{\sqrt{6}}\left[|1,1\ ;\ 1,-1\rangle + 2|1,1\ ;\ 0,0\rangle + |1,1\ ;\ -1,1\rangle\right] \tag{5}$$

then:

$$|2,-1\rangle = \frac{1}{\sqrt{2}}\left[|1,1\ ;\ 0,-1\rangle + |1,1\ ;\ -1,0\rangle\right] \tag{6}$$

and, finally:

$$|2,-2\rangle = |1,\ 1;\ -1,\ -1\rangle \tag{7}$$

1-b. The subspace $\mathcal{E}(J = 1)$

We shall now proceed to the subspace $\mathcal{E}(J = 1)$. The vector $|J = 1,\ M = 1\rangle$ must be a linear combination of the two basis kets $|1,1\ ;\ 1,0\rangle$ and $|1,1\ ;\ 0,1\rangle$ (the only ones for which $M = 1$):

$$|1,1\rangle = \alpha|1,1\ ;\ 1,0\rangle + \beta|1,1\ ;\ 0,1\rangle \tag{8}$$

with:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{9}$$

For it to be orthogonal to the vector $|2,1\rangle$, it is necessary [*cf.* (4)] that:

$$\alpha + \beta = 0 \tag{10}$$

We choose α and β to be real, and choose, by convention, α positive[1]. Under these conditions:

$$|1,1\rangle = \frac{1}{\sqrt{2}}\left[|1,1\ ;\ 1,0\rangle - |1,1\ ;\ 0,1\rangle\right] \tag{11}$$

Application of J_- here again enables us to deduce $|1,0\rangle$ and $|1,-1\rangle$. We easily find, using the same technique as above:

$$|1,0\rangle = \frac{1}{\sqrt{2}}\left[|1,1\ ;\ 1,-1\rangle - |1,1\ ;\ -1,1\rangle\right] \tag{12}$$

$$|1,-1\rangle = \frac{1}{\sqrt{2}}\left[|1,1\ ;\ 0,-1\rangle - |1,1\ ;\ -1,0\rangle\right] \tag{13}$$

[1]The component of the ket $|J,\ J\rangle$ on the ket $|j_1,\ j_2\ ;\ m_1 = j_1,\ m_2 = J - j_1\rangle$ is always chosen to be real and positive (*cf.* Complement B_X, § 2).

It is interesting to note that expansion (12) does not contain the vector $|1,1\ ;\ 0,0\rangle$, although it also corresponds to $M = 0$. It so happens that the corresponding Clebsch-Gordan coefficient is zero:

$$\langle 1,1\ ;\ 0,0|1,0\rangle = 0 \tag{14}$$

1-c. The vector $|J = 0,\ M = 0\rangle$

We are left with the calculation of the last vector of the $\{|J,\ M\rangle\}$ basis, associated with $J = M = 0$. This vector is a linear combination of the three basis kets for which $M = 0$:

$$|0,0\rangle = a|1,1\ ;\ 1,-1\rangle + b|1,1\ ;\ 0,0\rangle + c|1,1\ ;\ -1,1\rangle \tag{15}$$

with:

$$|a|^2 + |b|^2 + |c|^2 = 1 \tag{16}$$

It must also be orthogonal to $|2,\ 0\rangle$ [formula (5)] and $|1,\ 0\rangle$ [formula (12)]. This gives the two conditions:

$$a + 2b + c = 0 \tag{17a}$$

$$a - c = 0 \tag{17b}$$

These relations imply:

$$a = -b = c \tag{18}$$

We again choose a, b and c real, and agree to choose a positive (see note 1). We then obtain, using (16) and (18):

$$|0,0\rangle = \frac{1}{\sqrt{3}}\left[|1,1\ ;\ 1,-1\rangle - |1,1\ ;\ 0,0\rangle + |1,1\ ;\ -1,1\rangle\right] \tag{19}$$

This completes the construction of the $\{|J,\ M\rangle\}$ basis for the case $j_1 = j_2 = 1$.

Comment:

If the physical problem under study is that of a p^2 configuration of a two-particle system, the wave functions which represent the states of the initial basis are of the form:

$$\langle \mathbf{r}_1,\ \mathbf{r}_2|1,1\ ;\ m_1,\ m_2\rangle = R_{k_1,1}(r_1)R_{k_2,1}(r_2)Y_1^{m_1}(\theta_1,\varphi_1)Y_1^{m_2}(\theta_2,\varphi_2) \tag{20}$$

where $\mathbf{r}_1(r_1,\ \theta_1,\varphi_1)$ and $\mathbf{r}_2(r_2,\ \theta_2,\varphi_2)$ give the positions of the two particles. Since the radial functions are independent of the quantum numbers m_1 and m_2, the linear combinations that give the wave functions associated with the kets $|J,\ M\rangle$ are functions only of the angular dependence. For example, in the $\{|\mathbf{r}_1,\ \mathbf{r}_2\rangle\}$ representation, equation (19) can be written:

$$\begin{aligned}\langle \mathbf{r}_1,\ \mathbf{r}_2|0,\ 0\rangle = R_{k_1,1}(r_1)R_{k_2,1}(r_2)\frac{1}{\sqrt{3}}\big[&Y_1^1(\theta_1,\varphi_1)Y_1^{-1}(\theta_2,\varphi_2)\\ &-Y_1^0(\theta_1,\varphi_1)Y_1^0(\theta_2,\varphi_2) + Y_1^{-1}(\theta_1,\varphi_1)Y_1^1(\theta_2,\varphi_2)\big]\end{aligned} \tag{21}$$

2. Addition of an integral orbital angular momentum l and a spin 1/2

Now consider the addition of an orbital angular momentum ($j_1 = l$, an integer) and a spin 1/2 ($j_2 = 1/2$). This problem is encountered, for example, whenever one wants to study the total angular momentum of a spin 1/2 particle such as the electron.

The space $\mathcal{E}(l, 1/2)$ which we are considering here is $2(2l+1)$-dimensional. We already know a basis of this space[2]:

$$\{|l, 1/2\ ;\ m, \varepsilon\rangle\} \quad \text{with} \quad m = l,\, l-1,\, ...,\, -l \quad \text{and} \quad \varepsilon = \pm \tag{22}$$

formed of eigenstates of the observables $\mathbf{L}^2$, $\mathbf{S}^2$, L_z and S_z, where $\mathbf{L}$ and $\mathbf{S}$ are the orbital angular momentum and spin under consideration. We want to construct the eigenvectors $|J, M\rangle$ of $\mathbf{J}^2$ and J_z, where $\mathbf{J}$ is the total angular momentum of the system:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{23}$$

First of all, note that if l is zero, the solution to the problem is obvious. It is easy to show in this case that the vectors $|0, 1/2; 0, \varepsilon\rangle$ are also eigenvectors of $\mathbf{J}^2$ and J_z with eigenvalues such as $J = 1/2$ and $M = \varepsilon/2$. On the other hand, if l is not zero, there are two possible values of J :

$$J = l + \frac{1}{2} \quad , \quad l - \frac{1}{2} \tag{24}$$

2-a. The subspace $\mathcal{E}(J = l + 1/2)$

The $(2l+2)$ vectors $|J, M\rangle$ spanning the subspace $\mathcal{E}(J = l+1/2)$ can be obtained by using the general method of Chapter X. We have, first of all:

$$\left|l + \frac{1}{2}, l + \frac{1}{2}\right\rangle = \left|l, \frac{1}{2}; l, +\right\rangle \tag{25}$$

Through the action[3] of J_-, we obtain $\left|l + \frac{1}{2}, l - \frac{1}{2}\right\rangle$:

$$\begin{aligned}
\left|l + \frac{1}{2}, l - \frac{1}{2}\right\rangle &= \frac{1}{\hbar\sqrt{2l+1}} J_- \left|l + \frac{1}{2}, l + \frac{1}{2}\right\rangle \\
&= \frac{1}{\hbar\sqrt{2l+1}} (L_- + S_-) \left|l, \frac{1}{2}; l, +\right\rangle \\
&= \frac{1}{\hbar\sqrt{2l+1}} \left[\hbar\sqrt{2l} \left|l, \frac{1}{2}; l-1, +\right\rangle + \hbar \left|l, \frac{1}{2}; l, -\right\rangle\right] \\
&= \sqrt{\frac{2l}{2l+1}} \left|l, \frac{1}{2}; l-1, +\right\rangle + \frac{1}{\sqrt{2l+1}} \left|l, \frac{1}{2}; l, -\right\rangle
\end{aligned} \tag{26}$$

[2] If we wanted to conform strictly to the notation of Chapter X, we should write $\pm 1/2$, and not ε, in the basis kets. But we agreed in Chapters IV and IX to denote the eigenvectors of S_z in the spin state space by $|+\rangle$ and $|-\rangle$.

[3] To find the numerical coefficients appearing in the following equations, we can simply use the relation: $j(j+1) - m(m-1) = (j+m)(j-m+1)$.

We apply J_- again. An analogous calculation yields:

$$\left|l+\frac{1}{2},\, l-\frac{3}{2}\right\rangle = \frac{1}{\sqrt{2l+1}}\left[\sqrt{2l-1}\left|l,\,\frac{1}{2};\, l-2,\,+\right\rangle + \sqrt{2}\left|l,\,\frac{1}{2};\, l-1,\,-\right\rangle\right] \tag{27}$$

More generally, the vector $|l+1/2,\, M\rangle$ will be a linear combination of the only two basis vectors associated with M : $|l,\, 1/2;\, M-1/2,\, +\rangle$ and $|l,\, 1/2;\, M+1/2,\, -\rangle$ (M is, of course, half-integral). Comparing (25), (26) and (27), we can guess that this linear combination should be:

$$\left|l+\frac{1}{2},\, M\right\rangle = \frac{1}{\sqrt{2l+1}}\left[\sqrt{l+M+\frac{1}{2}}\left|l,\,\frac{1}{2};\, M-\frac{1}{2},\,+\right\rangle + \sqrt{l-M+\frac{1}{2}}\left|l,\,\frac{1}{2};\, M+\frac{1}{2},\,-\right\rangle\right] \tag{28}$$

with:

$$M = l+\frac{1}{2}\ ,\ \ l-\frac{1}{2}\ ,\ \ l-\frac{3}{2}\ ,\ \ ...,-l+\frac{1}{2}\ ,-\left(l+\frac{1}{2}\right) \tag{29}$$

Reasoning by recurrence, we can show this to be true, since application of J_- to both sides of (28) yields:

$$\begin{aligned}
\left|l+\frac{1}{2},\, M-1\right\rangle &= \frac{1}{\hbar\sqrt{\left(l+M+\frac{1}{2}\right)\left(l-M+\frac{3}{2}\right)}} J_- \left|l+\frac{1}{2},\, M\right\rangle \\
&= \frac{1}{\hbar\sqrt{\left(l+M+\frac{1}{2}\right)\left(l-M+\frac{3}{2}\right)}} \frac{1}{\sqrt{2l+1}} \times \\
&\times \left[\sqrt{l+M+\frac{1}{2}}\,\hbar\sqrt{\left(l+M-\frac{1}{2}\right)\left(l-M+\frac{3}{2}\right)}\left|l,\,\frac{1}{2};\, M-\frac{3}{2},\,+\right\rangle \right. \\
&\quad + \sqrt{l+M+\frac{1}{2}}\,\hbar\left|l,\,\frac{1}{2};\, M-\frac{1}{2},\,-\right\rangle \\
&\left. + \sqrt{l-M+\frac{1}{2}}\,\hbar\sqrt{\left(l+M+\frac{1}{2}\right)\left(l-M+\frac{1}{2}\right)}\left|l,\,\frac{1}{2};\, M-\frac{1}{2},\,-\right\rangle\right] \\
&= \frac{1}{\sqrt{2l+1}}\left[\sqrt{(l+M-\frac{1}{2}}\left|l,\,\frac{1}{2};\, M-\frac{3}{2},\,+\right\rangle \right. \\
&\left. + \sqrt{l-M+\frac{3}{2}}\left|l,\,\frac{1}{2};\, M-\frac{1}{2},\,-\right\rangle\right]
\end{aligned} \tag{30}$$

We indeed obtain the same expression as in (28), with M changed to $M-1$.

2-b. The subspace $\mathcal{E}(J = l - 1/2)$

We shall now try to determine the expression for the $2l$ vectors $|J,\ M\rangle$ associated with $J = l - 1/2$. The one which corresponds to the maximum value $l - 1/2$ of M is a normalized linear combination of $|l, 1/2\ ;\ l-1, +\rangle$ and $|l, 1/2\ ;\ l, -\rangle$, and it must be orthogonal to $|l + 1/2,\ l - 1/2\rangle$ [formula (26)]. Choosing the coefficient of $|l, 1/2\ ;\ l,\ -\rangle$ real and positive (*cf.* note 1), we easily find:

$$\left|l - \frac{1}{2},\ l - \frac{1}{2}\right\rangle = \frac{1}{\sqrt{2l+1}}\left[\sqrt{2l}\left|l, \frac{1}{2}\ ;\ l, -\right\rangle - \left|l, \frac{1}{2}\ ;\ l-1, +\right\rangle\right] \tag{31}$$

The operator J_- enables us to deduce successively all the other vectors of the family characterized by $J = l - 1/2$. Since there are only two basis vectors with a given value of M, and since $|l - 1/2,\ M\rangle$ is orthogonal to $|l + 1/2,\ M\rangle$, (28) leads us to expect that:

$$\left|l - \frac{1}{2},\ M\right\rangle = \frac{1}{\sqrt{2l+1}}\left[\sqrt{l + M + \frac{1}{2}}\left|l, \frac{1}{2}\ ;\ M + \frac{1}{2}, -\right\rangle - \sqrt{l - M + \frac{1}{2}}\left|l, \frac{1}{2}; M - \frac{1}{2}, +\right\rangle\right] \tag{32}$$

for:

$$M = l - \frac{1}{2}\ ,\ \ l - \frac{3}{2}\ ,\ \ ..., -l + \frac{3}{2}\ ,\ -\left(l - \frac{1}{2}\right) \tag{33}$$

By an argument analogous to the one in § 2-a, this formula can also be proved by recurrence.

Comments:

(*i*) The states $|l,\ 1/2;\ m,\ \varepsilon\rangle$ of a spin 1/2 particle can be represented by two-component spinors of the form:

$$\left[\psi_{l,\ \frac{1}{2};\ m,\ +}\right](\mathbf{r}) = R_{k,l}(r)Y_l^m(\theta,\varphi)\begin{pmatrix}1\\0\end{pmatrix} \tag{34a}$$

$$\left[\psi_{l,\ \frac{1}{2};\ m,\ -}\right](\mathbf{r}) = R_{k,l}(r)Y_l^m(\theta,\varphi)\begin{pmatrix}0\\1\end{pmatrix} \tag{34b}$$

The preceding calculations then show that the spinors associated with the states $|J,\ M\rangle$ can be written:

$$\left[\psi_{l+\frac{1}{2},M}\right](\mathbf{r}) = \frac{1}{\sqrt{2l+1}}R_{k,l}(r)\begin{pmatrix}\sqrt{l + M + \frac{1}{2}}\ Y_l^{M-\frac{1}{2}}(\theta,\varphi)\\ \sqrt{l - M + \frac{1}{2}}\ Y_l^{M+\frac{1}{2}}(\theta,\varphi)\end{pmatrix} \tag{35a}$$

$$\left[\psi_{l-\frac{1}{2},M}\right](\mathbf{r}) = \frac{1}{\sqrt{2l+1}}R_{k,l}(r)\begin{pmatrix}-\sqrt{l - M + \frac{1}{2}}\ Y_l^{M-\frac{1}{2}}(\theta,\varphi)\\ \sqrt{l + M + \frac{1}{2}}\ Y_l^{M+\frac{1}{2}}(\theta,\varphi)\end{pmatrix} \tag{35b}$$

(ii) In the particular case $l = 1$, formulas (25), (28), (31) and (32) yield:

$$\begin{aligned}
\left|\frac{3}{2}, \frac{3}{2}\right\rangle &= \left|1, \frac{1}{2}\ ;\ 1, +\right\rangle \\
\left|\frac{3}{2}, \frac{1}{2}\right\rangle &= \sqrt{\frac{2}{3}}\left|1, \frac{1}{2}\ ;\ 0, +\right\rangle + \frac{1}{\sqrt{3}}\left|1, \frac{1}{2}\ ;\ 1, -\right\rangle \\
\left|\frac{3}{2}, -\frac{1}{2}\right\rangle &= \frac{1}{\sqrt{3}}\left|1, \frac{1}{2}\ ;\ -1, +\right\rangle + \sqrt{\frac{2}{3}}\left|1, \frac{1}{2}\ ;\ 0, -\right\rangle \\
\left|\frac{3}{2}, -\frac{3}{2}\right\rangle &= \left|1, \frac{1}{2}\ ;\ -1, -\right\rangle
\end{aligned} \tag{36a}$$

and:

$$\begin{aligned}
\left|\frac{1}{2}, \frac{1}{2}\right\rangle &= \sqrt{\frac{2}{3}}\left|1, \frac{1}{2}\ ;\ 1, -\right\rangle - \frac{1}{\sqrt{3}}\left|1, \frac{1}{2}\ ;\ 0, +\right\rangle \\
\left|\frac{1}{2}, -\frac{1}{2}\right\rangle &= \frac{1}{\sqrt{3}}\left|1, \frac{1}{2}\ ;\ 0, -\right\rangle - \sqrt{\frac{2}{3}}\left|1, \frac{1}{2}\ ;\ -1, +\right\rangle
\end{aligned} \tag{36b}$$

References and suggestions for further reading:

Addition of an angular momentum l and an angular momentum $S = 1$: see "vectorial spherical harmonics" in the references of Chapter X.

Complement B_X

Clebsch-Gordan coefficients

1 General properties of Clebsch-Gordan coefficients 1051
1-a Selection rules 1051
1-b Orthogonality relations 1053
1-c Recurrence relations 1054
2 Phase conventions. Reality of Clebsch-Gordan coefficients 1054
2-a The coefficients $\langle j_1, j_2\ ;\ m_1, m_2 | J, J\rangle$; phase of the ket $|J, J\rangle$. 1055
2-b Other Clebsch-Gordan coefficients 1055
3 Some useful relations 1056
3-a The signs of some coefficients 1056
3-b Changing the order of j_1 and j_2 1057
3-c Changing the sign of M, m_1 and m_2 1057
3-d The coefficients $\langle j, j\ ;\ m, -m | 0, 0\rangle$ 1057

Clebsch-Gordan coefficients were introduced in Chapter X [*cf.* relation (C-66)]: they are the coefficients $\langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle$ involved in the expansion of the ket $|J, M\rangle$ on the $\{|j_1,\ j_2\ ;\ m_1, m_2\rangle\}$ basis:

$$|J, M\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle\, |j_1, j_2\ ;\ m_1, m_2\rangle \tag{1}$$

In this complement, we shall derive some interesting properties of Clebsch-Gordan coefficients, some of which were simply stated in Chapter X.

Note that, to define the $\langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle$ completely, equation (1) is not sufficient. The normalized vector $|J, M\rangle$ is fixed only to within a phase factor by the corresponding eigenvalues $J(J+1)\hbar^2$ and $M\hbar$, and a phase convention must be chosen in order to complete the definition. In Chapter X, we used the action of the J_- and J_+ operators to fix the relative phase of the $(2J+1)$ kets $|J, M\rangle$ associated with the same value of J. In this complement, we shall complete this choice of phase by adopting a convention for the phase of the kets $|J, J\rangle$. This will enable us to show that all the Clebsch-Gordan coefficients are then real.

However, before approaching, in § 2, the problem of the choice of the phase of the $\langle j_1,\ j_2;\ m_1,\ m_2 | J,\ M\rangle$, we shall, in § 1, study some of their most useful properties which do not depend on this phase convention. Finally, § 3 presents various relations which will be of use in other complements.

1. General properties of Clebsch-Gordan coefficients

1-a. Selection rules

Two important selection rules, which follow directly from the results of Chapter X concerning the addition of angular momenta, have already been given in that chapter [*cf.*

relations (C-67a) and (C-67b)]. We shall simply restate them here: the Clebsch-Gordan coefficient $\langle j_1, j_2\ ;\ m_1, m_2|J, M\rangle$ is necessarily zero if the following two conditions are not simultaneously satisfied:

$$M = m_1 + m_2 \tag{2}$$

$$|j_1 - j_2| \leq J \leq j_1 + j_2 \tag{3a}$$

Inequality (3a) is often called the "triangle selection rule", since it means that a triangle can be formed with three line segments of lengths j_1, j_2 and J (*cf.* Fig. 1). These three numbers therefore play symmetrical roles here, and (3a) can also be written in the form:

$$|J - j_1| \leq j_2 \leq J + j_1 \tag{3b}$$

or:

$$|J - j_2| \leq j_1 \leq J + j_2 \tag{3c}$$

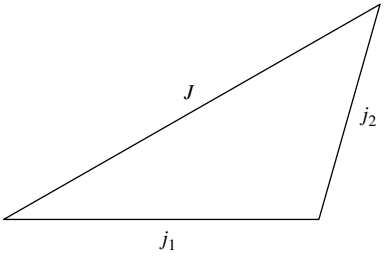


Figure 1: Triangle selection rule: the coefficient $\langle j_1, j_2; m_1, m_2|J, M\rangle$ can be different from zero only if it is possible to form a triangle with three line segments of lengths j_1, j_2, J.

Moreover, the general properties of angular momentum require that the ket $|J, M\rangle$ and, therefore, the coefficient $\langle j_1, j_2\ ;\ m_1, m_2|J, M\rangle$ exist only if M takes on one of the values:

$$M = J\ ,\ J-1\ ,\ J-2\ ,\ ..., -J \tag{4a}$$

Similarly, it is necessary that:

$$m_1 = j_1\ ,\ j_1 - 1\ ,\ ,..., -j_1 \tag{4b}$$
$$m_2 = j_2\ ,\ j_2 - 1\ ,\ ,..., -j_2 \tag{4c}$$

If this is not the case, the Clebsch-Gordan coefficients are not defined. However, in what follows, it will be convenient to assume that they exist for all m_1, m_2 and M, but that they are zero if at least one of conditions (4) is not satisfied. These relations thus play the role of new selection rules for the Clebsch-Gordan coefficients.

1-b. Orthogonality relations

Inserting the closure relation[1]:

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1, j_2 \ ; \ m_1, m_2\rangle \langle j_1, j_2 \ ; \ m_1, m_2| = 1 \tag{5}$$

in the orthogonality relation of the kets $|J, M\rangle$:

$$\langle J, M|J', M'\rangle = \delta_{JJ'}\delta_{MM'} \tag{6}$$

we obtain:

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle J, M|j_1, j_2 \ ; \ m_1, m_2\rangle \langle j_1, j_2 \ ; \ m_1, m_2|J', M'\rangle = \delta_{JJ'}\delta_{MM'} \tag{7a}$$

We shall see later [*cf.* relation (18b)] that the Clebsch-Gordan coefficients are real, which enables us to write this relation in the form:

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle j_1, j_2 \ ; \ m_1, m_2|J, M\rangle \langle j_1, j_2 \ ; \ m_1, m_2|J', M'\rangle = \delta_{JJ'}\delta_{MM'} \tag{7b}$$

Thus we obtain a first "orthogonality relation" for the Clebsch-Gordan coefficients. We note, moreover, that the summation which appears in it is performed over only one index: for the coefficients of the left-hand side to be different from zero, m_1 and m_2 must be related by (2).

Similarly, we insert the closure relation:

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^{J} |J, M\rangle \langle J, M| = 1 \tag{8}$$

in the orthogonality relation of the kets $|j_1, j_2 \ ; \ m_1, m_2\rangle$; we obtain:

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^{J} \langle j_1, j_2 \ ; \ m_1, m_2|J, M\rangle \langle J, M|j_1, j_2 \ ; \ m_1', m_2'\rangle = \delta_{m_1 m_1'}\delta_{m_2 m_2'} \tag{9a}$$

that is, with (18b) taken into account:

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^{J} \langle j_1, j_2 \ ; \ m_1, m_2|J, M\rangle \langle j_1, j_2 \ ; \ m_1', m_2'|J, M\rangle = \delta_{m_1 m_1'}\delta_{m_2 m_2'} \tag{9b}$$

Again, the summation is performed over only one index: since we must have $M = m_1 + m_2$, the summation over M reduces to a single term.

[1] This closure relation is valid for a given subspace $\mathcal{E}(k_1, k_2 \ ; \ j_1, j_2)$ (*cf.* Chap. X, § C-2).

1-c. Recurrence relations

In this section, we shall use the fact that the kets $|j_1, j_2\ ;\ m_1, m_2\rangle$ form a standard basis. Thus:

$$\begin{aligned} J_{1\pm}|j_1, j_2\ ;\ m_1, m_2\rangle &= \hbar\sqrt{j_1(j_1+1) - m_1(m_1 \pm 1)}\,|j_1, j_2\ ;\ m_1 \pm 1, m_2\rangle \\ J_{2\pm}|j_1, j_2\ ;\ m_1, m_2\rangle &= \hbar\sqrt{j_2(j_2+1) - m_2(m_2 \pm 1)}\,|j_1, j_2\ ;\ m_1, m_2 \pm 1\rangle \end{aligned} \tag{10}$$

Similarly, by construction, the kets $|J,\ M\rangle$ satisfy:

$$J_\pm|J, M\rangle = \hbar\sqrt{J(J+1) - M(M \pm 1)}\,|J, M \pm 1\rangle \tag{11}$$

We shall therefore apply the J_- operator to relation (1). Since $J_- = J_{1-} + J_{2-}$, we obtain (if $M > -J$):

$$\begin{aligned} \sqrt{J(J+1) - M(M-1)}\,|J, M-1\rangle = \sum_{m_1'=-j_1}^{j_1} \sum_{m_2'=-j_2}^{j_2} \langle j_1, j_2\ ;\ m_1', m_2'|J, M\rangle \\ \times \left[\sqrt{j_1(j_1+1) - m_1'(m_1'-1)}\,|j_1, j_2\ ;\ m_1' - 1, m_2'\rangle \right. \\ \left. + \sqrt{j_2(j_2+1) - m_2'(m_2'-1)}\,|j_1, j_2\ ;\ m_1', m_2' - 1\rangle\right] \end{aligned} \tag{12}$$

Multiplying this relation by the bra $\langle j_1, j_2\ ;\ m_1, m_2|$, we find:

$$\begin{aligned} &\sqrt{J(J+1) - M(M-1)}\,\langle j_1, j_2\ ;\ m_1, m_2|J, M-1\rangle \\ &\quad = \sqrt{j_1(j_1+1) - m_1(m_1+1)}\,\langle j_1, j_2\ ;\ m_1 + 1, m_2|J, M\rangle \\ &\quad + \sqrt{j_2(j_2+1) - m_2(m_2+1)}\,\langle j_1, j_2\ ;\ m_1, m_2 + 1|J, M\rangle \end{aligned} \tag{13}$$

If the value of M is equal to $-J$, we have $J_-|J, -J\rangle = 0$, and relation (13) remains valid if we use the convention, given above in § 1-b, according to which $\langle j_1, j_2\ ;\ m_1 + 1, m_2|J, M\rangle$ is zero if $|M| > J$.

Analogously, application of the operator $J_+ = J_{1+} + J_{2+}$ to relation (1) leads to:

$$\begin{aligned} &\sqrt{J(J+1) - M(M+1)}\,\langle j_1, j_2\ ;\ m_1, m_2|J, M+1\rangle \\ &\quad = \sqrt{j_1(j_1+1) - m_1(m_1-1)}\,\langle j_1, j_2\ ;\ m_1 - 1, m_2|J, M\rangle \\ &\quad + \sqrt{j_2(j_2+1) - m_2(m_2-1)}\,\langle j_1, j_2\ ;\ m_1, m_2 - 1|J, M\rangle \end{aligned} \tag{14}$$

(the left-hand side of this relation is zero if $M = J$); (13) and (14) are recurrence relations for the Clebsch-Gordan coefficients.

2. Phase conventions. Reality of Clebsch-Gordan coefficients

As we have seen, expressions (12) fix the relative phases of the kets $|J,\ M\rangle$ associated with the same value of J. To complete the definition of the Clebsch-Gordan coefficients involved in (1), we must choose the phase of the various kets $|J,\ J\rangle$. To this end, we shall begin by studying some properties of the coefficients $\langle j_1, j_2\ ;\ m_1, m_2|J, J\rangle$.

2-a. The coefficients $\langle j_1, j_2\ ;\ m_1, m_2|J, J\rangle$; phase of the ket $|J, J\rangle$

In the coefficient $\langle j_1, j_2\ ;\ m_1, m_2 - 1|J, J\rangle$, the maximum value of m_1 is $m_1 = j_1$. According to selection rule (2), m_2 is then equal to $J - j_1$, [whose modulus is well below j_2, according to (3b)]. As m_1 decreases from this maximum value j_1, one unit at a time, m_2 increases until it reaches its maximum value $m_2 = j_2$ [m_1 is then equal to $J - j_2$, whose modulus is well below j_1, according to (3c)]. In theory, therefore, $(j_1 + j_2 - J + 1)$ non-zero Clebsch-Gordan coefficients $\langle j_1, j_2\ ;\ m_1, m_2|J, J\rangle$ can exist. We are going to show that, indeed, none of them is ever zero.

If we set $M = J$ in (14), we obtain:

$$\langle j_1, j_2\ ;\ m_1 - 1, m_2|J, J\rangle = -\sqrt{\frac{j_2(j_2+1) - m_2(m_2-1)}{j_1(j_1+1) - m_1(m_1-1)}}\,\langle j_1, j_2\ ;\ m_1, m_2 - 1|J, J\rangle \tag{15}$$

The radical on the right-hand side of this relation is never zero, nor is it infinite, so long as the Clebsch-Gordan coefficients appearing there satisfy rules (4b) and (4c). Relation (15) therefore shows that if $\langle j_1, j_2\ ;\ j_1, J - j_1|J, J\rangle$ were equal to zero, $\langle j_1, j_2\ ;\ j_1 - 1, J - j_1 + 1|J, J\rangle$ would be zero as well, as would be all the succeeding coefficients $\langle j_1, j_2\ ;\ m_1, J - m_1|J, J\rangle$. Now, this is impossible, since the ket $|J,\ J\rangle$, which is normalized, cannot be zero. Therefore, all the coefficients $\langle j_1, j_2\ ;\ m_1, J - m_1|J, J\rangle$ (with $j_1 \geq m_1 \geq J - j_2$) are different from zero.

In particular, the coefficient $\langle j_1, j_2\ ;\ j_1, J - j_1|J, J\rangle$, in which m_1 takes on its maximum value, is not zero. To fix the phase of the ket $|J,\ J\rangle$, we shall require this coefficient to satisfy the condition:

$$\langle j_1, j_2\ ;\ j_1, J - j_1|J, J\rangle \qquad \text{real and positive} \tag{16}$$

Relation (15) then implies by recurrence that all the coefficients

$$\langle j_1, j_2\ ;\ m_1, J - m_1|J, J\rangle$$

are real [their sign being $(-1)^{j_1 - m_1}$].

Comment:

The phase convention we have chosen for the ket $|J, J\rangle$ gives the two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ asymmetrical roles. It actually depends on the order in which the quantum numbers j_1 and j_2 are arranged in the Clebsch-Gordan coefficients: if j_1 and j_2 are permuted, the phase of the ket $|J, J\rangle$ will be fixed by the condition:

$$\langle j_2, j_1\ ;\ j_2, J - j_2|J, J\rangle \qquad \text{real and positive} \tag{17}$$

which is not necessarily equivalent, *a priori*, to (16) [(16) and (17) may define different phases for the ket $|J, J\rangle$]. We shall return to this point in § 3-b.

2-b. Other Clebsch-Gordan coefficients

Relation (13) enables us to express, in terms of the $\langle j_1, j_2\ ;\ m_1, m_2|J, J\rangle$, all the coefficients $\langle j_1, j_2\ ;\ m_1, m_2|J, J - 1\rangle$; then, by the same method, all the other coefficients $\langle j_1, j_2\ ;\ m_1, m_2|J, J - 2\rangle$, etc. This relation, in which no imaginary numbers are involved, requires that all Clebsch-Gordan coefficients be real:

$$\langle j_1, j_2\ ;\ m_1, m_2|J, M\rangle^* = \langle j_1, j_2\ ;\ m_1, m_2|J, M\rangle \tag{18a}$$

which can also be written:

$$\langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle = \langle J, M | j_1, j_2\ ;\ m_1, m_2\rangle \tag{18b}$$

However, the signs of the $\langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle$ do not obey any simple rule for $M \neq J$.

3. Some useful relations

In this section, we give some useful relations, which complement those given in § 1. To prove them, we shall begin by studying the signs of a certain number of Clebsch-Gordan coefficients.

3-a. The signs of some coefficients

α. The coefficients $\langle j_1, j_2\ ;\ m_1, m_2 | j_1 + j_2, M\rangle$

Convention (16) requires the coefficient $\langle j_1, j_2\ ;\ j_1, j_2 | j_1 + j_2, j_1 + j_2\rangle$ to be real and positive; it is, moreover, equal to 1 (*cf.* Chap. X, § C-4-b-α). Setting $M = J = j_1 + j_2$ in (13), we then see that the coefficients $\langle j_1, j_2\ ;\ m_1, m_2 | j_1 + j_2, j_1 + j_2 - 1\rangle$ are positive. By recurrence, it is then easy to prove that:

$$\langle j_1, j_2\ ;\ m_1, m_2 | j_1 + j_2, M\rangle \geq 0 \tag{19}$$

β. Coefficients in which m_1 *has its maximum value*

Consider the coefficient $\langle j_1, j_2\ ;\ m_1, m_2 | J, M\rangle$. In theory, the maximum value of m_1 is $m_1 = j_1$. However, we then have $m_2 = M - j_1$, which, according to (4c), is possible only if $M - j_1 \geq -j_2$ that is:

$$M \geq j_1 - j_2 \tag{20}$$

If, on the other hand:

$$M \leq j_1 - j_2 \tag{21}$$

the maximum value of m_1 corresponds to the minimum value of m_2 ($m_2 = -j_2$), and is therefore equal to $m_1 = M + j_2$.

Let us show that all Clebsch-Gordan coefficients for which m_1 has its maximum value are non-zero and positive. To do so, we set $m_1 = j_1$, in (13); we find:

$$\sqrt{J(J+1) - M(M-1)}\,\langle j_1, j_2\ ;\ j_1, m_2 | J, M-1\rangle = \sqrt{j_2(j_2+1) - m_2(m_2+1)}\,\langle j_1, j_2\ ;\ j_1, m_2+1 | J, M\rangle \tag{22}$$

Using this relation, an argument by recurrence starting with (16) shows that all the coefficients $\langle j_1,\ j_2;\ j_1,\ M - j_1 | J,\ M\rangle$ are positive [and non-zero if M satisfies (20)]. Analogously, setting $m_2 = -j_2$ in (14), we could prove that all the coefficients $\langle j_1,\ j_2;\ M + j_2, -j_2 | J,\ M\rangle$ are positive [if M satisfies (21)].

γ. The coefficients $\langle j_1, j_2\ ;\ m_1, m_2|J, J\rangle$ *and* $\langle j_1, j_2\ ;\ m_1, m_2|J, -J\rangle$

We saw in §2-a that the sign of $\langle j_1, j_2\ ;\ m_1, m_2|J, J\rangle$ is $(-1)^{j_1-m_1}$. In particular:

$$\text{the sign of } \langle j_1, j_2\ ;\ J-j_2, j_2|J, J\rangle \text{ is } (-1)^{j_1+j_2-J} \tag{23}$$

To determine the sign of $\langle j_1, j_2\ ;\ m_1, m_2|J, -J\rangle$, we can set $M = -J$ in (13), whose left-hand side then goes to zero. We therefore see that the sign of $\langle j_1, j_2\ ;\ m_1, m_2|J, -J\rangle$ changes whenever m_1 (or m_2) varies by ± 1. Since, according to § β, $\langle j_1, j_2\ ;\ j_2 - J, -j_2|J, -J\rangle$ is positive, it follows that the sign of $\langle j_1, j_2\ ;\ m_1, m_2|J, -J\rangle$ is $(-1)^{m_2+j_2}$, and, in particular:

$$\text{the sign of } \langle j_1, j_2\ ; -j_1, -J+j_1|J, -J\rangle \text{ is } (-1)^{j_1+j_2-J} \tag{24}$$

3-b. Changing the order of j_1 and j_2

With the conventions we have chosen, the phase of the ket $|J,\ J\rangle$ depends on the order in which the two angular momenta j_1 and j_2 are arranged in the Clebsch-Gordan coefficients (*cf.* comment of § 2-a). If they are taken in the order j_1, j_2, the component of $|J,\ J\rangle$ along $|j_1, j_2\ ;\ j_1, J-j_1\rangle$ is positive, which means that the sign of the component along $|j_1, j_2\ ; J-j_2, j_2\rangle$ is $(-1)^{j_1+j_2-J}$, as is indicated by (23). On the other hand, if we pick the order j_2, j_1, relation (17) shows that the latter component is positive. Therefore, if we invert j_1 and j_2, the ket $|J, J\rangle$ is multiplied by $(-1)^{j_1+j_2-J}$. The same is true for the kets $|J, M\rangle$, which are constructed from $|J, J\rangle$ by the action of J_- in such a way that the order of j_1 and j_2 plays no role. Finally, the exchange of j_1 and j_2 leads to the relation:

$$\langle j_2, j_1\ ;\ m_2, m_1|J, M\rangle = (-1)^{j_1+j_2-J}\ \langle j_1, j_2\ ;\ m_1, m_2|J, M\rangle \tag{25}$$

3-c. Changing the sign of M, m_1 and m_2

In Chapter X and in this complement, we have constructed all the kets $|J,\ M\rangle$ (and, therefore, the Clebsch-Gordan coefficients) from the kets $|J,\ J\rangle$, by applying the operator J_-. We can take the opposite point of view, and start with the kets $|J, -J\rangle$, using the operator J_+. The reasoning which follows is exactly the same, and we find for the kets $|J, -M\rangle$ the same expansion coefficients on the kets $|j_1,\ j_2;\ -m_1, -m_2\rangle$ as for the $|J, M\rangle$ on the $|j_1,\ j_2;\ m_1,\ m_2\rangle$. The only differences that can appear are related to the phase conventions for the kets $|J,\ M\rangle$, since the analogue of (16) then requires $\langle j_1,\ j_2;\ -j_1, -J+j_1|J, -J\rangle$ to be real and positive. Now, according to (24), the sign of this coefficient is, in reality, $(-1)^{j_1+j_2-J}$. Consequently:

$$\langle j_1,\ j_2\ ;\ -m_1, -m_2|J, -M\rangle = (-1)^{j_1+j_2-J}\ \langle j_1,\ j_2\ ;\ m_1,\ m_2|J, M\rangle \tag{26}$$

In particular, if we set $m_1 = m_2 = 0$, we see that the coefficient $\langle j_1,\ j_2;\ 0, 0|J, 0\rangle$ is zero when $j_1 + j_2 - J$ is an odd number.

3-d. The coefficients $\langle j, j\ ;\ m, -m|0, 0\rangle$

According to (3a), J can be zero only if j_1 and j_2 are equal. We therefore substitute the values $j_1 = j_2 = j$, $m_1 = m$, $m_2 = -m-1$ and $J = M = 0$ into (13); we obtain:

$$\langle j, j\ ;\ m+1, -(m+1)|0, 0\rangle = -\langle j, j\ ;\ m, -m|0, 0\rangle \tag{27}$$

All the coefficients $\langle j, j\ ;\ m, -m|0, 0\rangle$ are therefore equal in modulus. Their signs change whenever m varies by one, and, since $\langle j, j\ ;\ j, -j|0, 0\rangle$ is positive, it is given by $(-1)^{j-m}$. Taking into account orthogonality relation (7b), which indicates that:

$$\sum_{m=-j}^{j} \langle j, j\ ;\ m, -m|0, 0\rangle^2 = 1 \tag{28}$$

we find:

$$\langle j, j\ ;\ m, -m|0, 0\rangle = \frac{(-1)^{j-m}}{\sqrt{2j+1}} \tag{29}$$

References

Messiah (1.17), app. C; Rose (2.19), Chap. III and app. I; Edmonds (2.21), Chap. 3; Sobel'man (11.12), Chap. 4, § 13.

Tables of Clebsch-Gordan coefficients: Condon and Shortley (11.13), Chap. III, § 14; Bacry (10-31), app. C.

Tables of $3j$ and $6j$ coefficients: Edmonds (2.21), Table 2; Rotenberg et al. (10.48).

Complement C_X

Addition of spherical harmonics

1 **The functions $\Phi_J^M(\Omega_1;\Omega_2)$** **1059**
2 **The functions $\mathbf{F}_l^m(\Omega)$** **1060**
3 **Expansion of a product of spherical harmonics; the integral of a product of three spherical harmonics** **1062**

In this complement, we use the properties of Clebsch-Gordan coefficients to prove relations that will be of use to us later, especially in Complements E_X and A_{XIII}: the spherical harmonic addition relations. With this aim in mind, we shall begin by introducing and studying the functions of two sets of polar angles Ω_1 and Ω_2, the $\Phi_J^M(\Omega_1;\Omega_2)$.

1. The functions $\Phi_J^M(\Omega_1;\Omega_2)$

Consider two particles (1) and (2), of state spaces $\mathcal{E}_{\mathbf{r}}^1$ and $\mathcal{E}_{\mathbf{r}}^2$ and orbital angular momenta $\mathbf{L}_1$ and $\mathbf{L}_2$. We choose for the space $\mathcal{E}_{\mathbf{r}}^1$ a standard basis, formed by the kets $\{|\varphi_{k_1,l_1,m_1}\rangle\}$, whose wave functions are:

$$\varphi_{k_1,l_1,m_1}(\mathbf{r}_1) = R_{k_1,l_1}(r_1)\ Y_{l_1}^{m_1}(\Omega_1) \tag{1}$$

(Ω_1 denotes the set of polar angles $\{\theta_1,\varphi_1\}$ of the first particle). Similarly, we choose for $\mathcal{E}_{\mathbf{r}}^2$ a standard basis, $\{|\varphi_{k_2,l_2,m_2}\rangle\}$. In all that follows, we shall confine the states of the two particles to the subspaces $\mathcal{E}(k_1,l_1)$ and $\mathcal{E}(k_2,l_2)$, where k_1, l_1, k_2 and l_2 are fixed; the radial functions and $R_{k_1,l_1}(r_1)$ and $R_{k_2,l_2}(r_2)$ play no role.

The angular momentum of the total system (1) + (2) is:

$$\mathbf{J} = \mathbf{L}_1 + \mathbf{L}_2 \tag{2}$$

According to the results of Chapter X, we can construct a basis of $\mathcal{E}(k_1,l_1)\otimes\mathcal{E}(k_2,l_2)$ of eigenvectors $|\Phi_J^M\rangle$ common to $\mathbf{J}^2$ [eigenvalue $J(J+1)\hbar^2$] and J_z (eigenvalue $M\hbar$). These vectors are of the form:

$$|\Phi_J^M\rangle = \sum_{m_1=-l_1}^{l_1}\sum_{m_2=-l_2}^{l_2} \langle l_1,l_2\ ;\ m_1,m_2|J,M\rangle\, |\varphi_{k_1,l_1,m_1}(1)\rangle\otimes|\varphi_{k_2,l_2,m_2}(2)\rangle \tag{3a}$$

the inverse change of basis being given by:

$$|\varphi_{k_1,l_1,m_1}(1)\rangle\otimes|\varphi_{k_2,l_2,m_2}(2)\rangle = \sum_{J=|l_1-l_2|}^{l_1+l_2}\sum_{M=-J}^{J} \langle l_1,l_2\ ;\ m_1,m_2|J,M\rangle\, |\Phi_J^M\rangle \tag{3b}$$

Relation (3a) shows that the angular dependence of the states $|\Phi_J^M\rangle$ is described by the functions:

$$\Phi_J^M(\Omega_1;\Omega_2) = \sum_{m_1}\sum_{m_2}\langle l_1,l_2\ ;\ m_1,m_2|J,M\rangle\, Y_{l_1}^{m_1}(\Omega_1)Y_{l_2}^{m_2}(\Omega_2) \tag{4a}$$

Similarly, relation (3b) implies that:

$$Y_{l_1}^{m_1}(\Omega_1)Y_{l_2}^{m_2}(\Omega_2) = \sum_{J=|l_1-l_2|}^{l_1+l_2} \sum_{M=-J}^{J} \langle l_1, l_2 \ ; \ m_1, m_2 | J, M \rangle \, \Phi_J^M(\Omega_1; \Omega_2) \tag{4b}$$

To the observables $\mathbf{L}_1$ and $\mathbf{L}_2$ correspond, for the wave functions, differential operators acting on the variables $\Omega_1 = \{\theta_1, \varphi_1\}$ and $\Omega_2 = \{\theta_2, \varphi_2\}$; in particular:

$$L_{1z} \Rightarrow \frac{\hbar}{i} \frac{\partial}{\partial \varphi_1} \tag{5a}$$

$$L_{2z} \Rightarrow \frac{\hbar}{i} \frac{\partial}{\partial \varphi_2} \tag{5b}$$

Since, by construction, the ket $|\Phi_J^M\rangle$ is an eigenvector of $J_z = L_{1z} + L_{2z}$, we can write:

$$\frac{\hbar}{i} \left(\frac{\partial}{\partial \varphi_1} + \frac{\partial}{\partial \varphi_2} \right) \Phi_J^M(\theta_1, \varphi_1; \theta_2, \varphi_2) = M\hbar \ \Phi_J^M(\theta_1, \varphi_1; \theta_2, \varphi_2) \tag{6}$$

Similarly, we have:

$$J_\pm |\Phi_J^M\rangle = \hbar \sqrt{J(J+1) - M(M \pm 1)} \, \left| \Phi_J^{M \pm 1} \right\rangle \tag{7}$$

which implies, with formulas (D-6) of Chapter VI taken into account:

$$\begin{aligned} \left\{ e^{\pm i\varphi_1} \left[\pm \frac{\partial}{\partial \theta_1} + i \cot \theta_1 \frac{\partial}{\partial \varphi_1} \right] \right. & \\ \left. + e^{\pm i\varphi_2} \left[\pm \frac{\partial}{\partial \theta_2} + i \cot \theta_2 \frac{\partial}{\partial \varphi_2} \right] \right\} & \Phi_J^M(\theta_1, \varphi_1; \theta_2, \varphi_2) \\ = \sqrt{J(J+1) - M(M \pm 1)} \ & \Phi_J^{M \pm 1}(\theta_1, \varphi_1; \theta_2, \varphi_2) \end{aligned} \tag{8}$$

2. The functions $\mathbf{F}_l^m(\Omega)$

We now introduce the function F_l^m defined by:

$$F_l^m(\theta, \varphi) \equiv F_l^m(\Omega) = \Phi_{J=l}^{M=m}(\Omega_1 = \Omega \ ; \ \Omega_2 = \Omega) \tag{9}$$

F_l^m is a function of a single pair of polar angles $\Omega = \{\theta, \varphi\}$, and can therefore characterize the angular dependence of a wave function associated with a single particle, of state space $\mathcal{E}_\mathbf{r}$ and angular momentum $\mathbf{L}$. In fact, we shall see that F_l^m is not a new function, but is simply proportional to the spherical harmonic Y_l^m.

To demonstrate this, we shall show that F_l^m is an eigenfunction of $\mathbf{L}^2$ and L_z with the eigenvalues $l(l+1)\hbar^2$ and $m\hbar$. We therefore begin by calculating the action of L_z on F_l^m. According to (9), F_l^m depends on θ and φ by way of $\Omega_1 = \{\theta_1, \varphi_1\}$ and $\Omega_2 = \{\theta_2, \varphi_2\}$, which are both taken equal to Ω. If we apply the differentiation theorem for functions of functions, we find:

$$\begin{aligned} L_z F_l^m(\theta, \varphi) &= \frac{\hbar}{i} \frac{\partial}{\partial \varphi} F_l^m(\theta, \varphi) \\ &= \frac{\hbar}{i} \left\{ \left[\frac{\partial}{\partial \varphi_1} + \frac{\partial}{\partial \varphi_2} \right] \Phi_{J=l}^{M=m}(\Omega_1; \Omega_2) \right\}_{\Omega_1 = \Omega_2 = \Omega} \end{aligned} \tag{10}$$

Relation (6) then yields:

$$L_z F_l^m(\theta,\varphi) = m\hbar\, F_l^m(\theta,\varphi) \tag{11}$$

which proves part of the result being sought. To calculate the action of $\mathbf{L}^2$ on F_l^m, we use the fact that:

$$\mathbf{L}^2 = \frac{1}{2}\left(L_+L_- + L_-L_+\right) + L_z^2 \tag{12}$$

Now, by using an argument analogous to the one that enabled us to write (10) and (11), relation (8) leads to:

$$L_\pm F_l^m(\theta,\varphi) = \hbar\sqrt{l(l+1) - m(m\pm 1)}\; F_l^{m\pm 1}(\theta,\varphi) \tag{13}$$

With this, (12) then yields:

$$\begin{aligned}\mathbf{L}^2 F_l^m(\theta,\varphi) &= \frac{\hbar^2}{2}\left\{[l(l+1) - m(m-1)] + [l(l+1) - m(m+1)] + 2m^2\right\} F_l^m(\theta,\varphi)\\ &= l(l+1)\hbar^2\, F_l^m(\theta,\varphi)\end{aligned} \tag{14}$$

F_l^m, which, according to (11), is an eigenfunction of L_z with the eigenvalue $m\hbar$, is therefore also an eigenfunction of $\mathbf{L}^2$ with the eigenvalue $l(l+1)\hbar^2$. Since $\mathbf{L}^2$ and L_z form a C.S.C.O. in the space of functions of θ and φ alone, F_l^m is necessarily proportional to the spherical harmonic Y_l^m. Relation (13) enables us to show easily that the proportionality coefficient does not depend on m, and we find:

$$F_l^m(\theta,\varphi) = \lambda(l)\; Y_l^m(\theta,\varphi) \tag{15}$$

We must now calculate this proportionality coefficient $\lambda(l)$. To do so, we shall choose a particular direction in space, the Oz direction ($\theta = 0$, φ indeterminate). In this direction, all the spherical harmonics Y_l^m are zero [1], except those corresponding to $m = 0$. When $m = 0$, the spherical harmonic $Y_l^m(\theta = 0, \varphi)$ is given by [*cf.* Complement $\mathrm{A_{VI}}$, relations (57) and (60)]:

$$Y_l^0(\theta = 0, \varphi) = \sqrt{\frac{2l+1}{4\pi}} \tag{16}$$

Substituting these results into (4a) and (9), we find:

$$F_l^{m=0}(\theta = 0, \varphi) = \langle l_1, l_2\ ;\ 0,\ 0 | l, 0\rangle\, \frac{\sqrt{(2l_1+1)(2l_2+1)}}{4\pi} \tag{17}$$

Furthermore, according to (15) and (16):

$$F_l^{m=0}(\theta = 0, \varphi) = \lambda(l)\sqrt{\frac{2l+1}{4\pi}} \tag{18}$$

We therefore have:

$$\lambda(l) = \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}}\, \langle l_1, l_2\ ;\ 0,\ 0 | l, 0\rangle \tag{19}$$

[1] Since Y_l^m is proportional to $\mathrm{e}^{im\varphi}$, they must be zero for the value of Y_l^m in the Oz direction to be defined uniquely; to see this, set $\theta = 0$ in (66), (67) and (69) of Complement $\mathrm{A_{VI}}$.

3. Expansion of a product of spherical harmonics; the integral of a product of three spherical harmonics

With (9), (15) and (19) taken into account, relations (4a) and (4b) imply that:

$$Y_l^m(\Omega) = \left[\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}}\,\langle l_1, l_2\ ;\ 0,0|l,0\rangle\right]^{-1} \times \sum_{m_1}\sum_{m_2}\langle l_1, l_2\ ;\ m_1, m_2|l,m\rangle\, Y_{l_1}^{m_1}(\Omega) Y_{l_2}^{m_2}(\Omega) \tag{20}$$

and:

$$Y_{l_1}^{m_1}(\Omega) Y_{l_2}^{m_2}(\Omega) = \sum_{l=|l_1-l_2|}^{l_1+l_2}\ \sum_{m=-l}^{l} \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}}\,\langle l_1, l_2\ ;\ 0,0|l,0\rangle \times \langle l_1, l_2\ ;\ m_1, m_2|l,m\rangle\, Y_l^m(\Omega) \tag{21}$$

This last relation (in which the summation over m is actually unnecessary, since the only non-zero terms necessarily satisfy $m = m_1 + m_2$) is called *the spherical harmonic addition relation*[2]. According to formula (26) of Complement B_X, the Clebsch-Gordan coefficient $\langle l_1, l_2\ ;\ 0,0|l,0\rangle$ is different from zero only if $l_1 + l_2 - l$ is even. The product $Y_{l_1}^{m_1}(\Omega) Y_{l_2}^{m_2}(\Omega)$ can therefore be expanded only in terms of spherical harmonics of orders:

$$l = l_1 + l_2\ \ ,\quad l_1 + l_2 - 2\ \ ,\quad l_1 + l_2 - 4\ \ ,\ ...,\ |l_1 - l_2| \tag{22}$$

In (21), the parity $(-1)^l$ of all the terms of the expansion on the right-hand side is thus indeed equal to $(-1)^{l_1+l_2}$, the parity of the product which constitutes the left-hand side.

We can use the spherical harmonic addition relation to calculate the integral:

$$I = \int Y_{l_1}^{m_1}(\Omega)\ Y_{l_2}^{m_2}(\Omega)\ Y_{l_3}^{m_3}(\Omega)\ \mathrm{d}\Omega \tag{23}$$

Substituting (21) into (23), we find expressions of the type:

$$K(l, m\ ;\ l_3, m_3) = \int Y_l^m(\Omega)\ Y_{l_3}^{m_3}(\Omega)\ \mathrm{d}\Omega \tag{24}$$

which, with the spherical harmonic complex conjugation relations and orthogonality relations taken into account [*cf.* Complement A_{VI}, relations (55) and (45)], are equal to:

$$K(l, m\ ;\ l_3, m_3) = (-1)^m\ \delta_{ll_3}\ \delta_{m,-m_3} \tag{25}$$

The value of I is therefore:

$$\int Y_{l_1}^{m_1}(\Omega) Y_{l_2}^{m_2}(\Omega)\, Y_{l_3}^{m_3}(\Omega)\,\mathrm{d}\Omega = (-1)^{m_3}\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l_3+1)}} \times \langle l_1, l_2\ ;\ 0,\ 0|l_3, 0\rangle\,\langle l_1, l_2\ ;\ m_1, m_2|l_3, -m_3\rangle \tag{26}$$

This integral is, consequently, different from zero only if:

[2] In the particular case in which $l_2 = 1$, $m_2 = 0[Y_1^0(\theta,\varphi) \propto \cos\theta]$, it yields formula (35) of Complement A_{VI}.

(*i*) $m_1 + m_2 + m_3 = 0$, as could have been predicted directly, since the integral over φ in (23) is $\int_0^{2\pi} \mathrm{d}\varphi\, \mathrm{e}^{i(m_1+m_2+m_3)\varphi} = \delta_{0,m_1+m_2+m_3}$.

(*ii*) a triangle can be formed with three line segments of lengths l_1, l_2 and l_3.

(*iii*) $l_1 + l_2 - l_3$ is even (necessary for $\langle l_1, l_2\ ;\ 0, 0 | l_3, 0\rangle$ to be different from zero), that is, if the product of the three spherical harmonics $Y_{l_1}^{m_1}$, $Y_{l_2}^{m_2}$ and $Y_{l_3}^{m_3}$ is an even function (obviously a necessary condition for its integral over all directions of space to be different from zero).

Relation (26) expresses, for the particular case of the spherical harmonics, a more general theorem, called the Wigner-Eckart theorem (*cf.* Complement $\mathrm{D_X}$).

Complement D_X

Vector operators: the Wigner-Eckart theorem

1 Definition of vector operators; examples 1066

2 The Wigner-Eckart theorem for vector operators 1067

2-a Non-zero matrix elements of V in a standard basis 1067

2-b Proportionality between the matrix elements of J and V inside a subspace $\mathcal{E}(k,j)$ 1068

2-c Calculation of the proportionality constant; the projection theorem 1070

3 Application: calculation of the Landé g_J factor of an atomic level 1072

3-a Rotational degeneracy; multiplets 1072

3-b Removal of the degeneracy by a magnetic field; energy diagram 1073

In Complement B_{VI} (*cf.* § 5-b), we defined the concept of a scalar operator: it is an operator A which commutes with the angular momentum $\mathbf{J}$ of the system under study. An important property of these operators was then given (*cf.* § 6-c-β of that complement): in a standard basis, $\{|k,j,m\rangle\}$, the non-zero matrix elements $\langle k,j,m|A|k',j',m'\rangle$ of a scalar operator must satisfy the conditions $j=j'$ and $m=m'$; in addition, these elements do not depend[1] on m, which allows us to write:

$$\langle k,j,m|A|k',j',m'\rangle = a_j(k,k')\ \delta_{jj'}\,\delta_{mm'} \tag{1}$$

In particular, if the values of k and j are fixed, which amounts to considering the "restriction" of A (*cf.* Complement B_{II}, § 3) to the subspace $\mathcal{E}(k,j)$ spanned by the $(2j+1)$ kets $|k,j,m\rangle$ $(m=-j,-j+1,...,+j)$, we obtain a very simple $(2j+1)\times(2j+1)$ matrix: it is diagonal and all its elements are equal.

Now consider another scalar operator B. The matrix corresponding to it in the subspace $\mathcal{E}(k,j)$ possesses the same property: it is proportional to the unit matrix. Therefore, the matrix corresponding to B can easily be obtained from the one associated with A, by multiplying all the (diagonal) elements by the same constant. We therefore see that the restrictions of two scalar operators A and B to a subspace $\mathcal{E}(k,j)$ are always proportional. Denoting by $P(k,j)$ the projector onto the subspace $\mathcal{E}(k,j)$, we can write this result in the form[2]:

$$P(k,j)\,B\,P(k,j) = \lambda(k,j)\,P(k,j)\,A\,P(k,j) \tag{2}$$

The aim of this complement is to study another type of operators that possesses properties analogous to the ones just recalled: vector operators. We shall see that if

[1] The proof of these properties was outlined in Complement B_{VI}. We shall return to this point in this complement (§ 3-a) when we study the matrix elements of a scalar Hamiltonian.

[2] For two given operators A and B, the proportionality coefficient generally depends on the subspace $\mathcal{E}(k,j)$ chosen; this is why we write $\lambda(k,j)$.

$\mathbf{V}$ and $\mathbf{V}'$ are vectorial, their matrix elements also obey selection rules, which we shall establish. Moreover, we shall show that the restrictions of $\mathbf{V}$ and $\mathbf{V}'$ to $\mathcal{E}(k,j)$ are always proportional:

$$P(k,j)\,\mathbf{V}'\,P(k,j) = \mu(k,j)\,P(k,j)\,\mathbf{V}\,P(k,j) \tag{3}$$

These results constitute the Wigner-Eckart theorem for vector operators.

Comment:

Actually, the Wigner-Eckart theorem is much more general. For example, it enables us to obtain selection rules for the matrix elements of $\mathbf{V}$ between two kets belonging to two different subspaces $\mathcal{E}(k,j)$ and $\mathcal{E}(k',j')$, or to relate these elements to the corresponding elements of $\mathbf{V}'$. The Wigner-Eckart theorem can also be applied to a whole class of operators, of which scalars and vectors merely represent special cases: the irreducible tensor operators (*cf.* exercise 8 of Complement G_X), which we shall not treat here.

1. Definition of vector operators; examples

In § 5-c of Complement B_{VI}, we showed that an observable $\mathbf{V}$ is a vector if its three components V_x, V_y and V_z in an orthonormal frame $Oxyz$ satisfy the following commutation relations:

$$[J_x, V_x] = 0 \tag{4a}$$
$$[J_x, V_y] = i\hbar V_z \tag{4b}$$
$$[J_x, V_z] = -i\hbar V_y \tag{4c}$$

as well as those obtained by cyclic permutation of the indices x, y and z.

To give an idea of what this means, we shall give some examples of vector operators.

(*i*) The angular momentum $\mathbf{J}$ is itself a vector; replacing $\mathbf{V}$ by $\mathbf{J}$ in formulas (4), we simply obtain the relations that define an angular momentum (*cf.* Chap. VI).

(*ii*) For a spinless particle whose state space is $\mathcal{E}_\mathbf{r}$, we have $\mathbf{J} = \mathbf{L}$. It is then simple to show that $\mathbf{R}$ and $\mathbf{P}$ are vector operators. We have, for example:

$$\begin{aligned} &[L_x, X] = [YP_z - ZP_y, X] = 0 \\ &[L_x, Y] = [-ZP_y, Y] = i\hbar Z \\ &[L_x, Z] = [YP_z, Z] = -i\hbar Y \end{aligned} \tag{5}$$

(*iii*) For a particle of spin $\mathbf{S}$, whose state space is $\mathcal{E}_\mathbf{r} \otimes \mathcal{E}_s$, $\mathbf{J}$ is given by $\mathbf{J} = \mathbf{L} + \mathbf{S}$. In this case, the operators $\mathbf{L}$, $\mathbf{S}$, $\mathbf{R}$, $\mathbf{P}$ are vectors. If we take into account the fact that all the spin operators (which act only in $\mathcal{E}_s$) commute with the orbital operators (which act only in $\mathcal{E}_\mathbf{r}$), the proof of these properties follows immediately from (*i*) and (*ii*).

On the other hand, operators of the type $\mathbf{L}^2$, $\mathbf{L}\cdot\mathbf{S}$, etc., are not vectors, but scalars [*cf.* comment (*i*) of Complement B_{VI}, § 5-c]. Other vector operators could, however, be constructed from those we have mentioned: $\mathbf{R} \times \mathbf{S}$, $(\mathbf{L}\cdot\mathbf{S})\,\mathbf{P}$, etc.

(*iv*) Consider the system (1) + (2), formed by the union of two systems: (1), of state space $\mathcal{E}_1$, and (2), of state space $\mathcal{E}_2$. If $\mathbf{V}(1)$ is an operator that acts only in $\mathcal{E}_1$, and

if this operator is a vector [that is, satisfies commutation relations (4) with the angular momentum $\mathbf{J}_1$ of the first system], then the extension of $\mathbf{V}(1)$ into $\mathcal{E}_1 \otimes \mathcal{E}_2$ is also a vector. For example, for a two-electron system, the operators $\mathbf{L}_1$, $\mathbf{R}_1$, $\mathbf{S}_2$, etc. are vectors.

2. The Wigner-Eckart theorem for vector operators

2-a. Non-zero matrix elements of V in a standard basis

We introduce the operators V_+, V_-, J_+ and J_- defined by:

$$\begin{aligned} V_\pm &= V_x \pm iV_y \\ J_\pm &= J_x \pm iJ_y \end{aligned} \tag{6}$$

Using relations (4), we can easily show that:

$$[J_x, V_\pm] = \mp\hbar V_z \tag{7a}$$
$$[J_y, V_\pm] = -i\hbar V_z \tag{7b}$$
$$[J_z, V_\pm] = \pm\hbar V_\pm \tag{7c}$$

from which we can deduce the commutation relations of $J_\pm$ and $V_\pm$:

$$[J_+, V_+] = 0 \tag{8a}$$
$$[J_+, V_-] = 2\hbar V_z \tag{8b}$$
$$[J_-, V_+] = -2\hbar V_z \tag{8c}$$
$$[J_-, V_-] = 0 \tag{8d}$$

Now consider the matrix elements of $\mathbf{V}$ in a standard basis. We shall see that the fact that $\mathbf{V}$ is a vector implies that a large number of them are zero. First of all, we shall show that the matrix elements $\langle k, j, m|V_z|k', j', m'\rangle$ are necessarily zero whenever m is different from m'. It suffices to note that V_z and J_z commute [which follows, after cyclic permutation of the indices x, y and z, from relation (4a)]. Therefore, the matrix elements of V_z between two vectors $|k, j, m\rangle$ corresponding to different eigenvalues $m\hbar$ of J_z are zero (*cf.* Chap. II, § D-3-a-β).

For the matrix elements $\langle k, j, m|V_\pm|k', j', m'\rangle$ of $V_\pm$, we shall show that they are different from zero only if $m - m' = \pm 1$. Equation (7c) indicates that:

$$J_z V_\pm = V_\pm J_z \pm \hbar V_\pm \tag{9}$$

Applying both sides of this relation to the ket $|k', j', m'\rangle$, we obtain:

$$\begin{aligned} J_z(V_\pm|k', j', m'\rangle) &= V_\pm J_z\,|k', j', m'\rangle \pm \hbar V_\pm\,|k', j', m'\rangle \\ &= (m' \pm 1)\hbar V_\pm\,|k', j', m'\rangle \end{aligned} \tag{10}$$

This relation indicates that $V_\pm\,|k', j', m'\rangle$ is an eigenvector[3] of J_z with the eigenvalue $(m' \pm 1)\hbar$. Since two eigenvectors of the Hermitian operator J_z associated with different

[3]It should not be concluded that $V_\pm\,|k, j, m\rangle$ is necessarily proportional to $|k, j, m \pm 1\rangle$. In fact, the argument we have given shows only that:

$$V_\pm\,|k, j, m\rangle = \sum_{k'} \sum_{j'} c_{k', j'}\,\big|k', j', m \pm 1\big\rangle\,.$$

For us to be able to omit, for example, the summation over j', it would be necessary for $V_\pm$ to commute with $\mathbf{J}^2$, which is not generally the case.

eigenvalues are orthogonal, it follows that the scalar product $\langle k, j, m|V_{\pm}|k', j', m'\rangle$ is zero if $m \neq m' \pm 1$.

Summing up, the selection rules obtained for the matrix elements of $\mathbf{V}$ are as follows:

$$V_z \Longrightarrow \Delta m = m - m' = 0 \tag{11a}$$

$$V_+ \Longrightarrow \Delta m = m - m' = +1 \tag{11b}$$

$$V_- \Longrightarrow \Delta m = m - m' = -1 \tag{11c}$$

From these results, we can easily deduce the forms of the matrices that represent the restrictions of the components of $\mathbf{V}$ inside a subspace $\mathcal{E}(k, j)$. The one associated with V_z is diagonal, and those associated with $V_{\pm}$ have matrix elements only just above and just below the principal diagonal.

2-b. Proportionality between the matrix elements of J and V inside a subspace $\mathcal{E}(k, j)$

α. Matrix elements of V_+ and V_-

Expressing the fact that the matrix element of the commutator (8a) between the bra $\langle k, j, m+2|$ and the ket $|k, j, m\rangle$ is zero, we have:

$$\langle k, j, m+2| J_+V_+ |k, j, m\rangle = \langle k, j, m+2| V_+J_+|k, j, m\rangle \tag{12}$$

On both sides of this relation and between the operators J_+ and V_+, we insert the closure relation:

$$\sum_{k',j',m'} |k', j', m'\rangle \langle k', j', m'| = 1 \tag{13}$$

We thus obtain the matrix elements $\langle k, j, m| J_+ |k', j', m'\rangle$ of J_+; by the very construction of the standard basis $\{|k, j, m\rangle\}$, they are different from zero only if $k = k'$, $j = j'$ and $m = m' + 1$. The summations over k', j' and m' are therefore unnecessary in this case, and (12) can be written:

$$\begin{aligned}\langle k, j, m+2| J_+ |k, j, m+1\rangle \langle k, j, m+1| V_+ |k, j, m\rangle \\ = \langle k, j, m+2| V_+ |k, j, m+1\rangle \langle k, j, m+1| J_+ |k, j, m\rangle\end{aligned} \tag{14}$$

that is:

$$\frac{\langle k, j, m+1| V_+ |k, j, m\rangle}{\langle k, j, m+1| J_+ |k, j, m\rangle} = \frac{\langle k, j, m+2| V_+ |k, j, m+1\rangle}{\langle k, j, m+2| J_+ |k, j, m+1\rangle} \tag{15}$$

(as long as the bras and kets appearing in this relation exist, that is, as long as $j - 2 \geq m \geq -j$, we can show immediately that neither of the denominators can go to zero). Writing the relation thus obtained for $m = -j, -j+1, \ldots j-2$, we get:

$$\begin{aligned}\frac{\langle k, j, -j+1| V_+ |k, j, -j\rangle}{\langle k, j, -j+1| J_+ |k, j, -j\rangle} &= \frac{\langle k, j, -j+2| V_+ |k, j, -j+1\rangle}{\langle k, j, -j+2| J_+ |k, j, -j+1\rangle} = \ldots \\ &= \frac{\langle k, j, m+1| V_+ |k, j, m\rangle}{\langle k, j, m+1| J_+ |k, j, m\rangle} = \ldots \\ &= \frac{\langle k, j, j| V_+ |k, j, j-1\rangle}{\langle k, j, j| J_+ |k, j, j-1\rangle}\end{aligned} \tag{16}$$

that is, if we call $\alpha_+(k,j)$ the common value of these ratios:

$$\langle k,j,m+1|\,V_+\,|k,j,m\rangle = \alpha_+(k,j)\,\langle k,j,m+1|\,J_+\,|k,j,m\rangle \tag{17}$$

where $\alpha_+(k,j)$ depends on k and on j, but not on m.

Selection rule (11b) implies that all the matrix elements $\langle k,j,m|\,V_+\,|k,j,m'\rangle$ and $\langle k,j,m|\,J_+\,|k,j,m'\rangle$ are zero if $\Delta m = m - m' \neq +1$. Therefore, whatever m and m', we have:

$$\langle k,j,m|\,V_+\,|k,j,m'\rangle = \alpha_+(k,j)\,\langle k,j,m|\,J_+\,|k,j,m'\rangle \tag{18a}$$

This result expresses the fact that all the matrix elements of V_+ inside $\mathcal{E}(k,j)$ are proportional to those of J_+.

An analogous argument can be made by taking the matrix element of the commutator (8d) between the bra $\langle k,j,m-2|$ and the ket $|k,j,m\rangle$ to be zero. We are thus led to:

$$\langle k,j,m|\,V_-\,|k,j,m'\rangle = \alpha_-(k,j)\,\langle k,j,m|\,J_-\,|k,j,m'\rangle \tag{18b}$$

an equation which expresses the fact that the matrix elements of V_- and J_- inside $\mathcal{E}(k,j)$ are proportional.

β. Matrix elements of V_z

To relate the matrix elements of V_z to those of J_z, we now place relation (8c) between the bra $\langle k,j,m|$ and the ket $|k,j,m\rangle$:

$$\begin{aligned}-2\hbar\,\langle k,j,m|\,V_z\,|k,j,m\rangle &= \langle k,j,m|\,(J_-V_+ - V_+J_-)\,|k,j,m\rangle \\ &= \hbar\sqrt{j(j+1)-m(m+1)}\,\langle k,j,m+1|\,V_+\,|k,j,m\rangle \\ &\quad - \hbar\sqrt{j(j+1)-m(m-1)}\,\langle k,j,m|\,V_+\,|k,j,m-1\rangle\end{aligned} \tag{19}$$

Using (18a), we get:

$$\begin{aligned}&\langle k,j,m|\,V_z\,|k,j,m\rangle \\ &\qquad = -\frac{1}{2}\alpha_+(k,j)\left\{\sqrt{j(j+1)-m(m+1)}\,\langle k,j,m+1|\,J_+\,|k,j,m\rangle\right. \\ &\qquad\quad \left. -\sqrt{j(j+1)-m(m-1)}\,\langle k,j,m|\,J_+\,|k,j,m-1\rangle\right\} \\ &\qquad = -\frac{\hbar}{2}\alpha_+(k,j)\,\{j(j+1)-m(m+1)-j(j+1)+m(m-1)\}\end{aligned} \tag{20}$$

that is:

$$\langle k,j,m|\,V_z\,|k,j,m\rangle = m\hbar\,\alpha_+(k,j) \tag{21}$$

Similarly, an analogous argument based on (8b) and (18b) leads to:

$$\langle k,j,m|\,V_z\,|k,j,m\rangle = m\hbar\,\alpha_-(k,j) \tag{22}$$

Relations (21) and (22) show that $\alpha_+(k,j)$ and $\alpha_-(k,j)$ are necessarily equal; from now on, we shall call their common value $\alpha(k,j)$:

$$\alpha(k,j) = \alpha_+(k,j) = \alpha_-(k,j) \tag{23}$$

In addition, these relations imply that:

$$\langle k,j,m|\,V_z\,|k,j,m'\rangle = \alpha(k,j)\,\langle k,j,m|\,J_z\,|k,j,m'\rangle \tag{24}$$

γ. Generalization to an arbitrary component of **V**

Any component of **V** is a linear combination of V_+, V_- and V_z. Consequently, using relation (23), we can summarize (18a), (18b) and (24) by writing:

$$\langle k,j,m|\,\mathbf{V}\,|k,j,m'\rangle = \alpha(k,j)\,\langle k,j,m|\,\mathbf{J}\,|k,j,m'\rangle \tag{25}$$

Therefore, *inside* $\mathcal{E}(k,j)$, *all the matrix elements of* **V** *are proportional to those of* **J**. This result expresses the Wigner-Eckart theorem, for a special case. Introducing the "restrictions" of **V** and **J** to $\mathcal{E}(k,j)$ (*cf.* Complement B_{II}, § 3), we can also write it:

$$P(k,j)\,\mathbf{V}\,P(k,j) = \alpha(k,j)\,P(k,j)\,\mathbf{J}\,P(k,j) \tag{26}$$

Comment:

Operator **J** commutes with $P(k,j)$ [*cf.* (27)]; since, moreover:

$$[P(k,j)]^2 = P(k,j)$$

we can omit either one of the two projectors $P(k,j)$ on the right-hand side of (26).

2-c. Calculation of the proportionality constant; the projection theorem

Consider the operator $\mathbf{J}\cdot\mathbf{V}$; its restriction to $\mathcal{E}(k,j)$ is $P(k,j)\mathbf{J}\cdot\mathbf{V}\,P(k,j)$. To transform this expression, we can use the fact that:

$$[\mathbf{J}, P(k,j)] = \mathbf{0} \tag{27}$$

a relation that can easily be verified by showing that the action of the commutators $[J_z, P(k,j)]$ and $[J_\pm, P(k,j)]$ on any ket of the $\{|k,j,m\rangle\}$ basis yields zero. Using (26), we then get:

$$\begin{aligned} P(k,j)\mathbf{J}\cdot\mathbf{V}\,P(k,j) &= \mathbf{J}.\,[P(k,j)\,\mathbf{V}\,P(k,j)] \\ &= \alpha(k,j)\,\mathbf{J}^2 P(k,j) \\ &= \alpha(k,j)\,j(j+1)\hbar^2\,P(k,j) \end{aligned} \tag{28}$$

The restriction to the space $\mathcal{E}(k,j)$ of the operator $\mathbf{J}\cdot\mathbf{V}$ is therefore equal to the identity operator[4] multiplied by $\alpha(k,j)\,j(j+1)\hbar^2$. Therefore, if $|\psi_{k,j}\rangle$ denotes an arbitrary normalized state belonging to the subspace $\mathcal{E}(k,j)$, the average value $\langle\mathbf{J}\cdot\mathbf{V}\rangle_{k,j}$ of $\mathbf{J}\cdot\mathbf{V}$ is independent of the ket $|\psi_{k,j}\rangle$ chosen, since:

$$\langle\mathbf{J}\cdot\mathbf{V}\rangle_{k,j} = \langle\psi_{k,j}|\mathbf{J}\cdot\mathbf{V}|\psi_{k,j}\rangle = \alpha(k,j)\,j(j+1)\hbar^2 \tag{29}$$

If we substitute this relation into (26), we see that[5], *inside the subspace* $\mathcal{E}(k,j)$:

$$\mathbf{V} = \frac{\langle\mathbf{J}\cdot\mathbf{V}\rangle_{k,j}}{\langle\mathbf{J}^2\rangle_{k,j}}\mathbf{J} = \frac{\langle\mathbf{J}\cdot\mathbf{V}\rangle_{k,j}}{j(j+1)\hbar^2}\mathbf{J} \tag{30}$$

[4] Since $\mathbf{J}\cdot\mathbf{V}$ is a scalar, the fact that its restriction is proportional to the identity operator was to be expected.

[5] We shall say that an operator relation is valid only inside a given subspace when it is actually valid only for the restrictions to this subspace of the operators being considered. To be completely rigorous, we should therefore have to place both sides of relation (30) between two projectors $P(k,j)$.

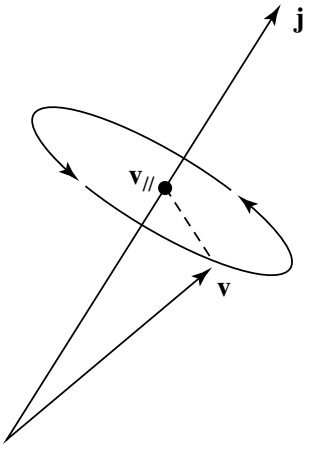


Figure 1: Classical interpretation of the projection theorem: since the vector $\mathbf{v}$ *rotates very rapidly about the total angular momentum* $\mathbf{j}$*, only its static component* $\mathbf{v}_{/\!/}$ *should be taken into account.*

This result is often called the "projection theorem". Whatever the physical system being studied, as long as we are concerned only with states belonging to the same subspace $\mathcal{E}(k,j)$, we can assume that all vector operators are proportional to $\mathbf{J}$.

We can give the following classical physical interpretation of this property (*cf.* § 1 of Complement $\mathrm{F_X}$): if $\mathbf{j}$ denotes the total angular momentum of any isolated physical system, all the physical quantities attached to the system rotate about $\mathbf{j}$, which is a constant vector (*cf.* Fig. 1). In particular, for a vector quantity $\mathbf{v}$, all that remains after averaging over time is its projection $\mathbf{v}_{/\!/}$ onto $\mathbf{j}$, that is, a vector parallel to $\mathbf{j}$, given by:

$$\mathbf{v}_{/\!/} = \frac{\mathbf{j}\cdot\mathbf{v}}{\mathbf{j}^2}\mathbf{j} \tag{31}$$

a formula which is indeed analogous to (30).

Comments:

(*i*) It cannot be deduced from (30) that, in the total state space [the direct sum of all the subspaces $\mathcal{E}(k,j)$], $\mathbf{V}$ and $\mathbf{J}$ are proportional. It must be noted that the proportionality constant $\alpha(k,j)$ (or $\langle \mathbf{J}\cdot\mathbf{V}\rangle_{k,j}$) depends on the subspace $\mathcal{E}(k,j)$ chosen. Moreover, any vector operator $\mathbf{V}$ may possess non-zero matrix elements between kets belonging to different subspaces while the corresponding elements of $\mathbf{J}$ are always zero.

(*ii*) Consider a second vector operator $\mathbf{W}$. Its restriction inside $\mathcal{E}(k,j)$ is proportional to $\mathbf{J}$, and therefore also to the restriction of $\mathbf{V}$. Therefore, *inside a subspace* $\mathcal{E}(k,j)$*, all vector operators are proportional.*

However, to calculate the proportionality coefficient between $\mathbf{V}$ and $\mathbf{W}$, we cannot simply replace $\mathbf{J}$ by $\mathbf{W}$ in (30) (which would give the value $\langle \mathbf{V}\cdot\mathbf{W}\rangle_{k,j} / \langle \mathbf{W}^2\rangle_{k,j}$). In the proof leading to relation (30), we used the fact that $\mathbf{J}$ commutes with $P(k,j)$ in (28), which is not generally the case for $\mathbf{W}$. To calculate this proportionality coefficient correctly, we note that, inside the subspace $\mathcal{E}(k,j)$:

$$\mathbf{W} = \frac{\langle \mathbf{J}\cdot\mathbf{W}\rangle_{k,j}}{\langle \mathbf{J}^2\rangle_{k,j}}\mathbf{J} \tag{32}$$

This yields, with (30) taken into account:

$$\mathbf{V} = \frac{\langle \mathbf{J} \cdot \mathbf{V} \rangle_{k,j}}{\langle \mathbf{J} \cdot \mathbf{W} \rangle_{k,j}} \mathbf{W} \tag{33}$$

3. Application: calculation of the Landé g_J factor of an atomic level

In this section, we shall use the Wigner-Eckart theorem to calculate the effect of a magnetic field $\mathbf{B}$ on the energy levels of an atom. We shall see that this theorem considerably simplifies the calculations and enables us to predict, in a very general way, that the magnetic field removes degeneracies, causing equidistant levels to appear (to first order in B). The energy difference of these states is proportional to B and to a constant g_J (the Landé factor) which we shall calculate.

Let $\mathbf{L}$ be the total orbital angular momentum of the electrons of an atom (the sum of their individual orbital angular momenta $\mathbf{L}_i$), and let $\mathbf{S}$ be their total spin angular momentum (the sum of their individual spins $\mathbf{S}_i$). The total internal angular momentum of the atom (assuming the spin of the nucleus to be zero) is:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{34}$$

In the absence of a magnetic field, we call H_0 the Hamiltonian of the atom; H_0 commutes[6] with $\mathbf{J}$. We shall assume that H_0, $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$ and J_z form a C.S.C.O., and we shall call $|E_0, L, S, J, M\rangle$ their common eigenvectors, of eigenvalues E_0, $L(L+1)\hbar^2$, $S(S+1)\hbar^2$, $J(J+1)\hbar^2$ and $M\hbar$, respectively.

> This hypothesis is valid for a certain number of light atoms for which the angular momentum coupling is of the $\mathbf{L} \cdot \mathbf{S}$ type (*cf.* Complement B_{XIV}). However, for other atoms, which have a different type of coupling (for example, the rare gases other than helium), this is not the case. Calculations based on the Wigner-Eckart theorem, similar to those presented here, can then be performed, and the central physical ideas remain the same. For the sake of simplicity, we shall confine ourselves here to the case in which L and S are actually good quantum numbers for the atomic state under study.

3-a. Rotational degeneracy; multiplets

Consider the ket $J_\pm |E_0, L, S, J, M\rangle$. According to the hypotheses set forth above, $J_\pm$ commutes with H_0; therefore, $J_\pm |E_0, L, S, J, M\rangle$ is an eigenvector of H_0 with the eigenvalue E_0. Furthermore, in accordance with the general properties of angular momenta and their addition, we have:

$$J_\pm |E_0, L, S, J, M\rangle = \hbar\sqrt{J(J+1) - M(M \pm 1)}\, |E_0, L, S, J, M \pm 1\rangle \tag{35}$$

This relation shows that, starting with a state $|E_0, L, S, J, M\rangle$, we can construct others with the same energy: those for which $-J \leq M \leq J$. It follows that the eigenvalue

[6] This general property follows from the invariance of the energy of the atom under a rotation of all the electrons, performed about an axis passing through the origin (which is the position of the nucleus, assumed to be motionless). H_0, which is invariant under rotation, therefore commutes with $\mathbf{J}$ (H_0 is a scalar operator; *cf.* Complement B_{VI}, § 5-b).

E_0 is necessarily at least $(2J+1)$-fold degenerate. This is an essential degeneracy, since it is related to the rotational invariance of H_0 (an accidental degeneracy may also be present). In atomic physics, the corresponding $(2J+1)$-fold degenerate energy level is called a multiplet. The eigensubspace associated with it, spanned by the kets $|E_0, L, S, J, M\rangle$ with $M = J, J-1, ..., -J$, will be written $\mathcal{E}(E_0, L, S, J)$.

3-b. Removal of the degeneracy by a magnetic field; energy diagram

In the presence of a magnetic field $\mathbf{B}$ parallel to Oz, the Hamiltonian becomes (*cf.* Complement D_{VII}):

$$H = H_0 + H_1 \tag{36}$$

with:

$$H_1 = \omega_L(L_z + 2S_z) \tag{37}$$

(the factor 2 before S_z arises from the electron spin gyromagnetic ratio). The "Larmor angular frequency" ω_L of the electron is defined in terms of its mass m and its charge q by:

$$\omega_L = -\frac{qB}{2m} = -\frac{\mu_B}{\hbar}B \tag{38}$$

(where $\mu_B = q\hbar/2m$ is the Bohr magneton).

To calculate the effect of the magnetic field on the energy levels of the atom, we shall consider only the matrix elements of H_1 inside the subspace $\mathcal{E}(E_0, L, S, J)$ associated with the multiplet under study. Perturbation theory, which will be explained in Chapter XI, justifies this procedure when B is not too large.

Inside the subspace $\mathcal{E}(E_0, L, S, J)$, we have, according to the projection theorem (§ 2-c):

$$\mathbf{L} = \frac{\langle \mathbf{L}\cdot\mathbf{J}\rangle_{E_0,L,S,J}}{J(J+1)\hbar^2}\mathbf{J} \tag{39a}$$

$$\mathbf{S} = \frac{\langle \mathbf{S}\cdot\mathbf{J}\rangle_{E_0,L,S,J}}{J(J+1)\hbar^2}\mathbf{J} \tag{39b}$$

where $\langle \mathbf{L}\cdot\mathbf{J}\rangle_{E_0,L,S,J}$ and $\langle \mathbf{S}\cdot\mathbf{J}\rangle_{E_0,L,S,J}$ denote respectively the average values of the operators $\mathbf{L}\cdot\mathbf{J}$ and $\mathbf{S}\cdot\mathbf{J}$ for the states of the system belonging to $\mathcal{E}(E_0, L, S, J)$. Now, we can write:

$$\mathbf{L}\cdot\mathbf{J} = \mathbf{L}\cdot(\mathbf{L}+\mathbf{S}) = \mathbf{L}^2 + \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \tag{40a}$$

as well as:

$$\mathbf{S}\cdot\mathbf{J} = \mathbf{S}\cdot(\mathbf{L}+\mathbf{S}) = \mathbf{S}^2 + \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \tag{40b}$$

It follows that:

$$\langle \mathbf{L}\cdot\mathbf{J}\rangle_{E_0,L,S,J} = L(L+1)\hbar^2 + \frac{\hbar^2}{2}\left[J(J+1) - L(L+1) - S(S+1)\right] \tag{41a}$$

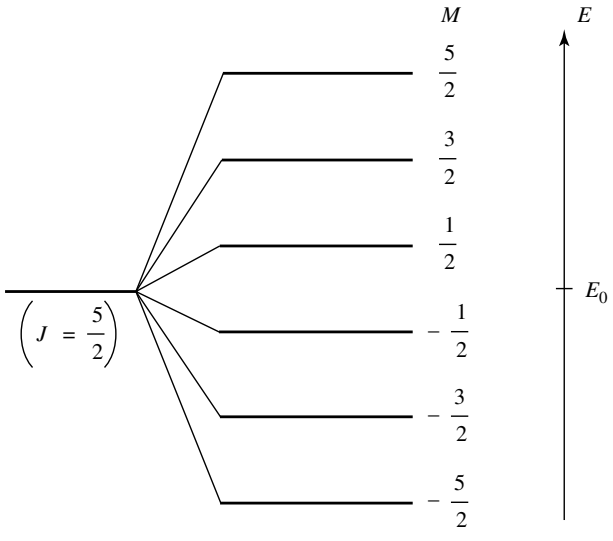


Figure 2: Energy diagram showing the removal of the $(2J + 1)$-fold degeneracy of a multiplet (here $J = 5/2$) by a static magnetic field $\mathbf{B}$. The distance between two adjacent levels is proportional to $|\mathbf{B}|$ and to the Landé g_J factor.

and:

$$\langle \mathbf{S} \cdot \mathbf{J} \rangle_{E_0,L,S,J} = S(S+1)\hbar^2 + \frac{\hbar^2}{2}\left[J(J+1) - L(L+1) - S(S+1)\right] \tag{41b}$$

Relations (41), substituted into (39) and then into (37), show that, inside the subspace $\mathcal{E}(E_0, L, S, J)$, the operator H_1 is given by:

$$H_1 = g_J \omega_L J_z \tag{42}$$

where the Landé g_J factor of the multiplet under consideration is equal to:

$$g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \tag{43}$$

Relation (42) implies that the eigenstates of the Hamiltonian H_1 inside the eigensubspace $\mathcal{E}(E_0, L, S, J)$ are simply the basis vectors $|E_0, L, S, J, M\rangle$, with the eigenvalues:

$$E_1(M) = g_J M \hbar \omega_L \tag{44}$$

We see that the magnetic field completely removes the degeneracy of the multiplet. As is shown by the diagram in Figure 2, a set of $(2J+1)$ equidistant levels appears, each one corresponding to one of the possible values of M. Such a diagram permits generalization of our earlier study of the polarization and frequency of optical lines emitted by a

fictitious atom with a single spinless electron (the "normal" Zeeman effect; *cf.* Complement D_{VII}), to the case of atoms with several electrons whose spins must be taken into account.

References and suggestions for further reading:

Tensor operators: Schiff (1.18), § 28; Messiah (1.17), Chap. XIII, § VI; Edmonds (2.21), Chap. 5; Rose (2.19), Chap. 5; Meijer and Bauer (2.18), Chap. 6.

Complement E_X

Electric multipole moments

1 Definition of multipole moments 1077
1-a Expansion of the potential on the spherical harmonics 1077
1-b Physical interpretation of multipole operators 1079
1-c Parity of multipole operators 1082
1-d Another way to introduce multipole moments 1083
2 Matrix elements of electric multipole moments 1084
2-a General expression for the matrix elements 1085
2-b Selection rules 1087
2-c Physical consequences 1087

Consider a system $\mathscr{S}$ composed of N charged particles placed in a given electrostatic potential $U(\mathbf{r})$. We shall show in this complement how to calculate the interaction energy of the system $\mathscr{S}$ with the potential $U(\mathbf{r})$ by introducing the electric multipole moments of $\mathscr{S}$. First of all, we shall begin by recalling how these moments are introduced in classical physics. Then we shall construct the corresponding quantum mechanical operators, and we shall see how, in a large number of cases, their use considerably simplifies the study of the electrical properties of a quantum mechanical system. This is because these operators possess general properties which are independent of the system being studied, satisfying in particular certain selection rules. For example, if the state of the system $\mathscr{S}$ being studied has an angular momentum j [i.e. is an eigenvector of $\mathbf{J}^2$ with the eigenvalue $j(j+1)\hbar^2$], we shall see that the average values of all multipole operators of order higher than $2j$ are necessarily zero.

1. Definition of multipole moments

1-a. Expansion of the potential on the spherical harmonics

For the sake of simplicity, we begin by studying a system $\mathscr{S}$ composed of a single particle, of charge q and position $\mathbf{r}$, placed in the potential $U(\mathbf{r})$. We shall then generalize the results obtained to N-particle systems.

α. Single particle

In classical physics, the potential energy of the particle is:

$$V(\mathbf{r}) = qU(\mathbf{r}) \tag{1}$$

Since the spherical harmonics form a basis for functions of θ and φ, we can expand $U(\mathbf{r})$ in the form:

$$U(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_{l,m}(r) Y_l^m(\theta, \varphi) \tag{2}$$

We shall assume the charges creating the electrostatic potential to be placed outside the region of space in which the particle being studied can be found. In this whole region, we then have:

$$\Delta U(\mathbf{r}) = 0 \tag{3}$$

Now, we know [*cf.* relation (A-15) of Chapter VII] that the Laplacian Δ is related to the differential operator $\mathbf{L}^2$ acting on the angular variables θ and φ by:

$$\Delta = \frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\mathbf{L}^2}{\hbar^2 r^2} \tag{4}$$

Also, the very definition of the spherical harmonics implies that:

$$\mathbf{L}^2 Y_l^m(\theta,\varphi) = l(l+1)\hbar^2 Y_l^m(\theta,\varphi) \tag{5}$$

It is therefore easy to calculate the Laplacian of expansion (2). If we write, using (3), that each of the terms thus obtained is zero, we get:

$$\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{l(l+1)}{r^2}\right] f_{l,m}(r) = 0 \tag{6}$$

This equation has two linearly independent solutions, r^l and $r^{-(l+1)}$. Since $U(\mathbf{r})$ is not infinite for $r = 0$, we must choose:

$$f_{l,m}(r) = \sqrt{\frac{4\pi}{2l+1}}c_{l,m}r^l \tag{7}$$

where the $c_{l,m}$ are coefficients that depend on the potential under consideration (the factor $\sqrt{4\pi/(2l+1)}$ is introduced for convenience, as will be seen later).

We can therefore write (2) in the form:

$$V(\mathbf{r}) = qU(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{l,m}\mathcal{Q}_l^m(\mathbf{r}) \tag{8}$$

where the functions $\mathcal{Q}_l^m(\mathbf{r})$ are defined by their expressions[1] in spherical coordinates:

$$\mathcal{Q}_l^m(\mathbf{r}) = q\sqrt{\frac{4\pi}{2l+1}}r^l\, Y_l^m(\theta,\varphi) \tag{9}$$

In quantum mechanics, the same type of expansion is possible; the potential energy operator of the particle is $V(\mathbf{R}) = qU(\mathbf{R})$, whose matrix elements in the $\{|\mathbf{r}\rangle\}$ representation are (*cf.* Complement B_{II}, § 4-b):

$$\langle\mathbf{r}|qU(\mathbf{R})|\mathbf{r}'\rangle = qU(\mathbf{r})\,\delta(\mathbf{r}-\mathbf{r}') \tag{10}$$

Expansion (8) then yields:

$$V(\mathbf{R}) = qU(\mathbf{R}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{l,m}Q_l^m \tag{11}$$

[1] Note the difference between the curly capitals of these classical functions and the Q_l^m for the quantum operators.

where the operators Q_l^m are defined by:

$$\begin{aligned}\langle \mathbf{r}|Q_l^m|\mathbf{r}'\rangle &= \mathcal{Q}_l^m(\mathbf{r})\,\delta(\mathbf{r}-\mathbf{r}') \\ &= q\sqrt{\frac{4\pi}{2l+1}}r^l\,Y_l^m(\theta,\varphi)\,\delta(\mathbf{r}-\mathbf{r}')\end{aligned} \tag{12}$$

The Q_l^m are called "electric multipole operators".

β. Generalization to N particles

Now consider N classical particles, with positions $\mathbf{r}_1$, $\mathbf{r}_2$, ..., $\mathbf{r}_N$ and charges q_1, q_2, ..., q_N. Their coupling energy with the external potential $U(\mathbf{r})$ is:

$$V(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N) = \sum_{n=1}^{N} q_n U(\mathbf{r}_n) \tag{13}$$

The argument of the preceding section can immediately be generalized to show that:

$$V(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{l,m}\,\mathcal{Q}_l^m(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N) \tag{14}$$

where the coefficients $c_{l,m}$ [which depend on the potential $U(\mathbf{r})$] have the same values as in the preceding section, and the functions $\mathcal{Q}_l^m$ are defined by their values in polar coordinates:

$$\mathcal{Q}_l^m(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N) = \sqrt{\frac{4\pi}{2l+1}}\sum_{n=1}^{N} q_n (r_n)^l\,Y_l^m(\theta_n,\varphi_n) \tag{15}$$

(θ_n and φ_n are the polar angles of $\mathbf{r}_n$). The multipole moments of the total system are therefore simply the sums of the moments associated with each of the particles.

Similarly, in quantum mechanics, the coupling energy of the N particles with the external potential is described by the operator:

$$V(\mathbf{R}_1,\mathbf{R}_2,...,\mathbf{R}_N) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{l,m}\,Q_l^m \tag{16}$$

with:

$$\begin{aligned}&\langle \mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N|Q_l^m|\mathbf{r}'_1,\mathbf{r}'_2,...,\mathbf{r}'_N\rangle \\ &\quad= \mathcal{Q}_l^m(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N)\,\delta(\mathbf{r}_1-\mathbf{r}'_1)\,\delta(\mathbf{r}_2-\mathbf{r}'_2)...\,\delta(\mathbf{r}_N-\mathbf{r}'_N)\end{aligned} \tag{17}$$

1-b. Physical interpretation of multipole operators

α. The operator Q_0^0; the total charge of the system

Since Y_0^0 is a constant ($Y_0^0 = 1/\sqrt{4\pi}$), definition (15) implies that:

$$\mathcal{Q}_0^0 = \sum_{n=1}^{N} q_n \tag{18}$$

The operator Q_0^0 is therefore a constant, equal to the total charge of the system.

The first term of expansion (14) therefore gives the coupling energy of the system with the potential $U(\mathbf{r})$, assuming all the particles to be situated at the origin O. This is obviously a good approximation if $U(\mathbf{r})$ does not vary very much in relative value over distances comparable to those separating the various particles from O (if the system $\mathscr{S}$ is centered at O, this distance is of the order of the dimensions of $\mathscr{S}$). Furthermore, there exists a special case in which expansion (14) is rigorously given by its first term: the case where the potential $U(\mathbf{r})$ is uniform, and therefore proportional to the spherical harmonic $l = 0$.

β. The operators Q_1^m; the electric dipole moment

According to (15) and the expression for the spherical harmonics Y_l^m [*cf.* Complement A_{VI}, equations (32)], we have:

$$\begin{cases} \mathcal{Q}_1^1 = -\dfrac{1}{\sqrt{2}} \sum\limits_n q_n \, (x_n + i y_n) \\ \mathcal{Q}_1^0 = \sum\limits_n q_n \, z_n \\ \mathcal{Q}_1^{-1} = \dfrac{1}{\sqrt{2}} \sum\limits_n q_n \, (x_n - i y_n) \end{cases} \tag{19}$$

These three quantities can be considered to be the components of a vector on the complex basis of three vectors $\mathbf{e}_1$, $\mathbf{e}_0$ and $\mathbf{e}_{-1}$:

$$\boldsymbol{\mathcal{D}} = -\mathcal{Q}_1^{-1}\mathbf{e}_1 + \mathcal{Q}_1^0\mathbf{e}_0 - \mathcal{Q}_1^1\mathbf{e}_{-1} \tag{20}$$

with:

$$\mathbf{e}_1 = -\frac{1}{\sqrt{2}}(\mathbf{e}_x + i\mathbf{e}_y); \quad \mathbf{e}_0 = \mathbf{e}_z; \quad \mathbf{e}_{-1} = \frac{1}{\sqrt{2}}(\mathbf{e}_x - i\mathbf{e}_y) \tag{21}$$

(where $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$ are the unit vectors of the Ox, Oy and Oz axes). The components of this vector $\boldsymbol{\mathcal{D}}$ on the $Oxyz$ axes are then:

$$\begin{aligned} \mathcal{Q}_1^x &= \frac{1}{\sqrt{2}} \left[\mathcal{D}_1^{-1} - \mathcal{D}_1^1\right] = \sum_n q_n \, x_n \\ \mathcal{Q}_1^y &= \frac{i}{\sqrt{2}} \left[\mathcal{D}_1^{-1} + \mathcal{D}_1^1\right] = \sum_n q_n \, y_n \\ \mathcal{Q}_1^z &= \mathcal{D}_1^0 \qquad\qquad\quad = \sum_n q_n \, z_n \end{aligned} \tag{22}$$

We recognize the three components of the total electric dipole moment of the system $\mathscr{S}$ with respect to the origin O:

$$\boldsymbol{\mathcal{D}} = \sum_{n=1}^{N} q_n \, \mathbf{r}_n \tag{23}$$

The operators Q_1^m are therefore actually the components of the electric dipole $\mathbf{D} = \sum_n q_n \, \mathbf{R}_n$.

Relations (19) enable us, moreover, to write the $l = 1$ terms of expansion (14) in the form:

$$\sum_{m=-1}^{+1} c_{1,m}\, \mathcal{Q}_1^m = -\frac{1}{\sqrt{2}}(c_{1,1} - c_{1,-1}) \sum_n q_n\, x_n - \frac{i}{\sqrt{2}}(c_{1,1} + c_{1,-1}) \sum_n q_n\, y_n \quad + \quad c_{1,0} \sum_n q_n\, z_n \tag{24}$$

We shall now show that the combinations of the coefficients $c_{1,m}$ that appear in this expression are none other than the components of the gradient of the potential $U(\mathbf{r})$ at $\mathbf{r} = \mathbf{0}$. If we take the gradient of expansion (8) of $U(\mathbf{r})$, the $l = 0$ term (which is constant) disappears; the $l = 1$ term can be put into a form analogous to (24) and yields:

$$[\boldsymbol{\nabla} U(\mathbf{r})]_{\mathbf{r}=\mathbf{0}} = -\frac{1}{\sqrt{2}}(c_{1,1} - c_{1,-1})\ \mathbf{e}_x - \frac{i}{\sqrt{2}}(c_{1,1} + c_{1,-1})\ \mathbf{e}_y + c_{1,0}\ \mathbf{e}_z \tag{25}$$

As for the $l > 1$ terms of (8), they are polynomials in x, y, z of degree higher than 1 (*cf.* §§ γ and δ below) which make no contribution to the gradient at $\mathbf{r} = \mathbf{0}$. The $l = 1$ term of expansion (14) can therefore be written, using (23) and (25):

$$\left(\sum_{n=1}^{N} q_n\, \mathbf{r}_n\right) \cdot (\boldsymbol{\nabla} U)_{\mathbf{r}=0} = -\boldsymbol{\mathcal{D}} \cdot \boldsymbol{\mathcal{E}}(\mathbf{r} = \mathbf{0}) \tag{26}$$

where:

$$\boldsymbol{\mathcal{E}}(\mathbf{r}) = -\boldsymbol{\nabla} U(\mathbf{r}) \tag{27}$$

is the electric field at point $\mathbf{r}$. Thus we recognize (26) as the well-known expression for the coupling energy between an electric dipole and the field $\boldsymbol{\mathcal{E}}$.

Comments:

(*i*) In physics, we often deal with systems whose total charge is zero (atoms, for example). $\mathcal{Q}_0^0$ is then equal to zero, and the first multipole operator appearing in expansion (14) is the electric dipole moment. This expansion can often be limited to the $l = 1$ terms [hence expression (26)], since the terms for which $l \geq 2$ are generally much smaller (this is the case, for example, if the electric field varies little over distances comparable to the distances of the particles from the origin; the $l \geq 2$ terms are, furthermore, rigorously zero in a special case: the case in which the electric field is uniform [*cf.* §§ γ and δ below)].

(*ii*) For a system $\mathscr{S}$ composed of two particles of opposite charge $+q$ and $-q$ (an electric dipole), the dipole moment $\boldsymbol{\mathcal{D}}$ is:

$$\boldsymbol{\mathcal{D}} = q(\mathbf{r}_1 - \mathbf{r}_2) \tag{28}$$

Its value is related to the position of the "relative particle" (*cf.* Chap. VII, § B) associated with the system $\mathscr{S}$; it therefore does not depend on the choice of the origin O. Actually this is a more general property: it is simple to show that the electric dipole moment of any electrically neutral system $\mathscr{S}$ is independent of the origin O chosen.

γ. The operators Q_2^m; the electric quadrupole moment

Using the explicit expression for the Y_2^m [*cf.* Complement A_{VI}, relations (33)], we could show without difficulty that:

$$\begin{cases} \mathcal{Q}_2^{\pm 2} = \dfrac{\sqrt{6}}{4} \sum\limits_n q_n \, (x_n \pm i y_n)^2 \\ \mathcal{Q}_2^{\pm 1} = \mp \dfrac{\sqrt{6}}{2} \sum\limits_n q_n \, z_n (x_n \pm i y_n) \\ \mathcal{Q}_2^{0} = \dfrac{1}{2} \sum\limits_n q_n \, (3 z_n^2 - r_n^2) \end{cases} \tag{29}$$

In this way, we obtain the five components of the electric quadrupole moment of the system $\mathscr{S}$. While the total charge of $\mathscr{S}$ is a scalar, and its dipole moment is a vector $\boldsymbol{\mathcal{D}}$, it can be shown that the quadrupole moment is a second-rank tensor. In addition, an argument similar to the one in § β would enable us to write the $l = 2$ terms of expansion (14) in the form:

$$\sum_{m=-2}^{+2} c_{2,m} \mathcal{Q}_2^m = \sum_{i,j} \left[\frac{\partial^2 U}{\partial x^i \, \partial x^j} \right]_{\mathbf{r}=\mathbf{0}} \sum_{n=1}^{N} q_n \, x_n^i x_n^j \tag{30}$$

(with x^i, $x^j = x, y$ or z). These terms describe the coupling between the electric quadrupole moment of the system $\mathscr{S}$ and the gradient of the field $\boldsymbol{\mathcal{E}}(\mathbf{r})$ at point $\mathbf{r} = \mathbf{0}$.

δ. Generalization: the electric l-pole moment

We could generalize the preceding arguments and show from the general expression for the spherical harmonics [*cf.* Complement A_{VI}, relations (26) or (30)] that:

- the quantities $\mathcal{Q}_l^m$ are polynomials (which are homogeneous in x, y, z) of degree l.
- the contribution to expansion (14) of the l terms involves lth order derivatives of the potential $U(\mathbf{r})$, evaluated at $\mathbf{r} = \mathbf{0}$.

Expression (14) for the potential can thus be seen to be a Taylor series expansion in the neighborhood of the origin. As the variation of the potential $U(\mathbf{r})$ in the region about $\mathscr{S}$ becomes more complicated, higher order terms must be retained in the expansion. For example, if $U(\mathbf{r})$ is constant, we have seen that the $l = 0$ term is the only one involved. If the field $\boldsymbol{\mathcal{E}}(\mathbf{r})$ is uniform, the $l = 1$ terms must be added to the expansion. If it is the gradient of the field $\boldsymbol{\mathcal{E}}$ that is uniform, we have $l \leq 2$, and so forth.

1-c. Parity of multipole operators

Finally, we shall consider the parity of the Q_l^m. We know that the parity of Y_l^m is $(-1)^l$ [*cf.* Chap. VI, relation (D-28)]. Therefore (*cf.* Complement F_{II}, § 2-a), the electric multipole operator Q_l^m has a definite parity, equal to $(-1)^l$, independent of m. This property will prove useful in what follows.

1-d. Another way to introduce multipole moments

We shall consider the same system of N charged particles as in § 1-a. However, instead of considering the interaction energy of this system with a given external potential $U(\mathbf{r})$, we shall try to calculate the potential $W(\boldsymbol{\rho})$ created by these charges at a distant point $\boldsymbol{\rho}$ (*cf.* Fig. 1). For the sake of simplicity, we shall use classical mechanics to treat this problem. The potential $W(\boldsymbol{\rho})$ is then:

$$W(\boldsymbol{\rho}) = \frac{1}{4\pi\varepsilon_0} \sum_{n=1}^{N} \frac{q_n}{|\boldsymbol{\rho} - \mathbf{r}_n|} \tag{31}$$

Now, when $|\boldsymbol{\rho}| > |\mathbf{r}_n|$, it can be shown that:

$$\frac{1}{|\boldsymbol{\rho} - \mathbf{r}_n|} = \frac{1}{\rho} \sum_{l=0}^{\infty} \left(\frac{r_n}{\rho} \right)^l P_l(\cos\, \alpha_n) \tag{32}$$

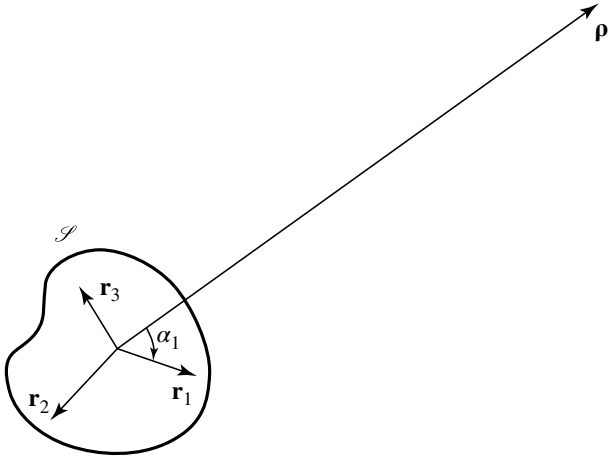


Figure 1: The potential $W(\boldsymbol{\rho})$ created at a distant point by a system $\mathscr{S}$ composed of N charged particles (of positions $\mathbf{r}_1, \mathbf{r}_2$, ...) can be expressed in terms of the multipole moments of $\mathscr{S}$.

where α_n denotes the angle $(\boldsymbol{\rho}, \mathbf{r}_n)$, and P_l is the lth-order Legendre polynomial. Using the spherical harmonic addition theorem (*cf.* Complement A_{VI}, § 2-e-γ), we can write:

$$P_l(\cos\, \alpha_n) = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} (-1)^m\, Y_l^{-m}(\theta_n, \varphi_n)\, Y_l^m(\Theta, \Phi) \tag{33}$$

(where Θ and Φ denote the polar angles of $\boldsymbol{\rho}$). Substituting (32) and (33) into (31), we finally obtain:

$$W(\boldsymbol{\rho}) = \frac{1}{4\pi\varepsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sqrt{\frac{4\pi}{2l+1}} (-1)^m\, \mathcal{Q}_l^{-m} \frac{1}{\rho^{l+1}} Y_l^m(\Theta, \Phi) \tag{34}$$

where $\mathcal{Q}_l^m(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ is defined by relation (15).

Relation (34) shows that the specification of the $\mathcal{Q}_l^m$ perfectly defines the potential created by the particle system in regions of space outside the system $\mathscr{S}$. This potential $W(\boldsymbol{\rho})$ can be seen to be the sum of an infinite number of terms:

(*i*) The $l = 0$ term gives the contribution of the total charge of the system. This term is isotropic (it does not depend on Θ and Φ) and can be written:

$$W_0(\boldsymbol{\rho}) = \frac{1}{4\pi\varepsilon_0}\frac{1}{\rho}\sum_n q_n \tag{35}$$

This is the $1/\rho$ potential which would be created by the charges if they were all situated at O. It is zero if the system is globally neutral.

(*ii*) The $l = 1$ term gives the contribution of the electric dipole moment $\boldsymbol{\mathcal{D}}$ of the system. By performing transformations analogous to those in § 1-b-β, it can be shown that this contribution can be written:

$$W_1(\boldsymbol{\rho}) = \frac{1}{4\pi\varepsilon_0}\frac{\boldsymbol{\mathcal{D}}\cdot\boldsymbol{\rho}}{\rho^3} \tag{36}$$

This potential decreases like $1/\rho^2$ when ρ increases.

(*iii*) The $l = 2, 3, ...$ terms give, in the same way, the contributions to the potential $W(\boldsymbol{\rho})$ of the successive multipole moments of the system under study. When ρ increases, each of these contributions decreases like $1/\rho^{l+1}$, and its angular dependence is described by an lth-order spherical harmonic. Moreover, we see from (34) and definition (15) that the potential due to the multipole moment $\mathcal{Q}_l$ is at most of the order of magnitude of $W_0(\rho) \times (d/\rho)^l$, where d is the maximum distance of the various particles of the system $\mathscr{S}$ from the origin. Therefore, if we are concerned with the potential at a point $\boldsymbol{\rho}$ such that $\rho \gg d$ (the potential at a distant point), the $W_1(\boldsymbol{\rho})$ terms decrease very rapidly when l increases, and we do not make a large error by retaining only the lowest values of l in (34).

Comment:

If we wanted to calculate the magnetic field created by a system of moving charges, we could introduce the magnetic multipole moments of the system in an analogous way: the magnetic dipole moment[2], the magnetic quadrupole moment, etc. The parities of the magnetic moments are the opposite of those of the corresponding electric moments: the magnetic dipole moment is even, the magnetic quadrupole moment is odd, and so on. This property arises from the fact that the electric field is a polar vector while the magnetic field is an axial vector.

2. Matrix elements of electric multipole moments

We shall again consider, for the sake of simplicity, a system composed of a single spinless particle. However, generalization to N-particle systems presents no theoretical difficulty.

[2]There is no magnetic multipole moment of order $l = 0$ (magnetic monopole). This result is related to the fact that the magnetic field, whose divergence is zero according to Maxwell's equations, has a conservative flux.

The state space $\mathcal{E}_{\mathbf{r}}$ of the particle is spanned by an orthonormal basis, $\{|\chi_{n,l,m}\rangle\}$, of common eigenvectors of $\mathbf{L}^2$ [eigenvalue $l(l+1)\hbar^2$] and L_z (eigenvalue $m\hbar$). We shall evaluate the matrix elements of a multipole operator Q_l^m in such a basis.

2-a. General expression for the matrix elements

α. Expansion of the matrix elements

From the general results of Chapter VII, we know that the wave functions associated with the states $|\chi_{n,l,m}\rangle$ are necessarily of the form:

$$\chi_{n,l,m}(\mathbf{r}) = R_{n,l}(r)\, Y_l^m(\theta,\varphi) \tag{37}$$

The matrix element of the operator Q_l^m can therefore be written, using (12):

$$\begin{aligned}
&\langle \chi_{n_1,l_1,m_1} \mid Q_l^m \mid \chi_{n_2,l_2,m_2}\rangle = \\
&= \int_0^\infty r^2 \mathrm{d}r \int_0^\pi \sin\theta\, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\ \chi^*_{n_1,l_1,m_1}(r,\theta,\varphi)\ \mathcal{Q}_l^m(r,\theta,\varphi)\ \chi_{n_2,l_2,m_2}(r,\theta,\varphi) \\
&= q\sqrt{\frac{4\pi}{2l+1}} \int_0^\infty r^2 \mathrm{d}r\, R^*_{n_1,l_1}(r)\, R_{n_2,l_2}(r) r^l \int_0^\pi \sin\theta\, \mathrm{d}\theta \\
&\qquad \times \int_0^{2\pi} \mathrm{d}\varphi\ Y_{l_1}^{m_1 *}(\theta,\varphi)\ Y_l^m(\theta,\varphi) Y_{l_2}^{m_2}(\theta,\varphi)
\end{aligned} \tag{38}$$

Thus, in the matrix element under consideration, we have a radial integral and an angular integral. The latter, furthermore, can be simplified; using the complex conjugation relation for spherical harmonics [*cf.* Chap. VI, relation (D-29)] and relation (26) of Complement $\mathrm{C_X}$ (Wigner-Eckart theorem for spherical harmonics), we can show that it can be written:

$$\begin{aligned}
&(-1)^{m_1} \int_0^\pi \sin\theta\, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\ Y_{l_1}^{-m_1}(\theta,\varphi)\ Y_l^m(\theta,\varphi) Y_{l_2}^{m_2}(\theta,\varphi) = \\
&\qquad = \sqrt{\frac{(2l+1)(2l_2+1)}{4\pi(2l_1+1)}}\, \langle l_2, l\ ;\ 0,0 \mid l_1, 0\rangle\, \langle l_2, l\ ;\ m_2, m \mid l_1, m_1\rangle
\end{aligned} \tag{39}$$

Finally, we obtain:

$$\begin{aligned}
&\langle \chi_{n_1,l_1,m_1} \mid Q_l^m \mid \chi_{n_2,l_2,m_2}\rangle = \\
&\qquad = \frac{1}{\sqrt{2l_1+1}}\, \langle \chi_{n_1,l_1} \parallel Q_l \parallel \chi_{n_2,l_2}\rangle\, \langle l_2, l;\, m_2, m \mid l_1, m_1\rangle
\end{aligned} \tag{40}$$

where the "reduced matrix element" $\langle \chi_{n_1,l_1} \parallel Q_l \parallel \chi_{n_2,l_2}\rangle$ of the lth-order electric multipole operator is defined by:

$$\begin{aligned}
\langle \chi_{n_1,l_1} \parallel Q_l \parallel \chi_{n_2,l_2}\rangle &= q\sqrt{2l_2+1}\, \langle l_2, l\ ;\ 0,0 \mid l_1, 0\rangle \\
&\quad \times \int_0^\infty \mathrm{d}r\, r^{l+2}\ R^*_{n_1,l_1}(r)\, R_{n_2,l_2}(r)
\end{aligned} \tag{41}$$

Relation (40) expresses, in the particular case of electric multipole operators, a general theorem whose application in the case of vector operators has already been illustrated (*cf.* Complement $\mathrm{D_X}$): the Wigner-Eckart theorem.

Comment:

We have confined ourselves here to a system $\mathscr{S}$ composed of a single spinless particle. Nevertheless, if we consider a system of N particles which may have spins, we can generalize the results we have obtained. To do so, we must introduce the total angular momentum $\mathbf{J}$ of the system (the sum of the orbital and spin angular momenta of the N particles), and denote by $|\chi_{n,j,m}\rangle$ the eigenvectors common to $\mathbf{J}^2$ and J_z. We can then derive a relation similar to (40), in which l_1 and l_2 are replaced by j_1 and j_2 (*cf.* Complement G_X, exercise 8). However, the quantum numbers j_1, j_2, m_1 and m_2 can then be either integral or half-integral, depending on the physical system being considered.

β. The reduced matrix element

The reduced matrix element $\langle \chi_{n_1,l_1} \parallel Q_l \parallel \chi_{n_2,l_2} \rangle$ is independent of m, m_1 and m_2. It involves the radial part $R_{n,l}(r)$ of the wave functions $\chi_{n,l,m}(r,\theta,\varphi)$. Its value therefore depends on the $\{|\chi_{n,l,m}\rangle\}$ basis chosen, and general properties can hardly be attributed to it. However, it can be noted that the Clebsch-Gordan coefficient $\langle l_2, l\ ;\ 0,0 \mid l_1, 0\rangle$ involved in (41) is zero if $l_1 + l_2 + l$ is odd (*cf.* Complement B_X, § 3-c); this implies that the reduced matrix element has the same property.

Comment:

This property is related to the $(-1)^l$ parity of the electric multipole operators Q_l^m. For the magnetic multipole operators, we have already pointed out that their parity is $(-1)^{l+1}$; therefore it is when $l_1 + l_2 + l$ is even that their matrix elements are zero.

γ. The angular part of the matrix element

In (40), the Clebsch-Gordan coefficient $\langle l_2, l;\ m_2, m \mid l_1, m_1\rangle$ arises solely from the angular integral appearing in the matrix element of Q_l^m [*cf.* (38)]. This coefficient depends only on the quantum numbers associated with the angular momenta of the states being considered and does not involve the radial dependence $R_{n,l}(r)$ of the wave functions. This is why it appears in the matrix elements of multipole operators whenever one chooses a basis of eigenvectors common to $\mathbf{L}^2$ and L_z (or $\mathbf{J}^2$ and J_z for a system of N particles which may have spins; *cf.* comment of § α above). Now, we know that such bases are frequently used in quantum mechanics, and, in particular, that the stationary states of a particle in a central potential $W(r)$ can be chosen in this form. The radial functions $R_{n,l}(r)$ associated with the stationary states thus depend on the potential $W(r)$ chosen; this is therefore also true for the reduced matrix element $\langle \chi_{n_1,l_1} \parallel Q_l \parallel \chi_{n_2,l_2} \rangle$. On the other hand, this is not the case for the angular dependence of the wave functions, and the same Clebsch-Gordan coefficient appears for all $W(r)$; this is why it plays a universal role.

2-b. Selection rules

According to the properties of Clebsch-Gordan coefficients (*cf.* Complement B_X, § 1), $\langle l_2, l; m_2, m \mid l_1, m_1 \rangle$ can be different from zero only if we have both:

$$m_1 = m_2 + m \tag{42}$$

$$|l_1 - l_2| \leq l \leq l_1 + l_2 \tag{43}$$

Therefore, relation (40) implies that if at least one of these conditions is not met, the matrix element $\langle \chi_{n_1,l_1,m_1} \mid Q_l^m \mid \chi_{n_2,l_2,m_2} \rangle$ is necessarily zero. We thus obtain selection rules that enable us, without calculations, to simplify considerably our search for the matrix which represents any multipole operator Q_l^m.

Furthermore, we saw in § 2-a-β that the reduced matrix element of a multipole operator obeys another selection rule:

- for an electric multipole operator:

$$l_1 + l_2 + l = \text{an even number} \tag{44a}$$

- for a magnetic multipole operator:

$$l_1 + l_2 + l = \text{an odd number} \tag{44b}$$

2-c. Physical consequences

α. *The average value of a multipole operator in a state of well-defined angular momentum*

Assume that the state $|\psi\rangle$ of the particle is one of the basis states $|\chi_{n_1,l_1,m_1}\rangle$. The average value $\langle Q_l^m \rangle$ of the operator Q_l^m is then:

$$\langle Q_l^m \rangle = \langle \chi_{n_1,l_1,m_1} \mid Q_l^m \mid \chi_{n_1,l_1,m_1} \rangle \tag{45}$$

Conditions (42) and (43) are written here:

$$m = 0 \tag{46}$$

$$0 \leq l \leq 2l_1 \tag{47}$$

Thus we obtain the following important rules:

- the average values, in a state $|\chi_{n_1,l_1,m_1}\rangle$, of all the operators Q_l^m are zero if $m \neq 0$:

$$\langle Q_l^m \rangle = 0 \quad \text{if} \quad m \neq 0 \tag{48}$$

- *the average values, in a state* $|\chi_{n_1,l_1,m_1}\rangle$, *of all operators of order* l *higher than* $2l_1$ *are zero*:

$$\langle Q_l^m \rangle = 0 \quad \text{if} \quad l > 2l_1 \tag{49}$$

If we now assume that the state $|\psi\rangle$, instead of being a state $|\chi_{n_1,l_1,m_1}\rangle$, is any superposition of such states, all corresponding to the same value of l_1, it is not difficult to show that rule (49) remains valid [but not rule (48), since, in general, matrix elements for which $m_1 \neq m_2$ then contribute to the average value $\langle Q_l^m \rangle$]. Relation (49) is therefore a very general one and can be applied whenever the system is in an eigenstate of $\mathbf{L}^2$.

Furthermore, relations (44) imply that the average value of an lth-order multipole operator can be different from zero only if:

- for an electric multipole operator:

$$l = \text{an even number} \tag{50a}$$

- for a magnetic multipole operator:

$$l = \text{an odd number} \tag{50b}$$

The preceding rules enable us to obtain, conveniently and without calculations, some simple physical results. For example, in an $l = 0$ state (like the ground state of the hydrogen atom), the dipole moments (electric or magnetic), quadrupole moments (electric or magnetic), etc. are always zero. For an $l = 1$ state, only the 0th-, 1st- and 2nd-order multipole operators can be non-zero; parity rules (50) indicate that they are the total charge and electric quadrupole of the system, as well as its magnetic dipole.

Comment:

The predictions obtained can be generalized to more complex systems (many-electron atoms for example). If the angular momentum of such a system is j (integral or half-integral) one can show that it suffices to replace, in (49), l_1 by j.

We shall apply, for example, rules (49) and (50) to the study of the electromagnetic properties of an atomic nucleus. We know that such a nucleus is a bound system composed of protons and neutrons, interacting through nuclear forces. If, in the ground state[3], the eigenvalue of the square of the angular momentum is $I(I+1)\hbar^2$, the quantum number I is called the nuclear spin.

The rules we have stated indicate that:

- if $I = 0$, the electromagnetic interactions of the nucleus are characterized by its total charge, all the other multipole moments being zero. This is the case, for example, for ^{4}He nuclei ("α-particles"), ^{20}Ne nuclei, etc.
- if $I = 1/2$, the nucleus has an electric charge and a magnetic dipole moment [parity rule (50a) excludes an electric dipole moment]. This is the case for the ^{3}He nucleus and the ^{1}H nucleus (i.e., the proton), as well as for all spin 1/2 particles (electrons, muons, neutrons, etc.).
- if $I = 1$, we must add the electric quadrupole moment to the charge and the magnetic dipole moment. This is the case for ^{2}H (deuterium), ^{6}Li, etc.

This argument can be generalized to any value of I. Actually, very few nuclei have spins greater than 3 or 4.

[3] In atomic physics, one generally consider the nucleus to be in its ground state: the energies involved, although high enough to excite the electronic cloud of the atom, are much too small to excite the nucleus.

β. Matrix elements between states of different quantum numbers

For arbitrary l_1, l_2, m_1 and m_2, the selection rules must be applied in their general forms, (42), (43) and (44). Consider, for example, a particle of charge q subjected to a central potential $V_0(r)$, whose stationary states are the states $|\chi_{n,l,m}\rangle$. Assume that we then add an additional electric field $\boldsymbol{\mathcal{E}}$, uniform and parallel to Oz. In the corresponding coupling Hamiltonian, the only non-zero term is the electric dipole term (*cf.* § 1-b-β):

$$\begin{aligned} V(\mathbf{R}) &= -\mathbf{D}\cdot\boldsymbol{\mathcal{E}} \\ &= -D_z\mathcal{E} \end{aligned} \tag{51}$$

As we saw in (22), the operator D_z is equal to the operator Q_1^0. Selection rules (42) and (43) then indicate that:

- the states $|\chi_{n,l,m}\rangle$ coupled by the additional Hamiltonian $V(\mathbf{R})$ necessarily correspond to the same value of m.
- the l-values of the two states necessarily differ by ± 1 [they cannot be equal, according to (44a)]. We can predict without calculation that a large number of matrix elements of $V(\mathbf{R})$ are zero. This considerably simplifies, for example, the study of the Stark effect (*cf.* Complement $\mathrm{E_{XII}}$), and that of the selection rules governing the emission spectrum of atoms (*cf.* Complement $\mathrm{A_{XIII}}$).

References and suggestions for further reading:

Cagnac and Pebay-Peyroula (11.2), annexe IV; Valentin (16 1), Chap. VIII; Jackson (7.5). Chaps. 4 and 16.

Complement $\mathsf{F_X}$

Evolution of two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ coupled by an interaction $a\mathbf{J}_1 \cdot \mathbf{J}_2$

1 Classical review 1092
1-a Equations of motion 1092
1-b Motion of $\mathcal{J}_1$ and $\mathcal{J}_2$ 1093
2 Quantum mechanical evolution of the average values $\langle \mathbf{J}_1 \rangle$ and $\langle \mathbf{J}_2 \rangle$ 1093
2-a Calculation of $\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_1 \rangle$ and $\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_2 \rangle$ 1093
2-b Discussion 1095
3 The special case of two spin 1/2's 1095
3-a Stationary states of the two-spin system 1095
3-b Calculation of $\langle \mathbf{S}_1 \rangle(t)$ 1096
3-c Discussion. Polarization of the magnetic dipole transitions . . 1098
4 Study of a simple model for the collision of two spin 1/2 particles 1102
4-a Description of the model 1102
4-b State of the system after collision 1103
4-c Discussion. Correlation introduced by the collision 1104

In a physical system, we must often consider the effect of a coupling between two partial angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$. These can, for example, be the angular momenta of two electrons in an atom, or the orbital and spin angular momenta of an electron. In the presence of such a coupling, $\mathbf{J}_1$ and $\mathbf{J}_2$ are no longer constants of the motion; only their sum:

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \tag{1}$$

commutes with the total Hamiltonian of the system.

We shall assume that the term of the Hamiltonian which introduces a coupling between $\mathbf{J}_1$ and $\mathbf{J}_2$ has the simple form:

$$W = a\,\mathbf{J}_1 \cdot \mathbf{J}_2 \tag{2}$$

where a is a real constant. Such a situation is frequently encountered in atomic physics. We shall see numerous examples in Chapter XII, when we use perturbation theory to study the effect of interactions involving electron or proton spins on the hydrogen atom spectrum. When the coupling has the form (2), classical theory predicts that the classical angular momenta $\mathcal{J}_1$ and $\mathcal{J}_2$ will precess about their resultant $\mathcal{J}$ with an angular velocity proportional to the constant a (*cf.* § 1 below). The "vector model" of the atom, which played a very important role in the history of the development of atomic physics, is founded on this result. In this complement, we shall show how, with the knowledge

of the common eigenstates of $\mathbf{J}^2$ and J_z, one can study the motion of the average values $\langle \mathbf{J}_1 \rangle$ and $\langle \mathbf{J}_2 \rangle$, and again derive, at least partially, the results of the vector model of the atom (§§ 2 and 3). In addition, this study will enable us to specify in simple cases the polarization of the electro-magnetic waves emitted or absorbed in magnetic dipole transitions. Finally, (§ 4), we shall take up the case in which the two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ are coupled only during a collision but not permanently. This case will serve as a simple illustration of the important concept of correlation between two systems.

1. Classical review

1-a. Equations of motion

If θ is the angle between the classical angular momenta $\boldsymbol{\mathcal{J}}_1$ and $\boldsymbol{\mathcal{J}}_2$ (Fig. 1), the coupling energy can be written:

$$\mathcal{W} = a\boldsymbol{\mathcal{J}}_1 \cdot \boldsymbol{\mathcal{J}}_2 = a\mathcal{J}_1\mathcal{J}_2 \cos\theta \tag{3}$$

Let $\mathcal{H}_0$ be the energy of the total system in the absence of coupling [$\mathcal{H}_0$ can represent, for example, the sum of the rotational kinetic energies of systems (1) and (2)]. We shall assume:

$$\mathcal{W} \ll \mathcal{H}_0 \tag{4}$$

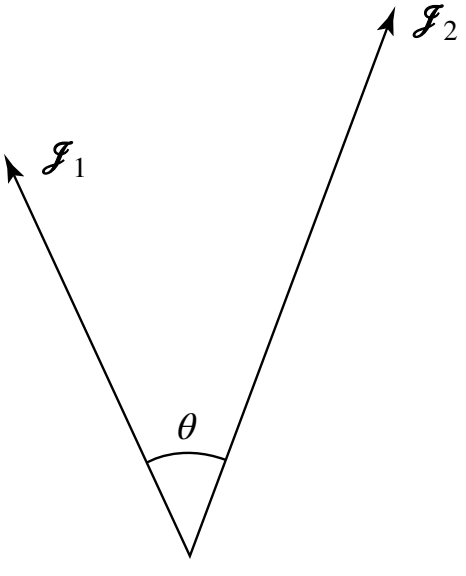


Figure 1: Two classical angular momenta $\boldsymbol{\mathcal{J}}_1$ and $\boldsymbol{\mathcal{J}}_2$ are coupled by an interaction term $\mathcal{W} = a\boldsymbol{\mathcal{J}}_1 \cdot \boldsymbol{\mathcal{J}}_2 = a\mathcal{J}_1\mathcal{J}_2 \cos\theta$.

Let us calculate the moment $\boldsymbol{\mathcal{M}}_1$ of the forces acting on system (1). Let $\mathbf{u}$ be a unit vector, and $\mathrm{d}\mathcal{W}$ the variation of the coupling energy when the system (1) is rotated through an angle $\mathrm{d}\alpha$ about $\mathbf{u}$. We know (the theorem of virtual work) that:

$$\boldsymbol{\mathcal{M}}_1 \cdot \mathbf{u} = -\frac{\mathrm{d}\mathcal{W}}{\mathrm{d}\alpha} \tag{5}$$

Starting with (3) and 5), we obtain, via a simple calculation:

$$\boldsymbol{\mathcal{M}}_1 = -a\boldsymbol{\mathcal{J}}_1 \times \boldsymbol{\mathcal{J}}_2 \tag{6a}$$

$$\boldsymbol{\mathcal{M}}_2 = -a\boldsymbol{\mathcal{J}}_2 \times \boldsymbol{\mathcal{J}}_1 \tag{6b}$$

and, consequently:

$$\frac{\mathrm{d}\boldsymbol{\mathcal{J}}_1}{\mathrm{d}t} = -a\boldsymbol{\mathcal{J}}_1 \times \boldsymbol{\mathcal{J}}_2 \tag{7a}$$

$$\frac{\mathrm{d}\boldsymbol{\mathcal{J}}_2}{\mathrm{d}t} = -a\boldsymbol{\mathcal{J}}_2 \times \boldsymbol{\mathcal{J}}_1 \tag{7b}$$

1-b. Motion of $\boldsymbol{\mathcal{J}}_1$ and $\boldsymbol{\mathcal{J}}_2$

Adding (7a) and (7b), we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\mathcal{J}}_1 + \boldsymbol{\mathcal{J}}_2) = \mathbf{0} \tag{8}$$

which shows that the total angular momentum $\boldsymbol{\mathcal{J}}_1 + \boldsymbol{\mathcal{J}}_2$ is indeed a constant of the motion. Furthermore, it can easily be deduced from (7a) and (7b) that:

$$\boldsymbol{\mathcal{J}}_1 \cdot \left(\frac{\mathrm{d}\boldsymbol{\mathcal{J}}_1}{\mathrm{d}t}\right) = \boldsymbol{\mathcal{J}}_2 \cdot \left(\frac{\mathrm{d}\boldsymbol{\mathcal{J}}_2}{\mathrm{d}t}\right) = 0 \tag{9}$$

and:

$$\boldsymbol{\mathcal{J}}_1 \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{J}}_2\right) + \left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{J}}_1\right) \cdot \boldsymbol{\mathcal{J}}_2 = \frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\mathcal{J}}_1 \cdot \boldsymbol{\mathcal{J}}_2\right) = 0 \tag{10}$$

The angle θ between $\boldsymbol{\mathcal{J}}_1$ and $\boldsymbol{\mathcal{J}}_2$, as well as the moduli of $\boldsymbol{\mathcal{J}}_1$ and $\boldsymbol{\mathcal{J}}_2$, therefore remain constant over time. Finally:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{J}}_1 = a\boldsymbol{\mathcal{J}}_2 \times \boldsymbol{\mathcal{J}}_1 = a(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{J}}_1) \times \boldsymbol{\mathcal{J}}_1 = a\boldsymbol{\mathcal{J}} \times \boldsymbol{\mathcal{J}}_1 \tag{11}$$

Since $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{J}}_1 + \boldsymbol{\mathcal{J}}_2$ is constant, the preceding equation shows that $\boldsymbol{\mathcal{J}}_1$ precesses about $\boldsymbol{\mathcal{J}}$ with an angular velocity equal to $a|\boldsymbol{\mathcal{J}}|$ (Fig. 2).

Under the effect of the coupling, $\boldsymbol{\mathcal{J}}_1$ and $\boldsymbol{\mathcal{J}}_2$ therefore precess about their resultant $\boldsymbol{\mathcal{J}}$ with an angular velocity proportional to $|\boldsymbol{\mathcal{J}}|$ and to the coupling constant a.

2. Quantum mechanical evolution of the average values $\langle \mathbf{J}_1 \rangle$ and $\langle \mathbf{J}_2 \rangle$

2-a. Calculation of $\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_1 \rangle$ and $\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_2 \rangle$

Recall, first of all, that if A is an observable of a quantum mechanical system of Hamiltonian H, we have (*cf.* Chap. III, § D-1-d):

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A \rangle(t) = \frac{1}{i\hbar}\langle [A, H] \rangle(t) \tag{12}$$

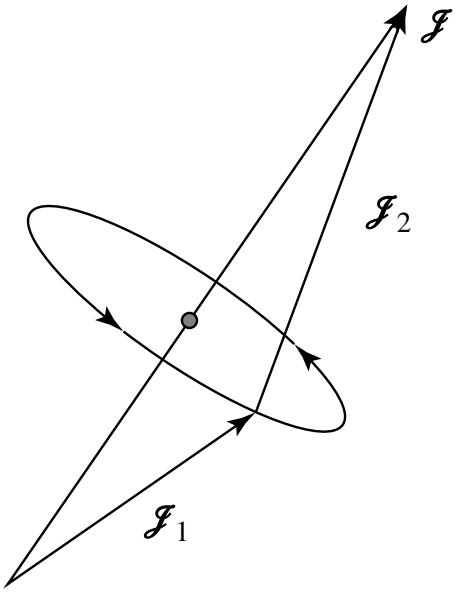


Figure 2: Under the effect of the coupling $\mathcal{W} = a\boldsymbol{\mathcal{J}}_1 \cdot \boldsymbol{\mathcal{J}}_2$, the angular momenta $\boldsymbol{\mathcal{J}}_1$ and $\boldsymbol{\mathcal{J}}_2$ precess about their resultant $\boldsymbol{\mathcal{J}}$, which is a constant of the motion.

In the present case, the Hamiltonian H is equal to:

$$H = H_0 + W \tag{13}$$

where H_0 is the sum of the energies of systems (1) and (2), and W is the coupling between $\mathbf{J}_1$ and $\mathbf{J}_2$ given in (2). In the absence of such a coupling, $\mathbf{J}_1$ and $\mathbf{J}_2$ are constants of the motion (they commute with H_0). Therefore, in the presence of the coupling, we have simply:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_1 \rangle = \frac{1}{i\hbar}\langle [\mathbf{J}_1, W] \rangle = \frac{a}{i\hbar}\langle [\mathbf{J}_1, \mathbf{J}_1 \cdot \mathbf{J}_2] \rangle \tag{14}$$

and an analogous expression for $\dfrac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_2 \rangle$. The calculation of the commutator appearing in formula 14) does not present any difficulty. We have, for example:

$$\begin{aligned} [J_{1x}, \mathbf{J}_1 \cdot \mathbf{J}_2] &= [J_{1x}, J_{1y}J_{2y}] + [J_{1x}, J_{1z}J_{2z}] \\ &= i\hbar J_{1z}J_{2y} - i\hbar J_{1y}J_{2z} \\ &= -i\hbar(\mathbf{J}_1 \times \mathbf{J}_2)_x \end{aligned} \tag{15}$$

From this, we see finally that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_1 \rangle = -a\langle \mathbf{J}_1 \times \mathbf{J}_2 \rangle \tag{16a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_2 \rangle = -a\langle \mathbf{J}_2 \times \mathbf{J}_1 \rangle \tag{16b}$$

2-b. Discussion

Note the close analogy between formulas (7a) and (7b) on the one hand and formulas (16a) and (16b) on the other. Adding (16a) and (16b), we again find that $\mathbf{J}$ is a constant of the motion, since:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_1 \rangle + \frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J}_2 \rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{J} \rangle = \mathbf{0} \tag{17}$$

However, we must recall that, in general:

$$\langle \mathbf{J}_1 \times \mathbf{J}_2 \rangle \neq \langle \mathbf{J}_1 \rangle \times \langle \mathbf{J}_2 \rangle \tag{18}$$

The motion of the average values is therefore not necessarily identical to the classical motion. To examine this point in greater detail, we shall now study a special case: that in which $\mathbf{J}_1$ and $\mathbf{J}_2$ are two spin 1/2's, which we shall denote by $\mathbf{S}_1$ and $\mathbf{S}_2$.

3. The special case of two spin 1/2's

The evolution of a quantum mechanical system can easily be calculated in the basis of eigenstates of the Hamiltonian of this system. Therefore, we shall begin by determining the stationary states of the two-spin system.

3-a. Stationary states of the two-spin system

Let:

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \tag{19}$$

be the total spin. Squaring both sides of (19), we obtain:

$$\mathbf{S}^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 \tag{20}$$

which enables us to write W in the form:

$$W = a\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{a}{2}\left[\mathbf{S}^2 - \mathbf{S}_1^2 - \mathbf{S}_2^2\right] = \frac{a}{2}\left[\mathbf{S}^2 - \frac{3}{2}\hbar^2\right] \tag{21}$$

(all vectors of the state space are eigenvectors of $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$ with the eigenvalue $3\hbar^2/4$).

In the absence of coupling, the Hamiltonian H_0 of the system is diagonal in the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis (with $\varepsilon_1 = \pm$, $\varepsilon_2 = \pm$) of eigenstates of S_{1z} and S_{2z}, as well as in the $\{|S, M\rangle\}$ basis (with $S = 0$ or 1, $-S \leq M \leq +S$) of eigenstates of $\mathbf{S}^2$ and S_z. The various vectors $|\varepsilon_1, \varepsilon_2\rangle$ or $|S, M\rangle$ are eigenvectors of H_0 with the same eigenvalue, which we shall take to be the energy origin.

When we take the coupling W into account, we see from formula (21) that the total Hamiltonian $H = H_0 + W$ is no longer diagonal in the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis. However, we may write:

$$(H_0 + W)|S, M\rangle = \frac{a\hbar^2}{2}\left[S(S+1) - \frac{3}{2}\right]|S, M\rangle \tag{22}$$

The stationary states of the two-spin system therefore separate into two levels (Fig. 3): the $S = 1$ level, three-fold degenerate with energy $E_1 = a\hbar^2/4$, and the non-degenerate $S = 0$ level, with energy $E_0 = -3a\hbar^2/4$. The splitting between the two levels is equal to $a\hbar^2$. If we set:

$$a\hbar^2 = \hbar\Omega \tag{23}$$

$\Omega/2\pi$ is the only non-zero Bohr frequency of the two-spin system.

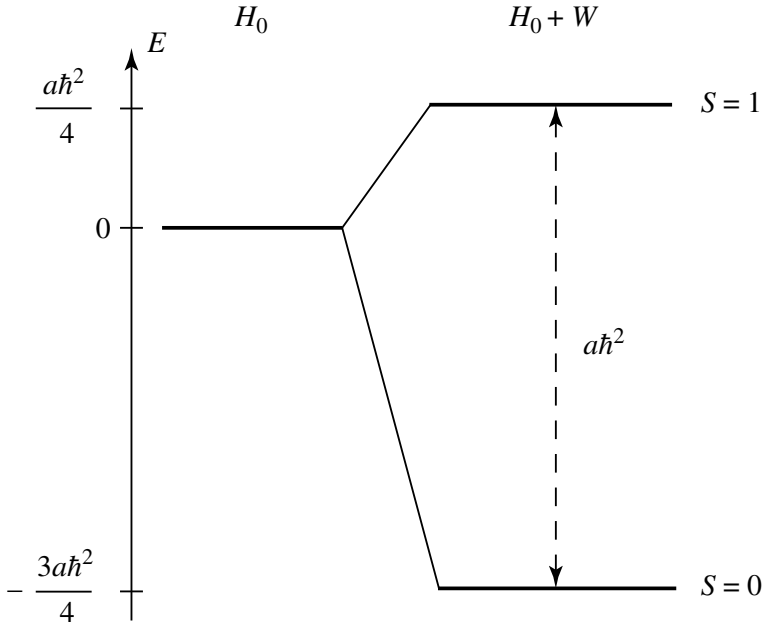


Figure 3: Energy levels of a system of two spin 1/2's. On the left-hand side of the figure, the coupling is assumed to be zero, and we obtain a single level which is four-fold degenerate. The coupling $W = a\mathbf{S}_1 \cdot \mathbf{S}_2$ splits it into two distinct levels, separated by an energy of $a\hbar^2$: the triplet level ($S = 1$, three-fold degenerate) and the singlet level ($S = 0$, non-degenerate).

3-b. Calculation of $\langle \mathbf{S}_1 \rangle(t)$

In order to find the evolution of $\langle \mathbf{S}_1 \rangle(t)$, we must first calculate the matrices representing S_{1x}, S_{1y} and S_{1z} (or, more simply, S_{1z} and $S_{1+} = S_{1x} + iS_{1y}$) in the $\{|S, M\rangle\}$ basis of stationary states. If we use expressions (B-22) and (B-23) of Chapter X, which give the expansion of the states $|S, M\rangle$ on the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis, it is possible to easily

calculate the action of S_{1z} or S_{1+} on the kets $|S, M\rangle$. We find:

$$\begin{cases} S_{1z}|1,1\rangle = \dfrac{\hbar}{2}|1,1\rangle \\ S_{1z}|1,0\rangle = \dfrac{\hbar}{2}|0,0\rangle \\ S_{1z}|1,-1\rangle = -\dfrac{\hbar}{2}|1,-1\rangle \\ S_{1z}|0,0\rangle = \dfrac{\hbar}{2}|1,0\rangle \end{cases} \tag{24}$$

and:

$$\begin{cases} S_{1+}|1,1\rangle = 0 \\ S_{1+}|1,0\rangle = \dfrac{\hbar}{\sqrt{2}}|1,1\rangle \\ S_{1+}|1,-1\rangle = \dfrac{\hbar}{\sqrt{2}}(|1,0\rangle + |0,0\rangle) \\ S_{1+}|0,0\rangle = -\dfrac{\hbar}{\sqrt{2}}|1,1\rangle \end{cases} \tag{25}$$

From this, we can immediately derive the matrices representing S_{1z} and S_{1+} in the basis of the four states $|S, M\rangle$ arranged in the order $|1,1\rangle$, $|1,0\rangle$, $|1,-1\rangle$ and $|0,0\rangle$:

$$(S_{1z}) = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{26}$$

$$(S_{1+}) = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{27}$$

Comment:

It can easily be shown that the restrictions of the S_{1z} and S_{1+} matrices to the $S = 1$ subspace are proportional respectively (with the same proportionality coefficient) to the matrices representing S_z and S_+ in the same subspace. This result could have been expected, in view of the Wigner-Eckart theorem relative to vector operators (*cf.* Complement $\mathrm{D_X}$).

Let:

$$|\psi(0)\rangle = \alpha|0,0\rangle + \beta_{-1}|1,-1\rangle + \beta_0|1,0\rangle + \beta_1|1,1\rangle \tag{28}$$

be the state of the system at the instant $t = 0$. From this we deduce the expression for $|\psi(t)\rangle$ (to within the factor $\mathrm{e}^{3ia\hbar t/4}$):

$$|\psi(t)\rangle = \alpha|0,0\rangle + [\beta_{-1}|1,-1\rangle + \beta_0|1,0\rangle + \beta_1|1,1\rangle]\,\mathrm{e}^{-i\Omega t} \tag{29}$$

It is then easy to obtain, using (26) and (27):

$$\langle S_{1z}\rangle(t) = \langle\psi(t) \mid S_{1z} \mid \psi(t)\rangle$$
$$= \frac{\hbar}{2}\left[|\beta_1|^2 - |\beta_{-1}|^2 + e^{i\Omega t}\,\alpha\beta_0^* + e^{-i\Omega t}\,\alpha^*\beta_0\right] \tag{30}$$

$$\langle S_{1+}\rangle(t) = \langle\psi(t) \mid S_{1+} \mid \psi(t)\rangle$$
$$= \frac{\hbar}{\sqrt{2}}\left[\beta_1^*\beta_0 + \beta_0^*\beta_{-1} - e^{i\Omega t}\,\beta_1^*\alpha + e^{-i\Omega t}\,\alpha^*\beta_{-1}\right] \tag{31}$$

The average values $\langle S_{1x}\rangle(t)$ and $\langle S_{1y}\rangle(t)$ can be expressed in terms of $\langle S_{1+}\rangle(t)$:

$$\langle S_{1x}\rangle(t) = \mathrm{Re}\,\langle S_{1+}\rangle(t) \tag{32}$$

$$\langle S_{1y}\rangle(t) = \mathrm{Im}\,\langle S_{1+}\rangle(t) \tag{33}$$

Analogous calculations enable us to obtain the three components of $\langle\mathbf{S}_2\rangle(t)$.

3-c. Discussion. Polarization of the magnetic dipole transitions

Studying the motion of $\langle\mathbf{S}_1\rangle(t)$ does more than compare the vector model of the atom with the predictions of quantum mechanics. It also enables us to specify the polarization of the electromagnetic waves emitted due to the motion of $\langle\mathbf{S}_1\rangle(t)$.

The Bohr frequency $\Omega/2\pi$ appears in the evolution of $\langle\mathbf{S}_1\rangle(t)$ because of the existence of non-zero matrix elements of S_{1x}, S_{1y}, or S_{1z} between the state $|0,0\rangle$ and one of the states $|1,M\rangle$ (with $M = -1,\ 0,\ +1$). In (28) or (29), we shall begin by assuming that, with α non-zero, only one of the three coefficients β_{-1}, β_0 or β_1 is different from zero. The examination of the motion of $\langle\mathbf{S}_1\rangle(t)$ in the three corresponding cases thus will enable us to specify the polarization of the radiation associated with the three magnetic dipole transitions:

$$|0,0\rangle \leftrightarrow |1,0\rangle, \quad |0,0\rangle \leftrightarrow |1,1\rangle \quad \text{and} \quad |0,0\rangle \leftrightarrow |1,-1\rangle$$

We can always choose α to be real; we shall set:

$$\beta_M = |\beta_M| e^{i\varphi_M} \quad (M = -1, 0, 1) \tag{34}$$

Comment:

Actually, the electromagnetic waves are radiated by the magnetic moments $\mathbf{M}_1$ and $\mathbf{M}_2$ associated with $\mathbf{S}_1$ and $\mathbf{S}_2$ (hence the name, magnetic dipole transitions). $\mathbf{M}_1$ and $\mathbf{M}_2$ are proportional respectively to $\mathbf{S}_1$ and $\mathbf{S}_2$. To be completely rigorous, we should then study the evolution of $\langle\mathbf{M}_1 + \mathbf{M}_2\rangle(t)$. Here we shall assume $\langle\mathbf{M}_1\rangle \gg \langle\mathbf{M}_2\rangle$. Such a situation is found, for example, in the ground state of the hydrogen atom: the hyperfine structure of this state is due to the coupling between the spin of the electron and that of the proton (*cf.* Chap. XII, § D). But the magnetic moment of the electron spin is much larger than that of the proton, so that the emission and absorption of electromagnetic waves at the hyperfine transition frequency are essentially governed by the motion of the electron spin. Taking $\langle\mathbf{M}_2\rangle$ into account as well would complicate the calculations without modifying the conclusions.

α. *The* $|0,0\rangle \leftrightarrow |1,0\rangle$ *transition* ($\beta_1 = \beta_{-1} = 0$)

If we take $\beta_1 = \beta_{-1} = 0$ in (30), (31), (32) and (33), we get:

$$\langle S_{1x}\rangle(t) = \langle S_{1y}\rangle(t) = 0$$

$$\langle S_{1z}\rangle(t) = \hbar\alpha\,|\beta_0|\cos(\Omega t - \varphi_0) \tag{35}$$

Furthermore, it can easily be seen that:

$$\langle S_x\rangle(t) = \langle S_y\rangle(t) = \langle S_z\rangle(t) = 0 \tag{36}$$

$\langle \mathbf{S}_1\rangle(t)$ and $\langle \mathbf{S}_2\rangle(t)$ are then permanently of opposite direction and vibrate along Oz at the frequency $\Omega/2\pi$ (Fig. 4).

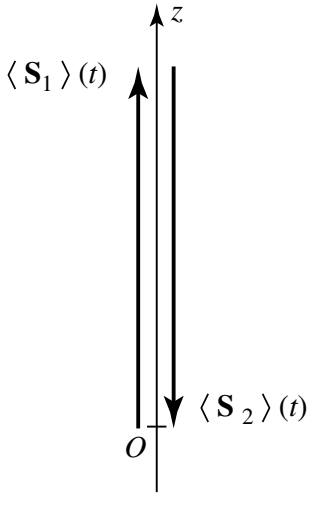


Figure 4: If the state of the two-spin system is a superposition of only the two stationary states $|0,0\rangle$ *and* $|1,0\rangle$, $\langle \mathbf{S}_1\rangle$ *and* $\langle \mathbf{S}_2\rangle$ *are always of opposite direction and vibrate along* Oz *at the frequency* $\Omega/2\pi$.

The electromagnetic waves emitted by $\langle \mathbf{S}_1\rangle$ therefore have a magnetic field[1] linearly polarized along Oz ("π polarization").

We see in this example that $(\langle \mathbf{S}_1\rangle)^2$ varies over time and is therefore not equal to $\langle \mathbf{S}_1^2\rangle$ (which is constant and equal to $3\hbar^2/4$). This represents an important difference with the classical situation studied in § 1, in which $\mathcal{J}_1$ maintains a constant length over time.

[1]Since these are magnetic dipole transitions, we are concerned with the magnetic field vector of the radiated wave. In the case of an electric dipole transition (*cf.* Complement $\mathrm{D_{VII}}$, § 2-c), on the other hand, we would be concerned with the radiated electric field.

β. The $|0,0\rangle \leftrightarrow |1,1\rangle$ transition ($\beta_0 = \beta_{-1} = 0$)

We find in this case:

$$\begin{cases} \langle S_{1z}\rangle(t) = \dfrac{\hbar}{2}|\beta_1|^2 \\ \langle S_{1x}\rangle(t) = -\dfrac{\hbar}{\sqrt{2}}\alpha|\beta_1|\cos(\Omega t - \varphi_1) \\ \langle S_{1y}\rangle(t) = -\dfrac{\hbar}{\sqrt{2}}\alpha|\beta_1|\sin(\Omega t - \varphi_1) \end{cases} \tag{37}$$

Furthermore, it can easily be verified that:

$$\begin{cases} \langle S_z\rangle(t) = \hbar|\beta_1|^2 \\ \langle S_x\rangle(t) = \langle S_y\rangle(t) = 0 \end{cases} \tag{38}$$

From this, it can be seen (Fig. 5) that $\langle \mathbf{S}_1\rangle(t)$ and $\langle \mathbf{S}_2\rangle(t)$ precess counterclockwise at an angular velocity Ω about their resultant $\langle \mathbf{S}\rangle$, which is parallel to Oz. The electromagnetic waves emitted by $\langle \mathbf{S}_1\rangle(t)$ in this case therefore have a right-hand circular polarization ("σ_+ polarization").

Note that here the motion obtained for the average values $\langle \mathbf{S}_1\rangle$ and $\langle \mathbf{S}_2\rangle$ is the classical motion.

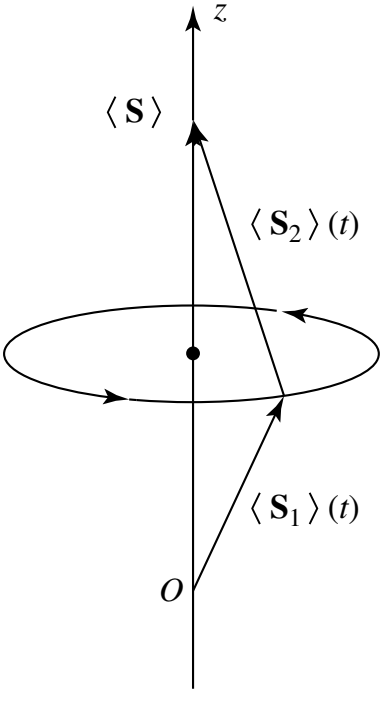


Figure 5: If the state of the two-spin system is a superposition of only the stationary states $|0,0\rangle$ and $|1,1\rangle$, $\langle \mathbf{S}_1\rangle$ and $\langle \mathbf{S}_2\rangle$ precess counterclockwise about their resultant $\langle \mathbf{S}\rangle$, with the angular velocity Ω.

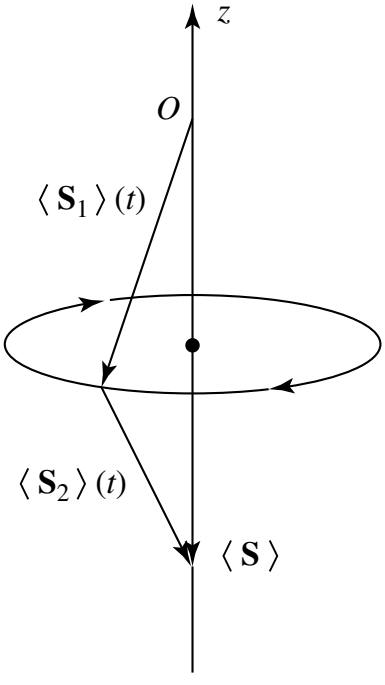


Figure 6: If the state of the two-spin system is a superposition of only the stationary states $|0,0\rangle$ and $|1,-1\rangle$, $\langle \mathbf{S}_1 \rangle$ and $\langle \mathbf{S}_2 \rangle$ still precess in the counterclockwise direction with the angular velocity Ω about their resultant $\langle \mathbf{S} \rangle$; however, the latter is now directed opposite to Oz.

γ. *The $|0,0\rangle \leftrightarrow |1,-1\rangle$ transition ($\beta_0 = \beta_1 = 0$)*

The calculations are closely analogous to those of the preceding section and lead to the following result (Fig. 6): $\langle \mathbf{S}_1 \rangle(t)$ and $\langle \mathbf{S}_2 \rangle(t)$ precess about Oz, again at the angular velocity Ω, but in the clockwise direction. It must be noted that $\langle S_z \rangle = -\hbar|\beta_{-1}|^2$ is now negative, so that while the direction of the precession of $\langle \mathbf{S}_1 \rangle$ and $\langle \mathbf{S}_2 \rangle$ about Oz is different from what it was in the preceding case, it remains the same relative to $\langle \mathbf{S} \rangle$. The electromagnetic waves emitted by $\langle \mathbf{S}_1 \rangle$ are now left-hand circularly polarized ("σ_- polarization").

δ. *General case*

In the general case (any α, β_{-1}, β_0 and β_1), we see from (30), (31), (32) and (33) that the components of $\langle \mathbf{S}_1 \rangle(t)$ on the three axes contain a static part and a part modulated at the frequency $\Omega/2\pi$. Since these three projected motions are sinusoidal motions of the same frequency, the tip of $\langle \mathbf{S}_1 \rangle(t)$ describes an ellipse in space. As the sum

$$\langle \mathbf{S}_1 \rangle(t) + \langle \mathbf{S}_2 \rangle(t) = \langle \mathbf{S} \rangle$$

remains constant, the tip of $\langle \mathbf{S}_2 \rangle(t)$ also describes an ellipse (Fig. 7).

Thus we find for the general case only part of the results of the vector model of the atom. We do find that, the larger the coupling constant a, the more rapidly $\langle \mathbf{S}_1 \rangle(t)$ and $\langle \mathbf{S}_2 \rangle(t)$ precess about $\langle \mathbf{S} \rangle$. However, as we saw clearly in the special case α studied

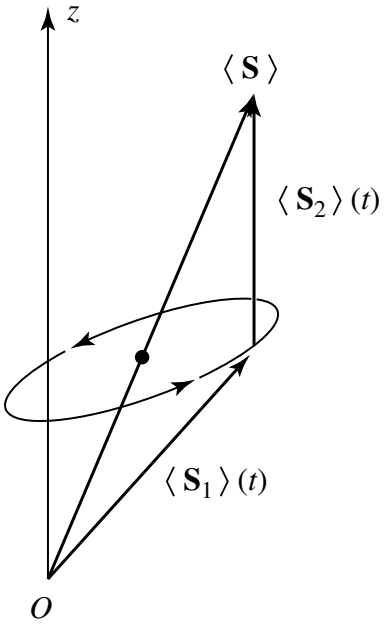


Figure 7: Motion of $\langle \mathbf{S}_1 \rangle(t)$ and $\langle \mathbf{S}_2 \rangle(t)$ in the general case, in which the state of the two-spin system is a superposition of the four stationary states $|0,0\rangle$, $|1,1\rangle$, $|1,0\rangle$ and $|1,-1\rangle$. The resultant $\langle \mathbf{S} \rangle$ is still constant but is not necessarily directed along Oz. $\langle \mathbf{S}_1 \rangle$ and $\langle \mathbf{S}_2 \rangle$ no longer have constant lengths, and their tips describe ellipses.

above, $|\langle \mathbf{S}_1 \rangle(t)|$ is not constant, and the tip of $\langle \mathbf{S}_1 \rangle(t)$ does not describe a circle in the general case.

4. Study of a simple model for the collision of two spin 1/2 particles

4-a. Description of the model

Consider two spin 1/2 particles, whose external degrees of freedom we shall treat classically and whose spin degrees of freedom we shall treat quantum mechanically. We shall assume that their trajectories are rectilinear (Fig. 8) and that the interaction between the two spins $\mathbf{S}_1$ and $\mathbf{S}_2$ is of the form $W = a\mathbf{S}_1 \cdot \mathbf{S}_2$, where the coupling constant a is a rapidly decreasing function of the distance r separating the two particles.

Since r varies over time, so does a. The shape of the variation of a with respect to t

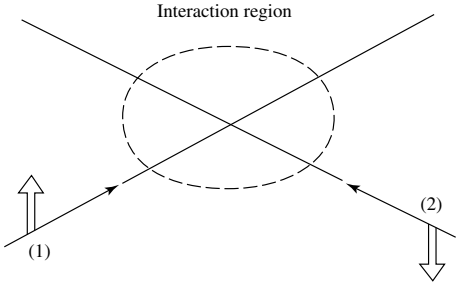


Figure 8: Collision between two spin 1/2 particles (1) and (2) whose orbital variables can be treated classically. The spin state of each particle is represented by a large arrow.

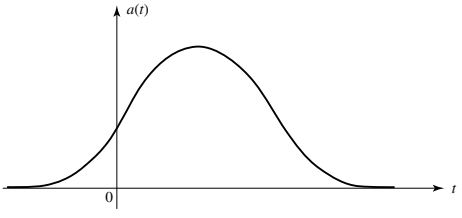


Figure 9: Shape of the variation of the coupling constant $a(t)$ during the collision.

is shown in Figure 9. The maximum corresponds to the time when the distance between the two particles is at a minimum. To simplify the reasoning, we shall replace the curve in Figure 9 by the one in Figure 10.

The problem we have here is the following: before the collision, that is, at $t = -\infty$, the spin state of the two-particle system is:

$$|\psi(-\infty)\rangle = |+,-\rangle \tag{39}$$

What is the state $|\psi(+\infty)\rangle$ of the system after the collision?

4-b. State of the system after collision

Since the Hamiltonian is zero for $t < 0$, we have:

$$\begin{aligned} |\psi(0)\rangle &= |\psi(-\infty)\rangle = |+,-\rangle \\ &= \frac{1}{\sqrt{2}}\left[|1,0\rangle + |0,0\rangle\right] \end{aligned} \tag{40}$$

The results of the preceding section concerning the eigenstates and eigenvalues of $W = a\mathbf{S}_1 \cdot \mathbf{S}_2$ are applicable between times 0 and T and enable us to calculate $|\psi(T)\rangle$:

$$|\psi(T)\rangle = \frac{1}{\sqrt{2}}\left[|1,0\rangle \mathrm{e}^{-iE_1T/\hbar} + |0,0\rangle \mathrm{e}^{-iE_0T/\hbar}\right] \tag{41}$$

Multiplying (41) by the global phase factor $\mathrm{e}^{i(E_0+E_1)T/2\hbar}$ (of no physical importance), setting $E_1 - E_0 = \hbar\Omega$ [*cf.* formula (23)], and returning to the $\{|\varepsilon_1, \varepsilon_2\rangle\}$ basis, we find:

$$|\psi(T)\rangle = \cos\frac{\Omega T}{2}|+,-\rangle - i\sin\frac{\Omega T}{2}|-,+\rangle \tag{42}$$

Finally, since the Hamiltonian is zero for $t > T$, we have:

$$|\psi(+\infty)\rangle = |\psi(T)\rangle \tag{43}$$

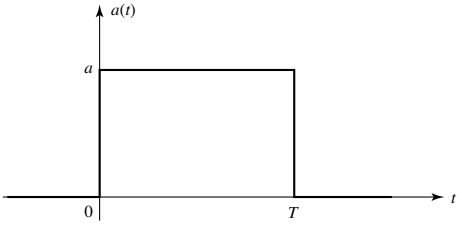


Figure 10: Simplified curve used to represent schematically the variation of the coupling constant $a(t)$ during the collision.

Comment:

The calculation could be performed for an arbitrary function $a(t)$ of the type shown in Figure 9. It would then be necessary to replace, in the preceding formula, $aT = \dfrac{\Omega T}{\hbar}$ by $\displaystyle\int_{-\infty}^{+\infty} a(t)\,\mathrm{d}t$ (*cf.* exercise 2 of Complement E_{XIII}).

4-c. Discussion. Correlation introduced by the collision

If the condition:

$$\frac{\Omega T}{2} = \frac{\pi}{2} + k\pi \quad , \quad k \text{ an integer } \gtrless 0 \tag{44}$$

is satisfied, we see from (42) that:

$$|\psi(+\infty)\rangle = |-,+\rangle \tag{45}$$

The orientation of the two spins, in this case, is exchanged during the collision.

On the other hand, if:

$$\frac{\Omega T}{2} = k\pi \quad , \quad k \text{ an integer } \gtrless 0 \tag{46}$$

we find that:

$$|\psi(+\infty)\rangle = |+,-\rangle = |\psi(-\infty)\rangle \tag{47}$$

In this case, the collision has no effect on the orientation of the spins.

For other values of T, we have:

$$|\psi(+\infty)\rangle = \alpha|+,-\rangle + \beta|-,+\rangle \tag{48}$$

with α and β simultaneously non-zero. The state of the two-spin system has been transformed by the collision into a linear superposition of the two states $|+,-\rangle$ and $|-,+\rangle$. $|\psi(+\infty)\rangle$ *is therefore no longer a tensor product,* although $|\psi(-\infty)\rangle$ was one: the interaction of the two spins has introduced *correlations* between them.

To see this, we shall analyze an experiment in which, after the collision, an observer [observer (1)] measures S_{1z}. According to formula (48) for $|\psi(+\infty)\rangle$, he has the probability $|\alpha|^2$ of finding $+\hbar/2$ and $|\beta|^2$ of finding $-\hbar/2$ [according to (42), $|\alpha|^2 + |\beta|^2 = 1$]. Assume that he finds $-\hbar/2$. Immediately after this measurement, the state of the total system is, according to the wave packet reduction postulate, $|-,+\rangle$. If, at this moment, a second observer [observer (2)] measures S_{2z}, he will always find $+\hbar/2$. Similarly, it can easily be shown that if observer (1) finds the result $+\hbar/2$, observer (2) will then always find $-\hbar/2$. Thus, the result obtained by observer (1) critically influences the result that observer (2) will obtain later, even if at the time of these two measurements, the particles are extremely far apart. This apparently paradoxical result (the Einstein-Podolsky-Rosen argument, *cf.* Chap. XXI, § F) reflects the existence of a strong correlation between the two spins, which has appeared because of their interaction during the collision.

Note, finally, that if we are concerned with only one of the two spins, it is impossible to describe its state after the collision by a state vector, since, according to formula (48), $|\psi(+\infty)\rangle$ is not a tensor product. Spin (1), for example, can be described in this case only by a density operator (*cf.* Complement $\mathrm{E_{III}}$). Let:

$$\rho = |\psi(+\infty)\rangle \langle\psi(+\infty)| \tag{49}$$

be the density operator of the total two-spin system. According to the results of Complement $\mathrm{E_{III}}$ (§ 5-b), the density operator of spin (1) can be obtained by taking the partial trace of ρ with respect to the spin variables of particle (2):

$$\rho(1) = \mathrm{Tr}_2\, \rho \tag{50}$$

Similarly:

$$\rho(2) = \mathrm{Tr}_1\, \rho \tag{51}$$

It is easy to calculate, from expression (48) for $|\psi(+\infty)\rangle$, the matrix representing ρ in the four-state basis, $\{|+,+\rangle, |+,-\rangle, |-,+\rangle, |-,-\rangle\}$, arranged in this order. We find:

$$\rho = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & |\alpha|^2 & \alpha\beta^* & 0 \\ 0 & \beta\alpha^* & |\beta|^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{52}$$

Applying (50) and (51), we then find:

$$\rho(1) = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix} \tag{53}$$

$$\rho(2) = \begin{pmatrix} |\beta|^2 & 0 \\ 0 & |\alpha|^2 \end{pmatrix} \tag{54}$$

Starting with expressions (53) and (54), we can form:

$$\rho' = \rho(1) \otimes \rho(2) \tag{55}$$

whose matrix representation can be written:

$$\rho' = \begin{pmatrix} |\alpha|^2 |\beta|^2 & 0 & 0 & 0 \\ 0 & |\alpha|^4 & 0 & 0 \\ 0 & 0 & |\beta|^4 & 0 \\ 0 & 0 & 0 & |\alpha|^2 |\beta|^2 \end{pmatrix} \tag{56}$$

We see that ρ' is different from ρ, reflecting the existence of correlations between the two spins.

References and suggestions for further readings:

The vector model of the atom: Eisberg and Resnick (1.3), Chap. 8, § 5; Cagnac and Pebay-Peyroula (11.2), Chaps. XVI, § 3B and XVII, §§ 3E and 4C.

The Einstein-Podolsky-Rosen paradox/argument: see references of Complement $\mathrm{D_{III}}$.

Complement G_X

Exercises

1. Consider a deuterium atom (composed of a nucleus of spin $I = 1$ and an electron). The electronic angular momentum is $\mathbf{J} = \mathbf{L}+\mathbf{S}$, where $\mathbf{L}$ is the orbital angular momentum of the electron and $\mathbf{S}$ is its spin. The total angular momentum of the atom is $\mathbf{F} = \mathbf{J} + \mathbf{I}$, where $\mathbf{I}$ is the nuclear spin. The eigenvalues of $\mathbf{J}^2$ and $\mathbf{F}^2$ are $J(J+1)\hbar^2$ and $F(F+1)\hbar^2$ respectively.

a. What are the possible values of the quantum numbers J and F for a deuterium atom in the $1s$ ground state?

b. Same question for deuterium in the $2p$ excited state.

2. The hydrogen atom nucleus is a proton of spin $I = 1/2$.

a. In the notation of the preceding exercise, what are the possible values of the quantum numbers J and F for a hydrogen atom in the $2p$ level?

b. Let $\{|n, l, m\rangle\}$ be the stationary states of the Hamiltonian H_0 of the hydrogen atom studied in § C of Chapter VII.

Let $\{|n, l, s, J, M_J\rangle\}$ be the basis obtained by adding $\mathbf{L}$ and $\mathbf{S}$ to form $\mathbf{J}$ ($M_J\hbar$ is the eigenvalue of J_z); and let $\{|n, l, s, J, I, F, M_F\rangle\}$ be the basis obtained by adding $\mathbf{J}$ and $\mathbf{I}$ to form $\mathbf{F}$ ($M_F\hbar$ is the eigenvalue of F_z).

The magnetic moment operator of the electron is:

$$\mathbf{M} = \mu_B(\mathbf{L} + 2\mathbf{S})/\hbar$$

In each of the subspaces $\mathcal{E}(n = 2, l = 1, s = 1/2, J, I = 1/2, F)$ arising from the $2p$ level and subtended by the $2F + 1$ vectors

$$\left|n = 2, l = 1, s = \frac{1}{2}, J, I = \frac{1}{2}, F, M_F\right\rangle$$

corresponding to fixed values of J and F, the projection theorem (*cf.* Complement D_X, §§ 2-c and 3) enables us to write:

$$\mathbf{M} = g_{JF}\mu_B\mathbf{F}/\hbar$$

Calculate the various possible values of the Landé factors g_{JF} corresponding to the $2p$ level.

3. Consider a system composed of two spin 1/2 particles whose orbital variables are ignored. The Hamiltonian of the system is:

$$H = \omega_1 S_{1z} + \omega_2 S_{2z}$$

where S_{1z} and S_{2z} are the projections of the spins $\mathbf{S}_1$ and $\mathbf{S}_2$ of the two particles onto Oz, and ω_1 and ω_2 are real constants.

a. The initial state of the system, at time $t = 0$, is:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}\,[|+\,-\rangle + |-\,+\rangle]$$

(with the notation of § B of Chapter X). At time t, $\mathbf{S}^2 = (\mathbf{S}_1 + \mathbf{S}_2)^2$ is measured. What results can be found, and with what probabilities?

b. If the initial state of the system is arbitrary, what Bohr frequencies can appear in the evolution of $\langle \mathbf{S}^2 \rangle$? Same question for $S_x = S_{1x} + S_{2x}$.

4. Consider a particle (a) of spin 3/2 which can disintegrate into two particles, (b) of spin 1/2 and (c) of spin 0. We place ourselves in the rest frame of (a). Total angular momentum is conserved during the disintegration.

a. What values can be taken on by the relative orbital angular momentum of the two final particles? Show that there is only one possible value if the parity of the relative orbital state is fixed. Would this result remain valid if the spin of particle (a) were greater than 3/2?

b. Assume that particle (a) is initially in the spin state characterized by the eigenvalue $m_a\hbar$ of its spin component along Oz. We know that the final orbital state has a definite parity. Is it possible to determine this parity by measuring the probabilities of finding particle (b) either in the state $|+\rangle$ or in the state $|-\rangle$ (you may use the general formulas of Complement A_X, § 2)?

5. Let $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3$ be the total angular momentum of three spin 1/2 particles (whose orbital variables will be ignored). Let $|\varepsilon_1, \varepsilon_2, \varepsilon_3\rangle$ be the eigenstates common to S_{1z}, S_{2z}, S_{3z}, of respective eigenvalues $\varepsilon_1\,\hbar/2$, $\varepsilon_2\,\hbar/2$, $\varepsilon_3\,\hbar/2$. Give a basis of eigenvectors common to $\mathbf{S}^2$ and S_z, in terms of the kets $|\varepsilon_1, \varepsilon_2, \varepsilon_3\rangle$. Do these two operators form a C.S.C.O.? (Begin by adding two of the spins in order to obtain a partial angular momentum, and then add it to the third one.)

6. Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be the intrinsic angular momenta of two spin 1/2 particles, $\mathbf{R}_1$ and $\mathbf{R}_2$, their position observables, and m_1 and m_2, their masses (with $\mu = \frac{m_1 m_2}{m_1 + m_2}$, the reduced mass). Assume that the interaction W between the two particles is of the form:

$$W = U(R) + V(R)\frac{\mathbf{S}_1 \cdot \mathbf{S}_2}{\hbar^2}$$

where $U(R)$ and $V(R)$ depend only on the distance $R = |\mathbf{R}_1 - \mathbf{R}_2|$ between the particles.

a. Let $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$ be the total spin of the two particles.

α. Show that:

$$P_1 = \frac{3}{4} + \frac{\mathbf{S}_1 \cdot \mathbf{S}_2}{\hbar^2}$$
$$P_0 = \frac{1}{4} + \frac{\mathbf{S}_1 \cdot \mathbf{S}_2}{\hbar^2}$$

are the projectors onto the total spin states $S = 1$ and $S = 0$ respectively.

β. Show from this that $W = W_1(R)\,P_1 + W_0(R)\,P_0$, where $W_1(R)$ and $W_0(R)$, are two fonctions of R, to be expressed in terms of $U(R)$ and $V(R)$.

b. Write the Hamiltonian H of the "relative particle" in the center of mass frame; $\mathbf{P}$ denotes the momentum of this relative particle. Show that H commutes with $\mathbf{S}^2$ and does not depend on S_z. Show from this that it is possible to study separately the eigenstates of H corresponding to $S = 1$ and $S = 0$.

Show that one can find eigenstates of H, with eigenvalue E, of the form:

$$|\psi_E\rangle = \lambda_{00}\,\left|\varphi_E^0\right\rangle\,|S=0, M=0\rangle + \sum_{M=-1}^{+1} \lambda_{1M}\,\left|\varphi_E^1\right\rangle\,|S=1, M\rangle$$

where λ_{00} and λ_{1M} are constants, and $\left|\varphi_E^0\right\rangle$ and $\left|\varphi_E^1\right\rangle$ are kets of the state space $\mathcal{E}_{\mathbf{r}}$ of the relative particle ($M\hbar$ is the eigenvalue of S_z). Write the eigenvalue equations satisfied by $\left|\varphi_E^0\right\rangle$ and $\left|\varphi_E^1\right\rangle$.

c. We want to study collisions between the two particles under consideration. Let $E = \hbar^2k^2/2\mu$ be the energy of the system in the center of mass frame. We assume in all that follows that, before the collision, one of the particles is in the $|+\rangle$ spin state, and the other one, in the $|-\rangle$ spin state. Let $\left|\psi_k^{\uparrow\downarrow}\right\rangle$ be the corresponding stationary scattering state (*cf.* Chap. VIII, § B). Show that:

$$\left|\psi_k^{\uparrow\downarrow}\right\rangle = \frac{1}{\sqrt{2}}\,\left|\varphi_k^0\right\rangle\,|S=0, M=0\rangle + \frac{1}{\sqrt{2}}\,\left|\varphi_k^1\right\rangle\,|S=1, M=0\rangle$$

where $\left|\varphi_k^0\right\rangle$ and $\left|\varphi_k^1\right\rangle$ are the stationary scattering states for a spinless particle of mass μ, scattered respectively by the potentials $W_0(R)$ and $W_1(R)$.

d. Let $f_0(\theta)$ and $f_1(\theta)$ be the scattering amplitudes that correspond to $\left|\varphi_k^0\right\rangle$ and $\left|\varphi_k^1\right\rangle$. Calculate, in terms of $f_0(\theta)$ and $f_1(\theta)$, the scattering cross section $\sigma_b(\theta)$ of the two particles in the θ direction, with simultaneous flip of the two spins (the spin which was in the $|+\rangle$ state goes into the $|-\rangle$ state, and vice versa).

e. Let δ_l^0 and δ_l^1 be the phase shifts of the l partial waves associated respectively with $W_0(R)$ and $W_1(R)$ (*cf.* Chap. VIII, § C-3). Show that the total scattering cross section σ_b, with simultaneous flip of the two spins, is equal to:

$$\sigma_b = \frac{\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2(\delta_l^1 - \delta_l^0)$$

7. We define the standard components of a vector operator $\mathbf{V}$ as the three operators:

$$\begin{cases} V_1^{(1)} = -\dfrac{1}{\sqrt{2}}(V_x + iV_y) \\ V_0^{(1)} = V_z \\ V_{-1}^{(1)} = \dfrac{1}{\sqrt{2}}(V_x - iV_y) \end{cases}$$

Using the standard components $V_p^{(1)}$ and $W_q^{(1)}$ of the two vector operators $\mathbf{V}$ and $\mathbf{W}$, we construct the operators:

$$\left[V^{(1)} \otimes W^{(1)}\right]_M^{(K)} = \sum_p \sum_q \langle 1,1;p,q \mid K,M\rangle\, V_p^{(1)} W_q^{(1)}$$

where the $\langle 1,1;p,q \mid K,M\rangle$ are the Clebsch-Gordan coefficients entering into the addition of two angular momenta 1 (these coefficients can be obtained from the results of § 1 of Complement A_X).

a. Show that $\left[V^{(1)} \otimes W^{(1)}\right]_0^{(0)}$ is proportional to the scalar product $\mathbf{V} \cdot \mathbf{W}$ of the two vector operators.

b. Show that the three operators $\left[V^{(1)} \otimes W^{(1)}\right]_M^{(1)}$ are proportional to the three standard components of the vector operator $\mathbf{V} \times \mathbf{W}$.

c. Express the five components $\left[V^{(1)} \otimes W^{(1)}\right]_M^{(2)}$ in terms of the various operators V_z, $V_\pm = V_x \pm iV_y$, W_z, $W_\pm = W_x \pm iW_y$.

d. We choose $\mathbf{V} = \mathbf{W} = \mathbf{R}$, where $\mathbf{R}$ is the position observable of a particle. Show that the five operators $\left[R^{(1)} \otimes R^{(1)}\right]_M^{(2)}$ are proportional to the five components Q_2^M of the electric quadrupole moment operator of this particle [*cf.* formula (29) of Complement E_X].

e. We choose $\mathbf{V} = \mathbf{W} = \mathbf{L}$, where $\mathbf{L}$ is the orbital angular momentum of the particle. Express the five operators $\left[L^{(1)} \otimes L^{(1)}\right]_M^{(2)}$ in terms of L_z, L_+, L_-. What are the selection rules satisfied by these five operators in a standard basis $\{|k,l,m\rangle\}$ of eigenstates common to $\mathbf{L}^2$ and L_z (in other words, on what conditions is the matrix element

$$\langle k,l,m| \left[L^{(1)} \otimes L^{(1)}\right]_M^{(2)} |k',l',m'\rangle$$

non-zero)?

8. Irreducible tensor operators; Wigner-Eckart theorem

The $2K+1$ operators $T_Q^{(K)}$, with K an integer ≥ 0 and $Q = -K, -K+1, \ldots +K$, are, by definition, the $2K+1$ components of an irreducible tensor operator of rank K if they satisfy the following commutation relations with the total angular momentum $\mathbf{J}$ of the physical system:

$$\left[J_z, T_Q^{(K)}\right] = \hbar Q T_Q^{(K)} \tag{1}$$

$$\left[J_+, T_Q^{(K)}\right] = \hbar\sqrt{K(K+1) - Q(Q+1)}\, T_{Q+1}^{(K)} \tag{2}$$

$$\left[J_-, T_Q^{(K)}\right] = \hbar\sqrt{K(K+1) - Q(Q-1)}\, T_{Q-1}^{(K)} \tag{3}$$

a. Show that a scalar operator is an irreducible tensor operator of rank $K = 0$, and that the three standard components of a vector operator (*cf.* exercise 7) are the components of an irreducible tensor operator of rank $K = 1$.

b. Let $\{|k, J, M\rangle\}$ be a standard basis of common eigenstates of $\mathbf{J}^2$ and J_z. By taking both sides of (1) to have the same matrix elements between $|k, J, M\rangle$ and $|k', J', M'\rangle$, show that $\left\langle k, J, M \mid T_Q^{(K)} \mid k', J', M' \right\rangle$ is zero if M is not equal to $Q + M'$.

c. Proceeding in the same way as above with relations (2) and (3), show that the $(2J+1)(2K+1)(2J'+1)$ matrix elements $\left\langle k, J, M \mid T_Q^{(K)} \mid k', J', M' \right\rangle$ corresponding to fixed values of k, J, K, k', J' satisfy recurrence relations identical to those satisfied by the $(2J+1)(2K+1)(2J'+1)$ Clebsch-Gordan coefficients $\langle J', K; M', Q \mid J, M\rangle$ (*cf.* Complement $\mathrm{B_X}$, §§ 1-c and 2) corresponding to fixed values of J, K, J'.

d. Show that:

$$\left\langle k, J, M \mid T_Q^{(K)} \mid k', J', M' \right\rangle = \alpha \langle J', K; M', Q \mid J, M\rangle \tag{4}$$

where α is a constant depending only on k, J, K, k', J', which is usually written in the form:

$$\alpha = \frac{1}{\sqrt{2J+1}} \left\langle k, J \parallel T^{(K)} \parallel k', J' \right\rangle$$

e. Show that, conversely, if $(2K+1)$ operators $T_Q^{(K)}$ satisfy relation (4) for all $|k, J, M\rangle$ and $|k', J', M'\rangle$, they satisfy relations (1), (2) and (3), that is, they constitute the $(2K+1)$ components of an irreducible tensor operator of rank K.

f. Show that, for a spinless particle, the electric multipole moment operators Q_l^m introduced in Complement $\mathrm{E_X}$ are irreducible tensor operators of rank l in the state space $\mathcal{E}_\mathbf{r}$ of this particle. Show that, in addition, when the spin degrees of freedom are taken into account, the operators Q_l^m remain irreducible tensor operators in the state space $\mathcal{E}_\mathbf{r} \otimes \mathcal{E}_s$ (where $\mathcal{E}_s$ is the spin state space).

g. Derive the selection rules satisfied by the Q_l^m in a standard basis $\{|k, l, J, M_J\rangle\}$ obtained by adding the orbital angular momentum $\mathbf{L}$ and the spin $\mathbf{S}$ of the particle to form the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ [$l(l+1)\hbar^2$, $J(J+1)\hbar^2$, $M_J\hbar$ are the eigenvalues of $\mathbf{L}^2$, $\mathbf{J}^2$, J_z respectively].

9. Let $A_{Q_1}^{(K_1)}$ be an irreducible tensor operator (exercise 8) of rank K_1 acting in a state space $\mathcal{E}_1$, and $B_{Q_2}^{(K_2)}$, an irreducible tensor operator of rank K_2 acting in a state space $\mathcal{E}_2$. With $A_{Q_1}^{(K_1)}$ and $B_{Q_2}^{(K_2)}$, we construct the operator:

$$C_Q^{(K)} = \left[A^{(K_1)} \otimes B^{(K_2)}\right]_Q^{(K)} = \sum_{Q_1 Q_2} \langle K_1, K_2; Q_1, Q_2 \mid K, Q\rangle \, A_{Q_1}^{(K_1)} B_{Q_2}^{(K_2)}$$

a. Using the recurrence relations for Clebsch-Gordan coefficients (*cf.* Complement $\mathrm{B_X}$), show that the $C_Q^{(K)}$ satisfy commutation relations (1), (2) and (3) of exercise 8 with the total angular momentum $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$ of the system. Show that the $C_Q^{(K)}$ are the components of an irreducible tensor operator of rank K.

b. Show that the operator $\sum_Q (-1)^Q A_Q^{(K)} B_{-Q}^{(K)}$ is a scalar operator (you may use the results of § 3-d of Complement B_X).

10. Addition of three angular momenta

Let $\mathcal{E}(1)$, $\mathcal{E}(2)$, $\mathcal{E}(3)$ be the state spaces of three systems (1), (2) and (3), of angular momenta $\mathbf{J}_1$, $\mathbf{J}_2$, $\mathbf{J}_3$. We shall write $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 + \mathbf{J}_3$ for the total angular momentum. Let $\{|k_a, j_a, m_a\rangle\}$, $\{|k_b, j_b, m_b\rangle\}$, $\{|k_c, j_c, m_c\rangle\}$ be the standard bases of $\mathcal{E}(1)$, $\mathcal{E}(2)$, $\mathcal{E}(3)$, respectively. To simplify the notation, we shall omit the indices k_a, k_b, k_c, as we did in Chapter X.

We are interested in the eigenstates and eigenvalues of the total angular momentum in the subspace $\mathcal{E}(j_a, j_b, j_c)$ subtended by the kets:

$$\{|j_a m_a\rangle\, |j_b m_b\rangle\, |j_c m_c\rangle\}$$
$$-j_a \leq m_a \leq j_a \;\;,\;\; -j_b \leq m_b \leq j_b \;\;,\;\; -j_c \leq m_c \leq j_c \tag{1}$$

We want to add j_a, j_b, j_c to form an eigenstate of $\mathbf{J}^2$ and J_z characterized by the quantum numbers j_f and m_f. We shall denote by:

$$|j_a, (j_b j_c) j_e; j_f m_f\rangle \tag{2}$$

such a normalized eigenstate obtained by first adding j_b to j_c to form an angular momentum j_e, then adding j_a to j_e to form the state $|j_f m_f\rangle$. One could also add j_a and j_b to form j_g and then add j_c to j_g to form the normalized state $|j_f m_f\rangle$, written:

$$|(j_a j_b) j_g, j_c; j_f m_f\rangle \tag{3}$$

a. Show that the system of kets (2), corresponding to the various possible values of j_e, j_f, m_f, forms an orthonormal basis in $\mathcal{E}(j_a, j_b, j_c)$. Same question for the system of kets (3), corresponding to the various values of j_g, j_f, m_f.

b. Show, by using the operators J_+ and J_-, that the scalar product of kets $\langle (j_a j_b) j_g, j_c; j_f m_f \mid j_a, (j_b j_c) j_e; j_f m_f\rangle$ does not depend on m_f, denoting such a scalar product by $\langle (j_a j_b) j_g, j_c; j_f \mid j_a, (j_b j_c) j_e; j_f\rangle$.

c. Show that:

$$|j_a, (j_b j_c) j_e; j_f m_f\rangle = \sum_{j_g} \langle (j_a j_b) j_g, j_c; j_f \mid j_a, (j_b j_c) j_e; j_f\rangle\, |(j_a, j_b) j_g, j_c; j_f m_f\rangle \tag{4}$$

d. Using the Clebsch-Gordan coefficients, write the expansions for vectors (2) and (3) on the basis (1). Show that:

$$\sum_{m_e} \langle j_b, j_c; m_b, m_c \mid j_e, m_e\rangle \langle j_a, j_e; m_a, m_e \mid j_f, m_f\rangle =$$
$$\sum_{j_g m_g} \langle j_a, j_b; m_a, m_b \mid j_g, m_g\rangle \langle j_g, j_c; m_g, m_c \mid j_f, m_f\rangle$$
$$\times \langle (j_a j_b) j_g, j_c; j_f \mid j_a, (j_b j_c) j_e; j_f\rangle \tag{5}$$

e. Starting with relation (5), prove, using the Clebsch-Gordan coefficient orthogonality relations, the following relations:

$$\sum_{m_a m_b m_e} \langle j_b, j_c; m_b, m_c | j_e, m_e \rangle \langle j_a, j_e; m_a, m_e | j_f, m_f \rangle \langle j_d, m_d | j_a, j_b; m_a, m_b \rangle$$
$$= \langle j_d, j_c; m_d, m_c | j_f, m_f \rangle \langle (j_a j_b) j_d, j_c; j_f | j_a, (j_b j_c) j_e; j_f \rangle \qquad (6)$$

as well as:

$$\langle (j_a j_b) j_d, j_c; j_f \mid j_a, (j_b j_c) j_e; j_f \rangle =$$
$$\frac{1}{2j_f + 1} \sum_{m_a m_b m_c m_d m_e m_f} \langle j_b, j_c; m_b, m_c \mid j_e, m_e \rangle$$
$$\times \langle j_a, j_e; m_a, m_e \mid j_f, m_f \rangle \langle j_d, m_d \mid j_a, j_b; m_a, m_b \rangle \langle j_f, m_f \mid j_d, j_c; m_d, m_c \rangle \qquad (7)$$

References

Exercises 8 et 9: see references of Complement D_X.
Exercise 10: Edmonds (2.21), Chap. 6; Messiah (1.17), § XIII-29 and App. C; Rose (2.19), App. 1.

Chapter XI

Stationary perturbation theory

A Description of the method 1116
A-1 Statement of the problem 1116
A-2 Approximate solution of the $H(\lambda)$ eigenvalue equation 1117
B Perturbation of a non-degenerate level 1120
B-1 First-order corrections 1120
B-2 Second-order corrections 1122
C Perturbation of a degenerate state 1124

The quantum mechanical study of conservative physical systems (that is, systems whose Hamiltonians are not explicitly time-dependent) is based on the eigenvalue equation of the Hamiltonian operator. We have already encountered two important examples of physical systems (the harmonic oscillator and the hydrogen atom) whose Hamiltonians are simple enough for their eigenvalue equations to be solved exactly. However, this happens in only a very small number of problems. In general, the equation is too complicated for us to be able to find its solutions in an analytic form[1]. For example, we do not know how to treat many-electron atoms, even helium, exactly. Besides, the hydrogen atom theory explained in Chapter VII (§ C) takes into account only the electrostatic interaction between the proton and the electron; when relativistic corrections (such as magnetic forces) are added to this principal interaction, the equation obtained for the hydrogen atom can no longer be solved analytically. We must then resort to solving it numerically with a computer. There exist, however, *approximation methods* that enable us to obtain analytically approximate solutions of the basic eigenvalue equation in certain cases. In this chapter, we shall study one of these methods, known as "stationary perturbation theory"[2]. (In Chapter XIII, we shall describe "time-dependent perturbation theory",

[1] Of course, this phenomenon is not limited to the domain of quantum mechanics. In all fields of physics, there are very few problems that can be treated completely analytically.

[2] Perturbation theory also exists in classical mechanics, where it is, in principle, entirely analogous to the one we shall describe here.

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

which is used to treat systems whose Hamiltonians contain explicitly time-dependent terms.)

Stationary perturbation theory is very widely used in quantum physics, since it reflects the physicist's usual approach to problems. In studying a phenomenon or a physical system, one begins by isolating the principal effects that are responsible for the main features of this phenomenon or this system. When they have been understood, one tries to explain the "finer" details by taking into account less important effects that were neglected in the first approximation. It is in treating these secondary effects that one commonly uses perturbation theory. In Chapter XII, we shall see, for example, the importance of perturbation theory in atomic physics: it will enable us to calculate the relativistic corrections in the case of the hydrogen atom. Similarly, Complement B_{XIV}, which is devoted to the helium atom, indicates how perturbation theory allows us to treat many-electron atoms. Numerous other applications of perturbation theory are given in the complements of this chapter and the following ones.

Let us mention, finally, another often used approximation method, the variational method, which we shall present in Complement E_{XI}. We shall briefly examine its applications in solid state physics (Complement F_{XI}) and in molecular physics (Complement G_{XI}).

A. Description of the method

A-1. Statement of the problem

Perturbation theory is applicable when the Hamiltonian H of the system being studied can be put in the form:

$$H = H_0 + W \tag{A-1}$$

where the eigenstates and eigenvalues of H_0 are known, and where W is much smaller than H_0. The operator H_0, which is time-independent, is called the "unperturbed Hamiltonian" and W, the "perturbation". If W is not time-dependent, we say that we are dealing with a "stationary perturbation"; this is the case we are considering in this chapter (the case of time-dependent perturbations will be studied in Chapter XIII). The problem is then to find the modifications produced in the energy levels of the system and in its stationary states by the addition of the perturbation W.

When we say that W is much smaller than H_0, this means that the matrix elements of W are much smaller[3] than those of H_0. To make this more explicit, we shall assume that W is proportional to a real parameter λ which is dimensionless and much smaller than 1:

$$W = \lambda \hat{W} \quad \text{with} \quad \lambda \ll 1 \tag{A-2}$$

(where $\hat{W}$ is an operator whose matrix elements are comparable to those of H_0). Perturbation theory consists of expanding the eigenvalues and eigenstates of H in powers of λ, keeping only a finite number of terms (often only one or two) of these expansions.

We shall assume the eigenstates and eigenvalues of the unperturbed Hamiltonian H_0 to be known. In addition, we shall assume, that *the unperturbed energies form a*

[3] More precisely, the important point is that the matrix elements of W are much smaller than the differences between eigenvalues of H_0 (*cf.* comment of § B-1-b).

discrete spectrum, and we shall label them by an integral index $p : E_p^0$. The corresponding eigenstates will be denoted by $|\varphi_p^i\rangle$, the additional index i permitting us, in the case of a degenerate eigenvalue E_p, to distinguish between the various vectors of an orthonormal basis of the associated eigensubspace. We therefore have:

$$H_0|\varphi_p^i\rangle = E_p^0|\varphi_p^i\rangle \tag{A-3}$$

where the set of vectors $|\varphi_p^i\rangle$ forms an orthonormal basis of the state space:

$$\langle\varphi_p^i|\varphi_{p'}^{i'}\rangle = \delta_{pp'}\,\delta_{ii'} \tag{A-4a}$$

$$\sum_p \sum_i |\varphi_p^i\rangle\,\langle\varphi_p^i| = 1 \tag{A-4b}$$

If we substitute (A-2) into (A-1), we can consider the Hamiltonian of the system to be continuously dependent on the parameter λ characterizing the intensity of the perturbation:

$$H(\lambda) = H_0 + \lambda\hat{W} \tag{A-5}$$

When λ is equal to zero, $H(\lambda)$ is equal to the unperturbed Hamiltonian H_0. The eigenvalues $E(\lambda)$ of $H(\lambda)$ generally depend on λ, and Figure 1 represents possible forms of their variations with respect to λ.

An eigenvector of $H(\lambda)$ is associated with each curve of Figure 1. For a given value of λ, these vectors form a basis of the state space [$H(\lambda)$ is an observable]. When λ is much smaller than 1, the eigenvalues $E(\lambda)$ and the eigenvectors $|\psi(\lambda)\rangle$ of $H(\lambda)$ remain very close to those of $H_0 = H(\lambda = 0)$, which they approach when $\lambda \longrightarrow 0$.

The operator $H(\lambda)$ may, of course, have one or several degenerate eigenvalues. For example, in Figure 1, the double curve represents a doubly degenerate energy (the one which approaches E_4^0 when $\lambda \longrightarrow 0$), which corresponds, for all λ, to a two-dimensional eigensubspace. It is also possible for several distinct eigenvalues $E(\lambda)$ to approach the same unperturbed energy[4] E_p^0 when $\lambda \longrightarrow 0$; this happens for instance for E_3^0 in Figure 1. In such a case, we say that the effect of the perturbation is to remove the degeneracy of the corresponding eigenvalue of H_0.

In the following section, we shall give an approximate solution of the eigenvalue equation of $H(\lambda)$ for $\lambda \ll 1$ [of course, we assume that we cannot solve this equation exactly; otherwise it would not be necessary to resort to perturbation theory to find the eigenstates and eigenvalues of $H = H(\lambda)$].

A-2. Approximate solution of the $H(\lambda)$ eigenvalue equation

We are looking for the eigenstates $|\psi(\lambda)\rangle$ and eigenvalues $E(\lambda)$ of the Hermitian operator $H(\lambda)$:

$$H(\lambda)\,|\psi(\lambda)\rangle = E(\lambda)\,|\psi(\lambda)\rangle \tag{A-6}$$

[4] Additional degeneracies may appear for particular non-zero values of λ (crossing at $\lambda = \lambda_1$ in Figure 1). We shall assume here that λ is small enough to avoid such a situation.

We shall assume[5] that $E(\lambda)$ and $|\psi(\lambda)\rangle$ can be expanded in powers of λ in the form:

$$E(\lambda) = \varepsilon_0 + \lambda\varepsilon_1 + \ldots + \lambda^q\varepsilon_q + \ldots \tag{A-7a}$$

$$|\psi(\lambda)\rangle = |0\rangle + \lambda|1\rangle + \ldots + \lambda^q\,|q\rangle + \ldots \tag{A-7b}$$

We then substitute these two expansions, as well as definition (A-5) of $H(\lambda)$, into equation (A-6):

$$\left(H_0 + \lambda\hat{W}\right)\left[\sum_{q=0}^{\infty}\lambda^q\,|q\rangle\right] = \left[\sum_{q'=0}^{\infty}\lambda^{q'}\varepsilon_{q'}\right]\left[\sum_{q=0}^{\infty}\lambda^q\,|q\rangle\right] \tag{A-8}$$

We require this equation to be satisfied for λ small but arbitrary. We must therefore equate the coefficients of successive powers of λ on both sides. This leads to:

- for 0th-order terms in λ:

$$H_0|0\rangle = \varepsilon_0|0\rangle \tag{A-9}$$

- for 1st-order terms:

$$(H_0 - \varepsilon_0)\,|1\rangle + \left(\hat{W} - \varepsilon_1\right)|0\rangle = 0 \tag{A-10}$$

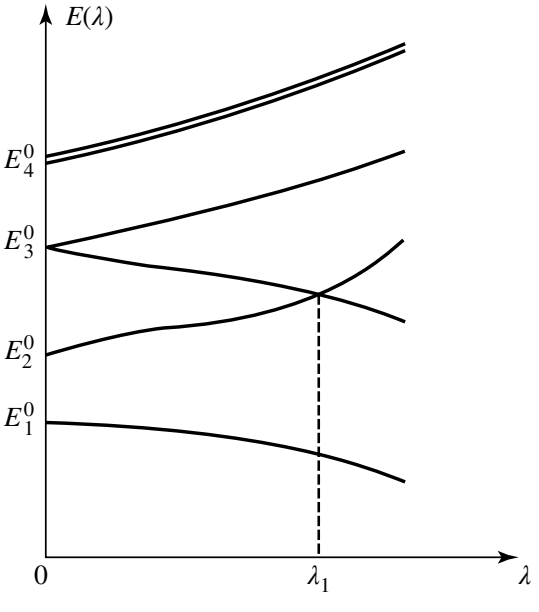


Figure 1: Variation of the eigenvalues $E(\lambda)$ of the Hamiltonian $H(\lambda) = H_0 + \lambda\hat{W}$ with respect to λ. Each curve corresponds to an eigenstate of $H(\lambda)$. For $\lambda = 0$, we obtain the spectrum of H_0. We have assumed here that the eigenvalues E_3^0 and E_4^0 are doubly degenerate; application of the perturbation $\lambda\hat{W}$ removes the degeneracy of E_3^0, but not that of E_4^0. An additional two-fold degeneracy appears for $\lambda = \lambda_1$.

[5]This is not obvious from a mathematical point of view, the basic problem being the convergence of the series (A-7).

- for 2nd-order terms:

$$(H_0 - \varepsilon_0)\,|2\rangle + \left(\hat{W} - \varepsilon_1\right)|1\rangle - \varepsilon_2|0\rangle = 0 \tag{A-11}$$

- for qth-order terms:

$$(H_0 - \varepsilon_0)\,|q\rangle + \left(\hat{W} - \varepsilon_1\right)|q-1\rangle - \varepsilon_2|q-2\rangle \ldots - \varepsilon_q|0\rangle = 0 \tag{A-12}$$

We shall confine ourselves here to the study of the first three equations, that is, we shall neglect, in expansions (A-7), terms of orders higher than 2 in λ.

We know that the eigenvalue equation (A-6) defines $|\psi(\lambda)\rangle$ only to within a constant factor. We can therefore choose the norm of $|\psi(\lambda)\rangle$ and its phase: *we shall require $|\psi(\lambda)\rangle$ to be normalized, and we shall choose its phase such that the scalar product $\langle 0|\psi(\lambda)\rangle$ is real.* To 0th order, this implies that the vector denoted by $|0\rangle$ must be normalized:

$$\langle 0|0\rangle = 1 \tag{A-13}$$

Its phase, however, remains arbitrary; we shall see in §§ B and C how it can be chosen in each particular case. To 1st order, the square of the norm of $|\psi(\lambda)\rangle$ can be written:

$$\begin{aligned}\langle\psi(\lambda)|\psi(\lambda)\rangle &= \big[\langle 0| + \lambda\langle 1|\big]\,\big[|0\rangle + \lambda|1\rangle\big] + O(\lambda^2)\\ &= \langle 0|0\rangle + \lambda\big[\langle 1|0\rangle + \langle 0|1\rangle\big] + O(\lambda^2)\end{aligned} \tag{A-14}$$

(where the symbol $O(\lambda^p)$ stands for all the terms of order higher than or equal to p). Using (A-13), we see that this expression is equal to 1 to first order if the λ term is zero. But the choice of phase indicates that the scalar product $\langle 0|1\rangle$ is real (since λ is real). We therefore obtain:

$$\langle 0|1\rangle = \langle 1|0\rangle = 0 \tag{A-15}$$

An analogous argument, for 2nd order in λ, yields:

$$\langle 0|2\rangle = \langle 2|0\rangle = -\frac{1}{2}\langle 1|1\rangle \tag{A-16}$$

and, for qth order:

$$\begin{aligned}\langle 0|q\rangle &= \langle q|0\rangle\\ &= -\frac{1}{2}\big[\langle q-1|1\rangle + \langle q-2|2\rangle + \ldots + \langle 2|q-2\rangle + \langle 1|q-1\rangle\big]\end{aligned} \tag{A-17}$$

When we confine ourselves to second order in λ, the perturbation equations are therefore (A-9), (A-10) and (A-11). With the conventions we have set, we must add conditions (A-13), (A-15) and (A-16).

Equation (A-9) expresses the fact that $|0\rangle$ is an eigenvector of H_0 with the eigenvalue ε_0. ε_0 therefore belongs to the spectrum of H_0. This was to be expected, since each eigenvalue of $H(\lambda)$, when $\lambda \longrightarrow 0$, approaches one of the unperturbed energies. We then

choose a particular value of ε_0, that is, an eigenvalue E_n^0 of H_0. As Figure 1 shows, there can exist one or several different energies $E(\lambda)$ of $H(\lambda)$ that approach E_n^0 when $\lambda \longrightarrow 0$.

Consider the set of eigenstates of $H(\lambda)$ corresponding to the various eigenvalues $E(\lambda)$ that approach E_n^0 when $\lambda \longrightarrow 0$. They span a vector subspace whose dimension clearly cannot vary discontinuously when λ varies in the neighborhood of zero. This dimension is consequently equal to the degeneracy g_n of E_n^0. In particular, if E_n^0 is non-degenerate, it can give rise only to a single energy $E(\lambda)$, and this energy is non-degenerate.

To study the influence of the perturbation W, we shall consider separately the case of non-degenerate, and degenerate levels of H_0.

B. Perturbation of a non-degenerate level

Consider a particular non-degenerate eigenvalue E_n^0 of the unperturbed Hamiltonian H_0. Associated with it is an eigenvector $|\varphi_n\rangle$ which is unique to within a constant factor. We want to determine the modifications in this unperturbed energy and in the corresponding stationary state produced by the addition of the perturbation W to the Hamiltonian.

To do so, we shall use perturbation equations (A-9) through (A-12), as well as conditions (A-13) and (A-15) through (A-17). For the eigenvalue of $H(\lambda)$ that approaches E_n^0 when $\lambda \longrightarrow 0$, we have:

$$\varepsilon_0 = E_n^0 \tag{B-1}$$

which, according to (A-9), implies that $|0\rangle$ must be proportional to $|\varphi_n\rangle$. The vectors $|0\rangle$ and $|\varphi_n\rangle$ are both normalized [*cf.* (A-13)], and we shall choose:

$$|0\rangle = |\varphi_n\rangle \tag{B-2}$$

Thus, when $\lambda \longrightarrow 0$, we again find the unperturbed state $|\varphi_n\rangle$ with the same phase.

We call $E_n(\lambda)$ the eigenvalue of $H(\lambda)$ which, when $\lambda \longrightarrow 0$, approaches the eigenvalue E_n^0 of H_0. We shall assume λ small enough for this eigenvalue to remain non-degenerate, that is, for a unique eigenvector $|\psi_n(\lambda)\rangle$ to correspond to it (in the case of the $n = 2$ level of Figure 1, this is satisfied if $\lambda < \lambda_1$). We shall now calculate the first terms of the expansion of $E_n(\lambda)$ and $|\psi_n(\lambda)\rangle$ in powers of λ.

B-1. First-order corrections

We shall begin by determining ε_1 and the vector $|1\rangle$ from equation (A-10) and condition (A-15).

B-1-a. Energy correction

Projecting equation (A-10) onto the vector $|\varphi_n\rangle$, we obtain:

$$\langle\varphi_n|\,(H_0 - \varepsilon_0)\,|1\rangle + \langle\varphi_n|\,(\hat{W} - \varepsilon_1)\,|0\rangle = 0 \tag{B-3}$$

The first term is zero, since $|\varphi_n\rangle = |0\rangle$ is an eigenvector of the Hermitian operator H_0 with the eigenvalue $E_n^0 = \varepsilon_0$. With (B-2) taken into account, equation (B-3) then yields:

$$\varepsilon_1 = \langle\varphi_n|\,\hat{W}\,|0\rangle = \langle\varphi_n|\,\hat{W}\,|\varphi_n\rangle \tag{B-4}$$

In the case of a non-degenerate state E_n^0, the eigenvalue $E_n(\lambda)$ of H which corresponds to E_n^0 can be written, to first order in the perturbation $W = \lambda\hat{W}$:

$$\boxed{E_n(\lambda) = E_n^0 + \langle\varphi_n| W |\varphi_n\rangle + O(\lambda^2)} \tag{B-5}$$

The first-order correction to a non-degenerate energy E_n^0 is simply equal to the average value of the perturbation term W in the unperturbed state $|\varphi_n\rangle$.

B-1-b. Eigenvector correction

The projection (B-3) obviously does not exhaust all the information contained in perturbation equation (A-10). We must now project this equation onto all the vectors of the $\{|\varphi_p^i\rangle\}$ basis other than $|\varphi_n\rangle$. We obtain, using (B-1) and (B-2):

$$\langle\varphi_p^i| (H_0 - E_n^0) |1\rangle + \langle\varphi_p^i| (\hat{W} - \varepsilon_1) |\varphi_n\rangle = 0 \qquad (p \neq n) \tag{B-6}$$

(since the eigenvalues E_p^0 other than E_n^0 can be degenerate, we must retain the degeneracy index i here). Since the eigenvectors of H_0 associated with different eigenvalues are orthogonal, the last term, $\varepsilon_1\langle\varphi_p^i|\varphi_n\rangle$, is zero. Furthermore, in the first term, we can let H_0 act on the left on $\langle\varphi_p^i|$. (B-6) then becomes:

$$\left(E_p^0 - E_n^0\right) \langle\varphi_p^i|1\rangle + \langle\varphi_p^i| \hat{W} |\varphi_n\rangle = 0 \tag{B-7}$$

which gives the coefficients of the desired expansion of the vector $|1\rangle$ on all the unperturbed basis states, except $|\varphi_n\rangle$:

$$\langle\varphi_p^i|1\rangle = \frac{1}{E_n^0 - E_p^0} \langle\varphi_p^i| \hat{W} |\varphi_n\rangle \qquad (p \neq n) \tag{B-8}$$

The last coefficient which we lack, $\langle\varphi_n|1\rangle$, is actually zero, according to condition (A-15), which we have not yet used [$|\varphi_n\rangle$, according to (B-2), coincides with $|0\rangle$]:

$$\langle\varphi_n|1\rangle = 0 \tag{B-9}$$

We therefore know the vector $|1\rangle$ since we know its expansion on the $\{|\varphi_p^i\rangle\}$ basis:

$$|1\rangle = \sum_{p\neq n} \sum_i \frac{\langle\varphi_p^i|\hat{W}|\varphi_n\rangle}{E_n^0 - E_p^0} |\varphi_p^i\rangle \tag{B-10}$$

Consequently, to first order in the perturbation $W = \lambda\hat{W}$, the eigenvector $|\psi_n(\lambda)\rangle$ of H corresponding to the unperturbed state $|\varphi_n\rangle$ can be written:

$$|\psi_n(\lambda)\rangle = |\varphi_n\rangle + \sum_{p\neq n} \sum_i \frac{\langle\varphi_p^i|W|\varphi_n\rangle}{E_n^0 - E_p^0} |\varphi_p^i\rangle + O(\lambda^2) \tag{B-11}$$

The first-order correction of the state vector is a linear superposition of all the unperturbed states other than $|\varphi_n\rangle$: the perturbation W is said to produce a "mixing" of the state $|\varphi_n\rangle$ with the other eigenstates of H_0. The contribution of a given state $|\varphi_p^i\rangle$ is

zero if the perturbation W has no matrix element between $|\varphi_n\rangle$ and $|\varphi_p^i\rangle$. In general, the stronger the coupling induced by W between $|\varphi_n\rangle$ and $|\varphi_p^i\rangle$ (characterized by the matrix element $\langle\varphi_p^i|W|\varphi_n\rangle$), and the closer the level E_p^0 to the level E_n^0 under study, the greater the mixing with $|\varphi_p^i\rangle$.

Comment:

We have assumed that the perturbation W is much smaller than the unperturbed Hamiltonian H_0, that is, that the matrix elements of W are much smaller than those of H_0. It appears here that this hypothesis is not sufficient: the first order correction of the state vector is small only if *the non-diagonal matrix elements of W are much smaller than the corresponding unperturbed energy differences.*

B-2. Second-order corrections

The second-order corrections can be extracted from perturbation equation (A-11) by the same method as above, with the addition of condition (A-16).

B-2-a. Energy correction

To calculate ε_2, we project equation (A-11) onto the vector $|\varphi_n\rangle$, using (B-1) and (B-2):

$$\langle\varphi_n|(H_0 - E_n^0)|2\rangle + \langle\varphi_n|(\hat{W} - \varepsilon_1)|1\rangle - \varepsilon_2 \langle\varphi_n|\varphi_n\rangle = 0 \tag{B-12}$$

For the same reason as in § B-1-a, the first term is zero. This is also the case for $\varepsilon_1\langle\varphi_n|1\rangle$, since, according to (B-9), $|1\rangle$ is orthogonal to $|\varphi_n\rangle$. We then get:

$$\varepsilon_2 = \langle\varphi_n|\,\hat{W}\,|1\rangle \tag{B-13}$$

that is, substituting expression (B-10) for the vector $|1\rangle$:

$$\varepsilon_2 = \sum_{p\neq n}\sum_i \frac{\left|\langle\varphi_p^i|\,\hat{W}\,|\varphi_n\rangle\right|^2}{E_n^0 - E_p^0} \tag{B-14}$$

This result enables us to write the energy $E_n(\lambda)$, to second order in the perturbation $W = \lambda\hat{W}$, in the form:

$$E_n(\lambda) = E_n^0 + \langle\varphi_n|\,W\,|\varphi_n\rangle + \sum_{p\neq n}\sum_i \frac{\left|\langle\varphi_p^i|\,W\,|\varphi_n\rangle\right|^2}{E_n^0 - E_p^0} + O(\lambda^3) \tag{B-15}$$

Comment:

The second-order energy correction for the state $|\varphi_n\rangle$ due to the presence of the state $|\varphi_p^i\rangle$ has the sign of $E_n^0 - E_p^0$. We can therefore say that, to second order, the closer the state $|\varphi_p^i\rangle$ to the state $|\varphi_n\rangle$, and the stronger the "coupling" $\left|\langle\varphi_p^i|\,W\,|\varphi_n\rangle\right|$, the more these two levels "repel" each other.

B-2-b. Eigenvector correction

By projecting equation (A-11) onto the set of basis vectors $\left|\varphi_p^i\right\rangle$ different from $|\varphi_n\rangle$, and by using conditions (A-16), we could obtain the expression for the ket $|2\rangle$, and therefore the eigenvector to second order. Such a calculation presents no theoretical difficulties, and we shall not give it here.

Comment:

In (B-4), the first-order energy correction is expressed in terms of the zeroth-order eigenvector. Similarly, in (B-13), the second-order energy correction involves the first-order eigenvector [which explains a certain similarity of formulas (B-10) and (B-14)]. This is a general result: by projecting (A-12) onto $|\varphi_n\rangle$, one makes the first term go to zero, which gives ε_q in terms of the corrections of order $q-1$, $q-2$, ... of the eigenvector. This is why we generally retain one more term in the energy expansion than in that of the eigenvector: for example, the energy is given to second order and the eigenvector to first order.

B-2-c. Upper limit of ε_2

If we limit the energy expansion to first order in λ, we can obtain an approximate idea of the error involved by evaluating the second-order term which is simple to obtain.

Consider expression (B-14) for ε_2. It contains a sum (which is generally infinite) of terms whose numerators are positive or zero. We denote by ΔE the absolute value of the difference between the energy E_n^0 of the level being studied and that of the closest level. For all n, we obviously have:

$$\left|E_n^0 - E_p^0\right| \geqslant \Delta E \tag{B-16}$$

This gives us an upper limit for the absolute value of ε_2:

$$|\varepsilon_2| \leqslant \frac{1}{\Delta E} \sum_{p\neq n} \sum_i \left|\left\langle \varphi_p^i \right| \hat{W} \left|\varphi_n\right\rangle\right|^2 \tag{B-17}$$

which can be written:

$$\begin{aligned} |\varepsilon_2| &\leqslant \frac{1}{\Delta E} \sum_{p\neq n} \sum_i \left\langle \varphi_n \right| \hat{W} \left|\varphi_p^i\right\rangle \left\langle \varphi_p^i \right| \hat{W} \left|\varphi_n\right\rangle \\ &\leqslant \frac{1}{\Delta E} \left\langle \varphi_n \right| \hat{W} \left[\sum_{p\neq n} \sum_i \left|\varphi_p^i\right\rangle \left\langle \varphi_p^i \right|\right] \hat{W} \left|\varphi_n\right\rangle \end{aligned} \tag{B-18}$$

The operator which appears inside the brackets differs from the identity operator only by the projector onto the state $|\varphi_n\rangle$, since the basis of unperturbed states satisfies the closure relation:

$$\left|\varphi_n\right\rangle \left\langle \varphi_n \right| + \sum_{p\neq n} \sum_i \left|\varphi_p^i\right\rangle \left\langle \varphi_p^i \right| = 1 \tag{B-19}$$

Inequality (B-18) therefore becomes simply:

$$|\varepsilon_2| \leqslant \frac{1}{\Delta E} \langle\varphi_n| \hat{W} [1 - |\varphi_n\rangle \langle\varphi_n|] \hat{W} |\varphi_n\rangle$$
$$\leqslant \frac{1}{\Delta E} \left[\langle\varphi_n| \hat{W}^2 |\varphi_n\rangle - \left(\langle\varphi_n| \hat{W} |\varphi_n\rangle \right)^2 \right] \tag{B-20}$$

Multiplying both sides of (B-20) by λ^2 we obtain an upper limit for the second-order term in the expansion of $E_n(\lambda)$, in the form:

$$\left|\lambda^2 \varepsilon_2\right| \leqslant \frac{1}{\Delta E} (\Delta W)^2 \tag{B-21}$$

where ΔW is the root-mean-square deviation of the perturbation W in the unperturbed state $|\varphi_n\rangle$. This indicates the order of magnitude of the error on the energy resulting from taking only the first-order correction into account.

C. Perturbation of a degenerate state

Now assume that the level E_n^0 whose perturbation we want to study is g_n-fold degenerate (where g_n is greater than 1, but finite). We denote by $\mathcal{E}_n^0$ the corresponding eigensubspace of H_0. In this case, the choice:

$$\varepsilon_0 = E_n^0 \tag{C-1}$$

does not suffice to determine the vector $|0\rangle$, since equation (A-9) can theoretically be satisfied by any linear combination of the g_n vectors $\left|\varphi_n^i\right\rangle$ ($i = 1, 2, ..., g_n$). We know only that $|0\rangle$ belongs to the eigensubspace spanned by them.

We shall see that, this time, under the action of the perturbation W, the level E_n^0 generally gives rise to several distinct "sublevels". Their number, f_n, is between 1 and g_n. If f_n is less than g_n, some of these sublevels are degenerate, since the total number of orthogonal eigenvectors of H associated with the f_n sublevels is always equal to g_n. To calculate the eigenvalues and eigenstates of the total Hamiltonian H, we shall limit ourselves, as usually done, to first order in λ for the energies and to zeroth order for the eigenvectors.

To determine ε_1 and $|0\rangle$, we can project equation (A-10) onto the g_n basis vectors $\left|\varphi_n^i\right\rangle$. Since the $\left|\varphi_n^i\right\rangle$ are eigenvectors of H_0 with the eigenvalue $E_n^0 = \varepsilon_0$, we obtain the g_n relations:

$$\left\langle\varphi_n^i\right| \hat{W} |0\rangle = \varepsilon_1 \left\langle\varphi_n^i|0\right\rangle \tag{C-2}$$

We now insert, between the operator $\hat{W}$ and the vector $|0\rangle$, the closure relation for the $\{|\varphi_p^i\rangle\}$ basis:

$$\sum_p \sum_{i'} \left\langle\varphi_n^i\right|\hat{W}\left|\varphi_p^{i'}\right\rangle\left\langle\varphi_p^{i'}|0\right\rangle = \varepsilon_1\left\langle\varphi_n^i|0\right\rangle \tag{C-3}$$

The vector $|0\rangle$, which belongs to the eigensubspace associated with E_n^0, is orthogonal to all the basis vectors $\left|\varphi_p^{i'}\right\rangle$ for which p is different from n. Consequently, on the left-hand

side of (C-3), the sum over the index p reduces to a single term ($p = n$), which gives:

$$\sum_{i'=1}^{g_n} \langle \varphi_n^i | \hat{W} | \varphi_n^{i'} \rangle \langle \varphi_n^{i'} | 0 \rangle = \varepsilon_1 \langle \varphi_n^i | 0 \rangle \tag{C-4}$$

We arrange the g_n^2 numbers $\langle \varphi_n^i | \hat{W} | \varphi_n^{i'} \rangle$ (where n is fixed and i, $i' = 1, 2, \ldots g_n$) in a $g_n \times g_n$ matrix of row index i and column index i'. This square matrix, which we shall denote by $(\hat{W}^{(n)})$ is, so to speak, cut out of the matrix which represents $\hat{W}$ in the $\{|\varphi_p^i\rangle\}$ basis: $(\hat{W}^{(n)})$ is the part which corresponds to $\mathcal{E}_n^0$. Equations (C-4) then show that the column vector of elements $\langle \varphi_n^i | 0 \rangle$ ($i = 1, 2, \ldots g_n$) is an eigenvector of $(\hat{W}^{(n)})$ with the eigenvalue ε_1.

System (C-4) can, moreover, be transformed into a *vector equation inside* $\mathcal{E}_n^0$. All we need to do is define the operator $\hat{W}^{(n)}$, the *restriction of*[6] $\hat{W}$ *to the subspace* $\mathcal{E}_n^0$. $\hat{W}^{(n)}$ acts only in $\mathcal{E}_n^0$, and it is represented in this subspace by the matrix of elements $\langle \varphi_n^i | \hat{W} | \varphi_n^{i'} \rangle$, that is, by $(\hat{W}^{(n)})$. System (C-4) is thus equivalent to the vector equation:

$$\hat{W}^{(n)} |0\rangle = \varepsilon_1 |0\rangle \tag{C-5}$$

[We stress the fact that the operator $\hat{W}^{(n)}$ is different from the operator $\hat{W}$ of which it is the restriction: equation (C-5) is an eigenvalue equation inside $\mathcal{E}_n^0$, and not in all space].

Therefore, *to calculate the eigenvalues (to first order) and the eigenstates (to zeroth order) of the Hamiltonian corresponding to a degenerate unperturbed state* E_n^0, *diagonalize the matrix* $(W^{(n)})$, *which represents the perturbation*[7] W, *inside the eigensubspace* $\mathcal{E}_n^0$ *associated with* E_n^0.

Let us examine more closely the first-order effect of the perturbation W on the degenerate state E_n^0. Let ε_1^j ($j = 1, 2, \ldots f_n^{(1)}$) be the various distinct roots of the characteristic equation of $(\hat{W}^{(n)})$. Since $(\hat{W}^{(n)})$ is Hermitian, its eigenvalues are all real, and the sum of their degrees of degeneracy is equal to g_n ($f_n^{(1)} \leqslant g_n$). Each eigenvalue introduces a different energy correction. Therefore, under the influence of the perturbation $W = \lambda \hat{W}$, the degenerate level splits, to first order, into $f_n^{(1)}$ distinct sublevels, whose energies can be written:

$$E_{n,j}(\lambda) = E_n^0 + \lambda \varepsilon_1^j \qquad j = 1, 2, \ldots f_n^{(1)} \leqslant g_n \tag{C-6}$$

If $f_n^{(1)} = g_n$, we say that, to first order, the perturbation W completely removes the degeneracy of the level E_n^0. If $f_n^{(1)} < g_n$, the degeneracy, to first order, is only partially removed (or not at all if $f_n^{(1)} = 1$).

We shall now choose an eigenvalue ε_1^j of $\hat{W}^{(n)}$. If this eigenvalue is non-degenerate, the corresponding eigenvector $|0\rangle$ is uniquely determined (to within a phase factor) by (C-5) [or by the equivalent system (C-4)]. There then exists a single eigenvalue $E(\lambda)$ of $H(\lambda)$ which is equal to $E_n^0 + \lambda \varepsilon_1^j$, to first order, and this eigenvalue is non-degenerate[8]. On

[6] If P_n is the projector onto the subspace $\mathcal{E}_n^0$, $\hat{W}^{(n)}$ can be written (Complement B_{II}, § 3): $\hat{W}^{(n)} = P_n \hat{W} P_n$.

[7] $(W^{(n)})$ is simply equal to $\lambda(\hat{W}^{(n)})$; this is why its eigenvalues yield directly the corrections $\lambda\varepsilon_1$.

[8] The proof of this point is analogous to the one that shows that a non-degenerate level of H_0 gives rise to a non-degenerate level of $H(\lambda)$ (*cf.* end of § A-2).

the other hand, if the eigenvalue ε_1^j of $\hat{W}^{(n)}$ being considered presents a q-fold degeneracy, (C-5) indicates only that $|0\rangle$ belongs to the corresponding q-dimensional subspace $\mathcal{F}_j^{(1)}$.

This property of ε_1^j can, actually, reflect two very different situations. One could distinguish between them by pursuing the perturbation calculation to higher orders of λ, and seeing whether the remaining degeneracy is removed. These two situations are the following:

(*i*) Suppose that there is only one exact energy $E(\lambda)$ that is equal, to first order, to $E_n^0 + \lambda\varepsilon_1^j$, and that this energy is q-fold degenerate [in Figure 1, for example, the energy $E(\lambda)$ that approaches E_4^0 when $\lambda \longrightarrow 0$ is two-fold degenerate, for any value of λ]. A q-dimensional eigensubspace then corresponds to the eigenvalue $E(\lambda)$, whatever λ, so that the degeneracy of the approximate eigenvalues will never be removed, to any order of λ.

In this case, the zeroth-order eigenvector $|0\rangle$ of $H(\lambda)$ cannot be completely specified, since the only condition imposed on $|0\rangle$ is that of belonging to a subspace which is the limit, when $\lambda \longrightarrow 0$, of the q-dimensional eigensubspace of $H(\lambda)$ corresponding to $E(\lambda)$. This limit is none other than the eigensubspace $\mathcal{F}_j^{(1)}$ of $(\hat{W}^{(n)})$ associated with the eigenvalue ε_1^j chosen.

This first case often arises when H_0 and W possess common symmetry properties, implying an essential degeneracy for $H(\lambda)$. Such a degeneracy then remains to all orders in perturbation theory.

(*ii*) It may also happen that several different energies $E(\lambda)$ are equal, to first order, to $E_n^0 + \lambda\varepsilon_1^j$ (the difference between these energies then appears in a calculation at second or higher orders).

In this case, the subspace $\mathcal{F}_j^{(1)}$ obtained to first order is only the direct sum of the limits, for $\lambda \longrightarrow 0$, of several eigensubspaces associated with these various energies $E(\lambda)$. In other words, all the eigenvectors of $H(\lambda)$ corresponding to these energies certainly approach kets of $\mathcal{F}_j^{(1)}$, but, inversely, a particular ket of $\mathcal{F}_j^{(1)}$ is not necessarily the limit $|0\rangle$ of an eigenket of $H(\lambda)$.

In this situation, going to higher order terms allows one, not only to improve the accuracy of the energies, but also to determine the zeroth-order kets $|0\rangle$. However, in practice, the partial information contained in equation (C-5) is often considered sufficient.

Comments:

(*i*) When we use the perturbation method to treat all the energies[9] of the spectrum of H_0, we must diagonalize the perturbation W inside each of the eigensubspaces $\mathcal{E}_n^0$ corresponding to these energies. It must be understood that this problem is much simpler than the initial problem, which is the complete diagonalization of the Hamiltonian in the entire state space. Perturbation theory enables us to ignore completely the matrix elements of W between vectors belonging to different subspaces $\mathcal{E}_n^0$. Therefore, instead of having to diagonalize a generally infinite matrix, we need only diagonalize, for each of the energies E_n^0 in which we are interested, a matrix of smaller dimensions, generally finite.

(*ii*) The matrix $(\hat{W}^{(n)})$ clearly depends on the $\{|\varphi_n^i\rangle\}$ basis initially chosen in this subspace $\mathcal{E}_n^0$ (although the eigenvalues and eigenkets of $\hat{W}^{(n)}$ obviously do not depend on it). Therefore, before we begin the perturbation calculation, it is advantageous to find a basis that simplifies as much as possible the form of

[9] The perturbation of a non-degenerate state, studied in § B, can be seen as a special case of that of a degenerate state.

$(W^{(n)})$ for this subspace, and, consequently, the search for its eigenvalues and eigenvectors (the simplest situation is obviously the one in which this matrix is obtained directly in a diagonal form). To find such a basis, we often use observables which commute both[10] with H_0 and W. Assume that we have an observable A which commutes with H_0 and W. Since H_0 and A commute, we can choose for the basis vectors $|\varphi_n^i\rangle$ eigenstates common to H_0 and A. Furthermore, since W commutes with A, its matrix elements are zero between eigenvectors of A associated with different eigenvalues. The matrix $(W^{(n)})$ then contains numerous zeros, which facilitates its diagonalization.

(*iii*) Just as for non-degenerate levels (*cf.* comment of § B-1-b), the method described in this section is valid only if the matrix elements of the perturbation W are much smaller than the differences between the energy of the level under study and those of the other levels (this conclusion would have been evident if we had calculated higher-order corrections). However, it is possible to extend this method to the case of a group of unperturbed levels that are very close to each other (but distinct) and very far from all the other levels of the system being considered. This means, of course, that the matrix elements of the perturbation W are of the same order of magnitude as the energy differences inside the group, but are negligible compared to the separation between a level in the group and one outside. We can then approximately determine the influence of the perturbation W by diagonalizing the matrix which represents $H = H_0 + W$ inside this group of levels. It is by relying on an approximation of this type that we can, in certain cases, reduce the study of a physical problem to that of a two-level system, such as those described in Chapter IV (§ C).

References and suggestions for further reading:

For other perturbation methods, see, for example:

Brillouin-Wigner series (an expansion which is simple for all orders but which involves the perturbed energies in the energy denominators): Ziman (2.26), § 3.1.

The resolvent method (an operator method which is well suited for the calculation of higher-order corrections): Messiah (1.17), Chap. XVI, § 111; Roman (2.3), § 4-5-d.

Method of Dalgarno and Lewis (which replaces the summations over the intermediate states by differential equations): Borowitz (1.7). § 14-5; Schiff (1.18), Chap. 8, § 33. Original references: (2.34), (2.35), (2.36).

The W.K.B. method, applicable to quasi-classical situations: Landau and Lifshitz (1.19), Chap. 7; Messiah (1.17), Chap. VI, § 11; Merzbacher (1.16), Chap. VII; Schiff (1.18), § 34; Borowitz (1.7), Chaps. 8 and 9.

The Hartree and Hartree-Fock methods: see Complement E_{XV}; Messiah (1.17), Chap. XVIII, § 11; Slater (11.8), Chaps. 8 and 9 (Hartree) and 17 (Hartree-Fock); Bethe and Jackiw (1.21), Chap. 4. See also references of Complement A_{XIV}.

[10] Recall that this does not imply that H_0 and W commute.

	A_{XI}, B_{XI}, C_{XI} and D_{XI}: illustrations of stationary perturbation theory using simple and important examples.
A_{XI}: A ONE-DIMENSIONAL HARMONIC OSCILLATOR SUBJECTED TO A PERTURBING POTENTIAL IN x, x^2, x^3	Study of a one-dimensional harmonic oscillator perturbed by a potential in x, x^2, x^3. Simple, advised for a first reading. The last example (perturbing potential in x^3) permits the study of the anharmonicity in the vibration of a diatomic molecule (a refinement on the model presented in Complement A_V).
B_{XI}: INTERACTION BETWEEN THE MAGNETIC DIPOLES OF TWO SPIN 1/2 PARTICLES	Can be considered as a worked example, illustrating perturbation theory for non-degenerate as well as degenerate states. Familiarizes the reader with the dipole-dipole interaction between magnetic moments of two spin 1/2 particules. Simple.
C_{XI}: VAN DER WAALS FORCES	Study of the long-distance forces between two neutral atoms using perturbation theory (Van der Waals forces). The accent is placed on the physical interpretation of the results. A little less simple than the two preceding complements: can be reserved for later study.
D_{XI}: THE VOLUME EFFECT: THE INFLUENCE OF THE SPATIAL EXTENSION OF THE NUCLEUS ON THE ATOMIC LEVELS	Study of the influence of the nuclear volume on the energy levels of hydrogen-like atoms. Simple. Can be considered as a sequel of Complement A_{VII}.
E_{XI}: THE VARIATIONAL METHOD	Presentation of another approximation method, the variational method. Important, since the applications of the variational method are very numerous.

F_{XI}, G_{XI}: two important applications of the variational method.

F_{XI}: ENERGY BANDS OF ELECTRONS IN SOLIDS: A SIMPLE MODEL

Introduction, using the strong-bonding approximation, of the concept of an allowed energy band for the electrons of a solid. Essential, because of its numerous applications. Moderetely difficult. The accent is placed on the interpretation of the results. The view point adopted is different from that of Complement O_{III} and somewhat simpler.

G_{XI}: A SIMPLE EXAMPLE OF THE CHEMICAL BOND: THE H_2^+ ION

Studies the phenomenon of the chemical bond for the simplest possible case, that of the (ionized) H_2^+ molecule. Shows how quantum mechanics explains the attractive forces between two atoms whose electronic wave fonctions overlap. Includes a proof of the virial theorem. Essential from the point of view of chemical physics. Moderately difficult.

H_{XI} : EXERCISES

Complement A_{XI}

A one-dimensional harmonic oscillator subjected to a perturbing potential in x, x^2, x^3

1 Perturbation by a linear potential 1131

1-a The exact solution 1132

1-b The perturbation expansion 1133

2 Perturbation by a quadratic potential 1133

3 Perturbation by a potential in x^3 1135

3-a The anharmonic oscillator 1135

3-b The perturbation expansion 1136

3-c Application: the anharmonicity of the vibrations of a diatomic molecule 1137

In order to illustrate the general considerations of Chapter XI by a simple example, we shall use stationary perturbation theory to study the effect of a perturbing potential in x, x^2 or x^3 on the energy levels of a one-dimensional harmonic oscillator (none of these levels is degenerate, *cf.* Chap. V).

The first two cases (a perturbing potential in x and in x^2) are exactly soluble. Consequently, we shall be able to verify in these two examples that the perturbation expansion coincides with the limited expansion of the exact solution with respect to the parameter that characterizes the strength of the perturbation. The last case (a perturbing potential in x^3) is very important in practice for the following reason. Consider a potential $V(x)$ which has a minimum at $x = 0$. To a first approximation, $V(x)$ can be replaced by the first term (in x^2) of its Taylor series expansion, in which case we are considering a harmonic oscillator and, therefore, an exactly soluble problem. The next term of the expansion of $V(x)$, which is proportional to x^3, then constitutes the first correction to this approximation. Calculation of the eflect of the term in x^3, consequently, is necessary whenever we want to study the anharmonicity of the vibrations of a physical system. It permits us, for example, to evaluate the deviations of the vibrational spectrum of diatomic molecules from the predictions of the (purely harmonic) model of Complement A_V.

1. Perturbation by a linear potential

We shall use the notation of Chapter V. Let:

$$H_0 = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2 \tag{1}$$

be the Hamiltonian of a one-dimensional harmonic oscillator of eigenvectors $|\varphi_n\rangle$ and eigenvalues[1]:

$$E_n^0 = \left(n + \frac{1}{2}\right)\hbar\omega \tag{2}$$

with $n = 0, 1, 2...$

We add to this Hamiltonian the perturbation:

$$W = \lambda\hbar\omega\hat{X} \tag{3}$$

where λ is a real dimensionless constant much smaller than 1, and $\hat{X}$ is given by formula (B-1) of Chapter V (since $\hat{X}$ is of the order of 1, $\hbar\omega\hat{X}$ is of the order of H_0 and plays the role of the operator $\hat{W}$ of Chapter XI). The problem consists of finding the eigenstates $|\psi_n\rangle$ and eigenvalues E_n of the Hamiltonian:

$$H = H_0 + W \tag{4}$$

1-a. The exact solution

We have already studied an example of a linear perturbation in X: when the oscillator, assumed to be charged, is placed in a uniform electric field $\mathcal{E}$, we must add to H_0 the electrostatic Hamiltonian:

$$W = -q\mathcal{E}X = -q\mathcal{E}\sqrt{\frac{\hbar}{m\omega}}\hat{X} \tag{5}$$

where q is the charge of the oscillator. The effect of such a term on the stationary states of the harmonic oscillator was studied in detail in Complement F_V. It is therefore possible to use the results of this complement to determine the eigenstates and eigenvalues of the Hamiltonian H given by (4) if we perform the substitution:

$$\lambda\hbar\omega \longleftrightarrow -q\mathcal{E}\sqrt{\frac{\hbar}{m\omega}} \tag{6}$$

Expression (39) of F_V thus yields immediately:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{\lambda^2}{2}\hbar\omega \tag{7}$$

Similarly, we see from (40) of F_V (after having replaced P by its expression in terms of the creation and annihilation operators $a^\dagger$ and a):

$$|\psi_n\rangle = \mathrm{e}^{-\frac{\lambda}{\sqrt{2}}(a^\dagger - a)}\,|\varphi_n\rangle \tag{8}$$

The expansion of the exponential then yields:

$$\begin{aligned}|\psi_n\rangle &= \left[1 - \frac{\lambda}{\sqrt{2}}(a^\dagger - a) + \ldots\right]|\varphi_n\rangle \\ &= |\varphi_n\rangle - \lambda\sqrt{\frac{n+1}{2}}\,|\varphi_{n+1}\rangle + \lambda\sqrt{\frac{n}{2}}\,|\varphi_{n-1}\rangle + \ldots\end{aligned} \tag{9}$$

[1] To specify that we are considering the unperturbed Hamiltonian, as in Chapter XI, we add the index 0 to the eigenvalue of H_0.

1-b. The perturbation expansion

We replace $\hat{X}$ by $\frac{1}{\sqrt{2}}(a^\dagger + a)$ in (3) [*cf.* formula (B-7a) of Chapter V].
We obtain:

$$W = \lambda \frac{\hbar\omega}{\sqrt{2}} \left(a^\dagger + a\right) \tag{10}$$

W then mixes the state $|\varphi_n\rangle$ only with the two states $|\varphi_{n+1}\rangle$ and $|\varphi_{n-1}\rangle$. The only non-zero matrix elements of W are, consequently:

$$\begin{aligned} \langle\varphi_{n+1}|\, W\, |\varphi_n\rangle &= \lambda\sqrt{\frac{n+1}{2}}\,\hbar\omega \\ \langle\varphi_{n-1}|\, W\, |\varphi_n\rangle &= \lambda\sqrt{\frac{n}{2}}\hbar\omega \end{aligned} \tag{11}$$

According to general expression (B-15) of Chapter XI, we have:

$$E_n = E_n^0 + \langle\varphi_n|\, W\, |\varphi_n\rangle + \sum_{n'\neq n} \frac{\left|\langle\varphi_{n'}|\, W\, |\varphi_n\rangle\right|^2}{E_n^0 - E_{n'}^0} + \dots \tag{12}$$

Substituting (11) into (12) and replacing $E_n^0 - E_{n'}^0$ by $(n-n')\hbar\omega$, we immediately obtain:

$$\begin{aligned} E_n &= E_n^0 + 0 - \frac{\lambda^2(n+1)}{2}\hbar\omega + \frac{\lambda^2 n}{2}\hbar\omega + \dots \\ &= \left(n + \frac{1}{2}\right)\hbar\omega - \frac{\lambda^2}{2}\hbar\omega + \dots \end{aligned} \tag{13}$$

This shows that the perturbation expansion of the eigenvalue to second order in λ coincides[2] with the exact solution (7).

Similarly, general formula (B-11) of Chapter XI:

$$|\psi_n\rangle = |\varphi_n\rangle + \sum_{n'\neq n} \frac{\langle\varphi_{n'}|\, W\, |\varphi_n\rangle}{E_n^0 - E_{n'}^0}\, |\varphi_{n'}\rangle + \dots \tag{14}$$

yields here:

$$|\psi_n\rangle = |\varphi_n\rangle - \lambda\sqrt{\frac{n+1}{2}}\; |\varphi_{n+1}\rangle + \lambda\sqrt{\frac{n}{2}}\; |\varphi_{n-1}\rangle + \dots \tag{15}$$

an expression which is identical to expansion (9) of the exact solution.

2. Perturbation by a quadratic potential

We now assume W to have the following form:

$$W = \frac{1}{2}\rho\,\hbar\omega\hat{X}^2 = \frac{1}{2}\rho\, m\omega^2 X^2 \tag{16}$$

[2] It can be shown that all terms of order higher than 2 in the perturbation expansion are zero.

where ρ is a real dimensionless parameter much smaller than 1. H can then be written:

$$H = H_0 + W = \frac{P^2}{2m} + \frac{1}{2}m\omega^2(1+\rho)X^2 \tag{17}$$

In this case, the effect of the perturbation is simply to change the spring constant of the harmonic oscillator. If we set:

$$\omega'^2 = \omega^2(1+\rho) \tag{18}$$

we see that H is still a harmonic oscillator Hamiltonian, whose angular frequency has become ω'.

In this section, we shall confine ourselves to the study of the eigenvalues of H. According to (17) and (18), they can be written simply:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega' = \left(n + \frac{1}{2}\right)\hbar\omega\sqrt{1+\rho} \tag{19}$$

that is, expanding the radical:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega\left[1 + \frac{\rho}{2} - \frac{\rho^2}{8} + \ldots\right] \tag{20}$$

Let us now find result (20) by using stationary perturbation theory. Expression (16) can also be written:

$$\begin{aligned} W &= \frac{1}{4}\rho\,\hbar\omega\left(a^\dagger + a\right)^2 = \frac{1}{4}\rho\,\hbar\omega\left(a^{\dagger 2} + a^2 + aa^\dagger + a^\dagger a\right) \\ &= \frac{1}{4}\rho\,\hbar\omega\left[a^{\dagger 2} + a^2 + 2a^\dagger a + 1\right] \end{aligned} \tag{21}$$

From this, it can be seen that the only non-zero matrix elements of W associated with $|\varphi_n\rangle$ are:

$$\begin{aligned} \langle\varphi_n|\,W\,|\varphi_n\rangle &= \frac{1}{2}\rho\left(n + \frac{1}{2}\right)\hbar\omega \\ \langle\varphi_{n+2}|\,W\,|\varphi_n\rangle &= \frac{1}{4}\rho\big[(n+1)(n+2)\big]^{1/2}\hbar\omega \\ \langle\varphi_{n-2}|\,W\,|\varphi_n\rangle &= \frac{1}{4}\rho\big[n(n-1)\big]^{1/2}\hbar\omega \end{aligned} \tag{22}$$

When we use this result to evaluate the varions terms of (12), we find:

$$\begin{aligned} E_n &= E_n^0 + \frac{\rho}{2}\left(n + \frac{1}{2}\right)\hbar\omega - \frac{\rho^2}{16}(n+1)(n+2)\frac{\hbar\omega}{2} + \frac{\rho^2}{16}n(n-1)\frac{\hbar\omega}{2} + \ldots \\ &= E_n^0 + \left(n + \frac{1}{2}\right)\hbar\omega\frac{\rho}{2} - \left(n + \frac{1}{2}\right)\hbar\omega\frac{\rho^2}{8} + \ldots \\ &= \left(n + \frac{1}{2}\right)\hbar\omega\left[1 + \frac{\rho}{2} - \frac{\rho^2}{8} + \ldots\right] \end{aligned} \tag{23}$$

which indeed coincides with expansion (20).

3. Perturbation by a potential in x^3

We now add to H_0 the perturbation:

$$W = \sigma\hbar\omega\hat{X}^3 \tag{24}$$

where σ is a real dimensionless number much smaller than 1.

3-a. The anharmonic oscillator

Figure 1 represents the variation with respect to x of the total potential $\frac{1}{2}m\omega^2x^2 + W(x)$ in which the particle is moving. The dashed line gives the parabolic potential $\frac{1}{2}m\omega^2x^2$ of the "unperturbed" harmonic oscillalor. We have chosen $\sigma < 0$, so that the total potential (the solid curve in the figure) increases less rapidly for $x > 0$ than for $x < 0$.

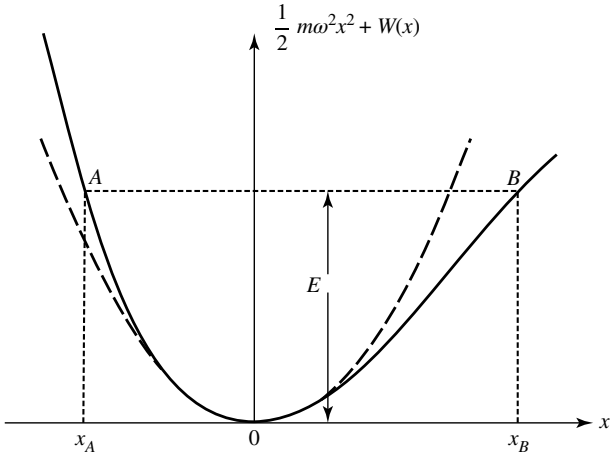


Figure 1: Variation of the potential associated with an anharmonic oscillator with respect to x. We treat the difference between the real potential (solid line) and the harmonic potential (dashed line) of the unperturbed Hamiltonian as a perturbation (x_A and x_B are the limits of the classical motion of energy E).

When the problem is treated in classical mechanics, the particle with total energy E is found to oscillate between two points, x_A and x_B (Fig. 1), which are no longer symmetric with respect to O. This motion, while it remains periodic, is no longer sinusoidal: there appears, in the Fourier expansion of $x(t)$, a whole series of harmonics of the fundamental frequency. This is why such a system is called an "anharmonic oscillator" (its motion is no longer harmonic). Finally, let us point out that the period of the motion is no longer independent of the energy E, as was the case for the harmonic oscillator.

3-b. The perturbation expansion

α. Matrix elements of the perturbation W

We replace $\hat{X}$ by $\frac{1}{\sqrt{2}}(a^\dagger + a)$ in (24). Using relations (B-9) and (B-17) of Chapter V, we obtain, after a simple calculation:

$$W = \frac{\sigma\hbar\omega}{2^{3/2}}\left[a^{\dagger 3} + a^3 + 3Na^\dagger + 3(N+1)a\right] \tag{25}$$

where $N = a^\dagger a$ was defined in Chapter V [formula (B-13)].

From this can immediately be deduced the only non-zero matrix elements of W associated with $|\varphi_n\rangle$:

$$\begin{aligned}
\langle\varphi_{n+3}|\, W\, |\varphi_n\rangle &= \sigma\left[\frac{(n+3)(n+2)(n+1)}{8}\right]^{\frac{1}{2}}\hbar\omega \\
\langle\varphi_{n-3}|\, W\, |\varphi_n\rangle &= \sigma\left[\frac{n(n-1)(n-2)}{8}\right]^{\frac{1}{2}}\hbar\omega \\
\langle\varphi_{n+1}|\, W\, |\varphi_n\rangle &= 3\sigma\left(\frac{n+1}{2}\right)^{\frac{3}{2}}\hbar\omega \\
\langle\varphi_{n-1}|\, W\, |\varphi_n\rangle &= 3\sigma\left(\frac{n}{2}\right)^{\frac{3}{2}}\hbar\omega
\end{aligned} \tag{26}$$

β. Calculation of the energies

We substitute results (26) into the perturbation expansion of E_n, see relation (12). Since the diagonal element of W is zero, there is no first-order correction. The four matrix elements (26) enter, however, into the second-order correction. A simple calculation thus yields:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{15}{4}\sigma^2\left(n + \frac{1}{2}\right)^2\hbar\omega - \frac{7}{16}\sigma^2\hbar\omega + \ldots \tag{27}$$

The effect of W is therefore to lower the levels (whatever the sign of σ). The larger n, the greater the shift (Fig. 2). The difference between two adjacent levels is equal to:

$$E_n - E_{n-1} = \hbar\omega\left[1 - \frac{15}{2}\sigma^2 n\right] \tag{28}$$

It is no longer independent of n, as it was for the harmonic oscillator. The energy states are no longer equidistant and move closer together as n increases.

γ. Calculation of the eigenstates

Substituting relations (26) into expansion (14), we easily obtain:

$$|\psi_n\rangle = |\varphi_n\rangle - 3\sigma\left(\frac{n+1}{2}\right)^{\frac{3}{2}}|\varphi_{n+1}\rangle + 3\sigma\left(\frac{n}{2}\right)^{\frac{3}{2}}|\varphi_{n-1}\rangle - \frac{\sigma}{3}\left[\frac{(n+3)(n+2)(n+1)}{8}\right]^{\frac{1}{2}}|\varphi_{n+3}\rangle + \frac{\sigma}{3}\left[\frac{n(n-1)(n-2)}{8}\right]^{\frac{1}{2}}|\varphi_{n-3}\rangle + \ldots \tag{29}$$

Under the effect of the perturbation W, the state $|\varphi_n\rangle$ is therefore mixed with the states $|\varphi_{n+1}\rangle$, $|\varphi_{n-1}\rangle$, $|\varphi_{n+3}\rangle$ and $|\varphi_{n-3}\rangle$.

3-c. Application: the anharmonicity of the vibrations of a diatomic molecule

In Complement A_V, we showed that a heteropolar diatomic molecule could absorb or emit electromagnetic waves whose frequency coincides with the vibrational frequency of the two nuclei of the molecule about their equilibrium position. If we denote by x the displacement $r - r_e$ of the two nuclei from their equilibrium position r_e, the electric dipole moment of the molecule can be written:

$$D(x) = d_0 + d_1 x + \ldots \tag{30}$$

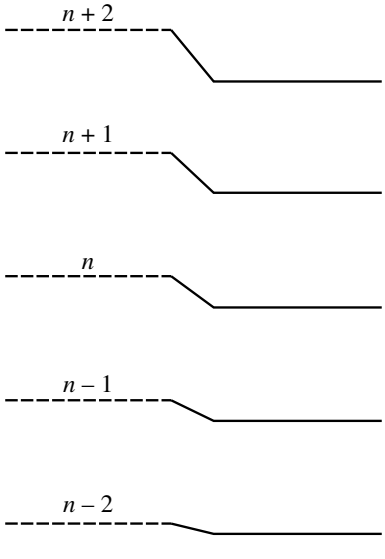


Figure 2: Energy levels of H_0 (dashed lines) and of H (solid lines). Under the effect of the perturbation W, each level of H_0 is lowered, and the higher n, the greater the shift.

The vibrational frequencies of this dipole are therefore the Bohr frequencies which can appear in the expression for $\langle X \rangle(t)$. For a harmonic oscillator, the selection rules satisfied by X are such that only one Bohr frequency can be involved, the frequency $\omega/2\pi$ (*cf.* Complement A_V).

When we take the perturbation W into account, the states $|\varphi_n\rangle$ of the oscillator are "mixed" [*cf.* expression (29)], and X can connect states $|\psi_n\rangle$ and $|\psi_{n'}\rangle$ for which $n' - n \neq \pm 1$: new frequencies can thus be absorbed or emitted by the molecule.

To analyze this phenomenon more closely, we shall assume that the molecule is initially in its vibrational ground state $|\psi_0\rangle$ (this is practically always the case at ordinary temperatures T since, in general, $\hbar\omega \gg kT$). By using expression (29), we can calculate, to first order[3] in σ, the matrix elements of $\hat{X}$ between the state $|\psi_0\rangle$ and an arbitrary state $|\psi_n\rangle$. A simple calculation thus yields the following matrix elements (all the others are zero to first order in σ):

$$\langle \psi_1 | \hat{X} | \psi_0 \rangle = \frac{1}{\sqrt{2}} \tag{31a}$$

$$\langle \psi_2 | \hat{X} | \psi_0 \rangle = \frac{1}{\sqrt{2}}\sigma \tag{31b}$$

$$\langle \psi_0 | \hat{X} | \psi_0 \rangle = -\frac{3}{2}\sigma \tag{31c}$$

From this, we can find the transition frequencies observable in the absorption spectrum of the ground state. We naturally find the frequency:

$$\nu_1 = \frac{E_1 - E_0}{h} \tag{32a}$$

which appears with the greatest intensity since, according to (31a), $\langle \psi_1 | \hat{X} | \psi_0 \rangle$ is of zeroth-order in σ. Then, with a much smaller intensity [*cf.* formula (31b)], we find the frequency:

$$\nu_2 = \frac{E_2 - E_0}{h} \tag{32b}$$

which is often called the second harmonic (although it is not rigorously equal to twice ν_1).

Comment:

Result (31c) means that the average value of $\hat{X}$ is not zero in the ground state. This can easily be understood from Figure 1, since the oscillatory motion is no longer symmetric about O. If σ is negative (the case in Figure 1), the oscillator spends more time in the $x > 0$ region than in the $x < 0$ region, and the average value of X must be positive. We thus understand the sign appearing in (31c).

The preceding calculation reveals only one new line in the absorption spectrum. Actually, the perturbation calculation could be pursued to higher orders in σ, taking into account higher order terms in expansion (30) of the dipole moment $D(x)$, as well as

[3] It would not be correct to keep terms of order higher than 1 in the calculation, since expansion (29) is valid only to first order in σ.

terms in x^4, x^5... in the expansion of the potential in the neighborhood of $x = 0$. All the frequencies:

$$\nu_n = \frac{E_n - E_0}{h} \tag{33}$$

with $n = 3, 4, 5...$ would then be present in the absorption spectrum of the molecule (with intensities decreasing very rapidly when n increases). This would finally give, for this spectrum, the form shown in Figure 3. This is what is actually observed.

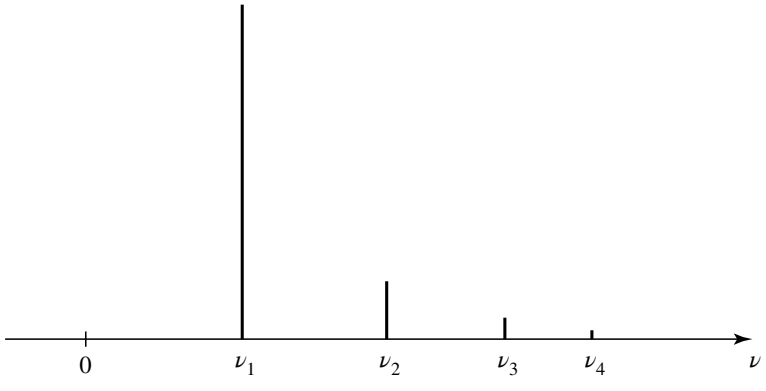


Figure 3: Form of the vibrational spectrum of a heteropolar diatomic molecule. A series of "harmonic" frequencies ν_2, $\nu_3 \ldots \nu_n \ldots$ appear in addition to the fundamental frequency ν_1. This results from the anharmonicity of the potential, as well as higher order terms in the power series expansion in x (the distance between the two atoms) of the molecular dipole moment $D(x)$. Note that the corresponding lines are not quite equidistant and that their intensity decreases rapidly when n increases.

Note that the various spectral lines of Figure 3 are not equidistant since, according to formula (28):

$$\nu_1 - 0 = \frac{E_1 - E_0}{h} = \frac{\omega}{2\pi}\left(1 - \frac{15}{2}\sigma^2\right) \tag{34}$$

$$\nu_2 - \nu_1 = \frac{E_2 - E_1}{h} = \frac{\omega}{2\pi}\left(1 - 15\sigma^2\right) \tag{35}$$

$$\nu_3 - \nu_2 = \frac{E_3 - E_2}{h} = \frac{\omega}{2\pi}\left(1 - \frac{45}{2}\sigma^2\right) \tag{36}$$

which gives the relation:

$$(\nu_2 - \nu_1) - \nu_1 = (\nu_3 - \nu_2) - (\nu_2 - \nu_1) = -\frac{15\omega}{4\pi}\sigma^2 \tag{37}$$

Thus we see that the study of the precise positions of the lines of the absorption spectrum makes it possible to find the parameter σ.

Comments:

(*i*) The constant ξ appearing in (52) of Complement F_{VII} can be evaluated by using formula (27) of the present complement. Comparing these two expressions and replacing n by v in (27), we obtain:

$$\xi = -\frac{15}{4}\sigma^2 \tag{38}$$

Now, the perturbing potential in F_{VII} is equal to $-gx^3$, while here we have chosen it equal to $\sigma\hbar\omega\hat{x}^3$, that is, equal to:

$$\sigma\left(\frac{m^3\omega^5}{\hbar}\right)^{\frac{1}{2}} x^3 \tag{39}$$

We therefore have:

$$\sigma = -g\left(\frac{\hbar}{m^3\omega^5}\right)^{\frac{1}{2}} \tag{40}$$

which, substituted into (38), finally yields:

$$\xi = -\frac{15}{4}\,\frac{g^2\hbar}{m^3\omega^5} \tag{41}$$

(*ii*) In the expansion of the potential in the neighborhood of $x = 0$, the term in x^4 is much smaller than the term in x^3 but it corrects the energies to first order, while the term in x^3 enters only in second order (*cf.* § 3-b-β above). It is therefore necessary to evaluate these two corrections simultaneously (they may be comparable) when the spectrum of Figure 3 is studied more precisely.

References and suggestions for further reading:

Anharmonicity of the vibrations of a diatomic molecule: Herzberg (12.4), vol. I, Chap. III, § 2.

Complement B_{XI}

Interaction between the magnetic dipoles of two spin 1/2 particles

1 The interaction Hamiltonian *W* 1141
1-a The form of the Hamiltonian *W*. Physical interpretation . . . 1141
1-b An equivalent expression for *W* 1142
1-c Selection rules 1143
2 Effects of the dipole-dipole interaction on the Zeeman sublevels of two fixed particles 1144
2-a Case where the two particles have different magnetic moments 1144
2-b Case where the two particles have equal magnetic moments . 1147
2-c Example: the magnetic resonance spectrum of gypsum 1149
3 Effects of the interaction in a bound state 1149

In this complement, we intend to use stationary perturbation theory to study the energy levels of a system of two spin 1/2 particles placed in a static field $\mathbf{B}_0$ and coupled by a magnetic dipole-dipole interaction.

Such systems do exist. For example, in a gypsum monocrystal ($CaSO_4$, $2H_20$), the two protons of each crystallization water molecule occupy fixed positions, and the dipole-dipole interaction between them leads to a fine structure in the nuclear magnetic resonance spectrum.

In the hydrogen atom, there also exists a dipole-dipole interaction between the electron spin and the proton spin. In this case, however, the two particles are moving relative to each other, and we shall see that the effect of the dipole-dipole interaction vanishes due to the symmetry of the $1s$ ground state. The hyperfine structure observed in this state is thus due to other interactions (contact interaction; *cf.* Chap. XII, §§ B-2 and D-2 and Complement A_{XII}).

1. The interaction Hamiltonian *W*

1-a. The form of the Hamiltonian *W*. Physical interpretation

Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be the spins of particles (1) and (2), and $\mathbf{M}_1$ and $\mathbf{M}_2$ their corresponding magnetic moments:

$$\begin{aligned} \mathbf{M}_1 &= \gamma_1 \, \mathbf{S}_1 \\ \mathbf{M}_2 &= \gamma_2 \, \mathbf{S}_2 \end{aligned} \tag{1}$$

[where γ_1 and γ_2 are the gyromagnetic ratios of (1) and (2)].

We call W the interaction of the magnetic moment $\mathbf{M}_2$ with the field created by $\mathbf{M}_1$ at (2). If $\mathbf{n}$ denotes the unit vector of the line joining the two particles and r, the distance between them (Fig. 1), W can be written:

$$W = \frac{\mu_0}{4\pi} \gamma_1 \gamma_2 \frac{1}{r^3} \left[\mathbf{S}_1 \cdot \mathbf{S}_2 - 3 \left(\mathbf{S}_1 \cdot \mathbf{n} \right) \left(\mathbf{S}_2 \cdot \mathbf{n} \right) \right] \tag{2}$$

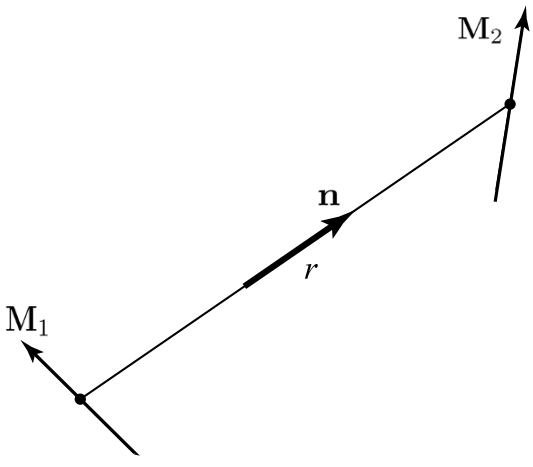


Figure 1: Relative disposition of the magnetic moments $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *of particles (1) and (2) (r is the distance between the two particles, and* $\mathbf{n}$ *is the unit vector of the straight line between them).*

The calculation which enables us to obtain expression (2) is in every way analogous to the one that will be presented in Complement C_{XI} and which leads to the expression for the interaction between two electric dipoles.

1-b. An equivalent expression for *W*

Let θ and φ be the polar angles of $\mathbf{n}$. If we set:

$$\xi(r) = -\frac{\mu_0}{4\pi}\,\frac{\gamma_1\gamma_2}{r^3} \tag{3}$$

we get:

$$\begin{aligned} W = \xi(r)\Big\{ & 3\left[S_{1z}\cos\theta + \sin\theta\left(S_{1x}\cos\varphi + S_{1y}\sin\varphi\right)\right] \\ & \times \left[S_{2z}\cos\theta + \sin\theta\left(S_{2x}\cos\varphi + S_{2y}\sin\varphi\right)\right] - \mathbf{S}_1\cdot\mathbf{S}_2\Big\} \\ = \xi(r)\Big\{ & 3\left[S_{1z}\cos\theta + \frac{1}{2}\sin\theta\left(S_{1+}\mathrm{e}^{-i\varphi} + S_{1-}\mathrm{e}^{i\varphi}\right)\right] \\ & \times \left[S_{2z}\cos\theta + \frac{1}{2}\sin\theta\left(S_{2+}\mathrm{e}^{-i\varphi} + S_{2-}\mathrm{e}^{i\varphi}\right)\right] - \mathbf{S}_1\cdot\mathbf{S}_2\Big\} \end{aligned} \tag{4}$$

that is:

$$W = \xi(r)\left[T_0 + T_0' + T_1 + T_{-1} + T_2 + T_{-2}\right] \tag{5}$$

where:

$$\begin{cases} T_0 &= \left(3\cos^2\theta - 1\right) S_{1z}\, S_{2z} \\ T_0' &= -\dfrac{1}{4}\left(3\cos^2\theta - 1\right)\left(S_{1+}\, S_{2-} + S_{1-}\, S_{2+}\right) \\ T_1 &= \dfrac{3}{2}\sin\theta\cos\theta\, \mathrm{e}^{-i\varphi}\left(S_{1z}\, S_{2+} + S_{1+}\, S_{2z}\right) \\ T_{-1} &= \dfrac{3}{2}\sin\theta\cos\theta\, \mathrm{e}^{i\varphi}\left(S_{1z}\, S_{2-} + S_{1-}\, S_{2z}\right) \\ T_2 &= \dfrac{3}{4}\sin^2\theta\, \mathrm{e}^{-2i\varphi} S_{1+}\, S_{2+} \\ T_{-2} &= \dfrac{3}{4}\sin^2\theta\, \mathrm{e}^{2i\varphi} S_{1-}\, S_{2-} \end{cases} \tag{6}$$

Each of the terms T_q (or T_q') appearing in (5) is, according to (6), the product of a function of θ and φ proportional to the second-order spherical harmonic Y_2^q and an operator acting only on the spin degrees of freedom [the space and spin operators appearing in (6) are second-rank tensors; W, for this reason, is often called the "tensor interaction"].

1-c. Selection rules

r, θ and φ are the spherical coordinates of the relative particle associated with the system of two particles (1) and (2). The operator W acts only on these variables and on the spin degrees of freedom of the two particles. Let $\{|\varphi_{n,l,m}\rangle\}$ be a standard basis in the state space $\mathcal{E}_{\mathbf{r}}$ of the relative particle, and $\{|\varepsilon_1, \varepsilon_2\rangle\}$, the basis of eigenvectors common to S_{1z} and S_{2z} in the spin state space ($\varepsilon_1 = \pm, \varepsilon_2 = \pm$). The state space in which W acts is spanned by the $\{|\varphi_{n,l,m}\rangle \otimes |\varepsilon_1, \varepsilon_2\rangle\}$ basis, in which it is very easy, using expressions (5) and (6), to find the selection rules satisfied by the matrix elements of W.

α. Spin degrees of freedom

- T_0 changes neither ε_1 nor ε_2.
- T_0' "flips" both spins:

$$|+,-\rangle \longrightarrow |-,+\rangle \quad \text{and} \quad |-,+\rangle \longrightarrow |+,-\rangle$$

- T_1 flips one of the two spins up:

$$|-,\varepsilon_2\rangle \longrightarrow |+,\varepsilon_2\rangle \quad \text{or} \quad |\varepsilon_1,-\rangle \longrightarrow |\varepsilon_1,+\rangle$$

- Similarly, T_{-1} flips one of the two spins down:

$$|+,\varepsilon_2\rangle \longrightarrow |-,\varepsilon_2\rangle \quad \text{or} \quad |\varepsilon_1,+\rangle \longrightarrow |\varepsilon_1,-\rangle$$

- Finally, T_2 and T_{-2} flip both spins up and down, respectively:

$$|-,-\rangle \longrightarrow |+,+\rangle \quad \text{and} \quad |+,+\rangle \longrightarrow |-,-\rangle$$

β. Orbital degrees of freedom

When we calculate the matrix element of $\xi(r)T_q$ between the state $|\varphi_{n,l,m}\rangle$ and the state $|\varphi_{n',l',m'}\rangle$, the following angular integral appears:

$$\int Y_{l'}^{m'*}(\theta,\varphi)\; Y_2^q(\theta,\varphi)\; Y_l^m(\theta,\varphi)\,\mathrm{d}\Omega \tag{7}$$

which, according to the results of Complement C_X, is different from zero only for:

$$l' = l,\; l-2,\; l+2 \tag{8a}$$
$$m' = m+q \tag{8b}$$

Note that the case $l = l' = 0$, although not in contradiction with (8), is excluded because we must always be able to form a triangle with l, l' and 2, which is impossible when $l = l' = 0$. We must have then:

$$l, l' \geq 1 \tag{8c}$$

2. Effects of the dipole-dipole interaction on the Zeeman sublevels of two fixed particles

In this section, we shall assume the two particles to be fixed in space. We shall therefore quantize only the spin degrees of freedom, considering the quantities r, θ and φ as given parameters.

The two particles are placed in a static field $\mathbf{B}_0$ parallel to Oz. The Zeeman Hamiltonian H_0, describing the interaction of the two spin magnetic moments with $\mathbf{B}_0$, can then be written:

$$H_0 = \omega_1 S_{1z} + \omega_2 S_{2z} \tag{9}$$

with:

$$\omega_1 = -\gamma_1 B_0$$
$$\omega_2 = -\gamma_2 B_0 \tag{10}$$

In the presence of the dipole-dipole interaction W, the total Hamiltonian H of the system becomes:

$$H = H_0 + W \tag{11}$$

We shall assume the field B_0 to be large enough and treat W as a perturbation of H_0.

2-a. Case where the two particles have different magnetic moments

α. Zeeman levels and the magnetic resonance spectrum in the absence of interaction

According to (9), we have:

$$H_0|\varepsilon_1,\varepsilon_2\rangle = \frac{\hbar}{2}\left(\varepsilon_1\omega_1 + \varepsilon_2\omega_2\right)|\varepsilon_1,\varepsilon_2\rangle \tag{12}$$

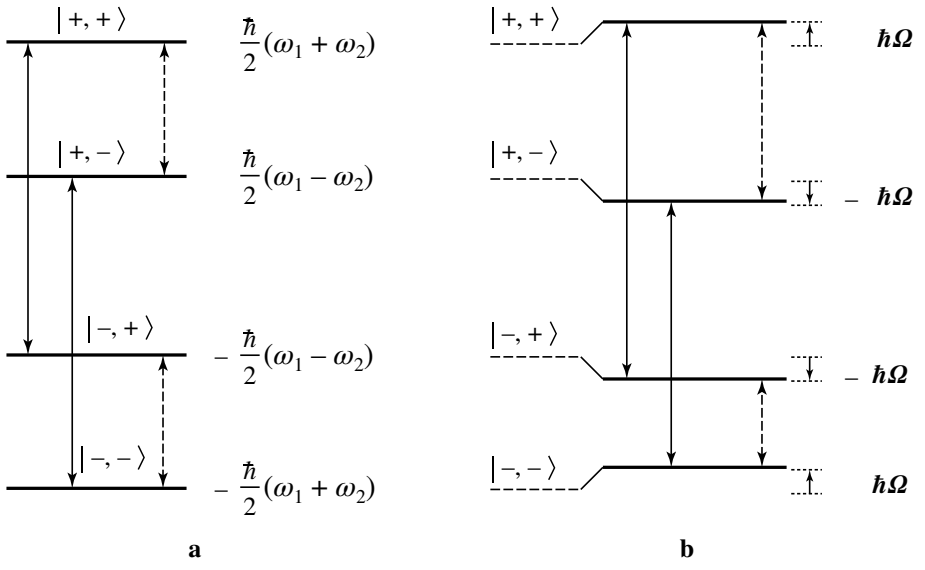


Figure 2: Energy levels of two spin 1/2 particles, placed in a static field $\mathbf{B}_0$ *parallel to* Oz*. The two Larmor angular frequencies,* $\omega_1 = -\gamma_1 B_0$ *and* $\omega_2 = -\gamma_2 B_0$*, are assumed to be different.*
For figure a, the energy levels are calculated without taking account of the dipole-dipole interaction W *between the two spins.*
For figure b, we take this interaction into account. The levels undergo a shift whose approximate value, to first order in W*, is indicated on the right-hand side of the figure. The solid-line arrows join the levels between which* S_{1x} *has a non-zero matrix element, and the shorter dashed-line arrows those for which* S_{2x} *does.*

Figure 2a represents the energy levels of the two-spin system in the absence of the dipole-dipole interaction (we have assumed $\omega_1 > \omega_2 > 0$). Since $\omega_1 \neq \omega_2$, these levels are all non-degenerate.

If we apply a radio-frequency field $\mathbf{B}_1 \cos \omega t$ parallel to Ox, we obtain a series of magnetic resonance lines. The frequencies of these resonances correspond to the various Bohr frequencies which can appear in the evolution of $\langle \gamma_1 S_{1x} + \gamma_2 S_{2x} \rangle$ (the radio-frequency field interacts with the component along Ox of the total magnetic moment). The solid-line (dashed-line) arrows of Figure 2a join levels between which $S_{1x}(S_{2x})$ has a non-zero matrix element. Thus we see that there are two distinct Bohr angular frequencies, equal to ω_1 and ω_2 (Fig. 3a), which correspond simply to the resonances of the individual spins, (1) and (2).

β. Modifications created by the interaction

Since all the levels of Figure 2a are non-degenerate, the effect of W can be obtained to first order by calculating the diagonal elements of W, $\langle \varepsilon_1, \varepsilon_2 | W | \varepsilon_1, \varepsilon_2 \rangle$. It is clear from expressions (5) and (6) that only the term T_0 makes a non-zero contribution to this

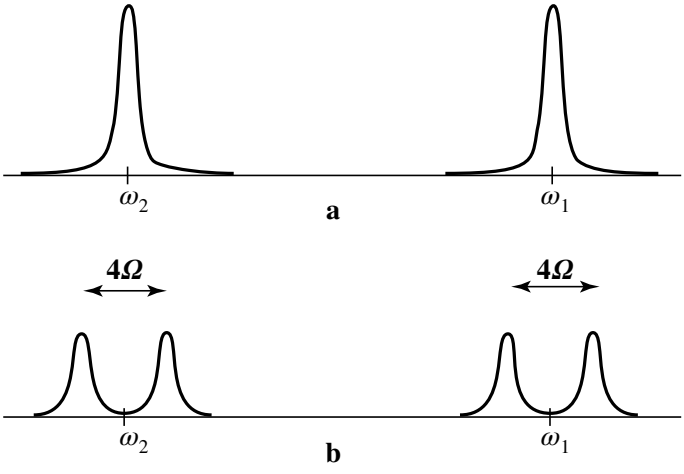


Figure 3: The Bohr frequencies appearing in the evolution of $\langle S_{1x} \rangle$ and $\langle S_{2x} \rangle$ give the positions of the magnetic resonance lines that can be observed for the two-spin system (the transitions corresponding to the arrows of Figure 2). In the absence of a dipole-dipole interaction, two resonances are obtained, each one corresponding to one of the two spins (fig. a). The dipole-dipole interaction is expressed by a splitting of each of the two preceding lines (fig. b).

matrix element, which is then equal to:

$$\langle \varepsilon_1, \varepsilon_2 | W | \varepsilon_1, \varepsilon_2 \rangle = \xi(r) \left(3\cos^2 \theta - 1 \right) \frac{\varepsilon_1 \varepsilon_2 \hbar^2}{4} = \varepsilon_1 \varepsilon_2 \, \hbar\Omega \tag{13}$$

with:

$$\Omega = \frac{\hbar}{4} \xi(r) \left(3\cos^2 \theta - 1 \right) = \frac{-\hbar \mu_0}{16\pi} \frac{\gamma_1 \gamma_2}{r^3} \left(3\cos^2 \theta - 1 \right) \tag{14}$$

Since W is much smaller than H_0, we have:

$$\Omega \ll \omega_1 - \omega_2 \tag{15}$$

From this we can immediately deduce the level shifts to first order in W: $\hbar\Omega$ for $|+,+\rangle$ and $|-,-\rangle$, and $-\hbar\Omega$ for $|+,-\rangle$ and for $|-,+\rangle$ (Fig. 2b).

What now happens to the magnetic resonance spectrum of Figure 3a? If we are concerned only with lines whose intensities are of zeroth order in W (that is, those that approach the lines of Figure 2a when W approaches zero), then to calculate the Bohr frequencies appearing in $\langle S_{1x} \rangle$ and $\langle S_{2x} \rangle$ we simply use the zeroth-order expressions for the eigenvectors[1]. It is then the same transitions which are involved (compare the arrows of Figures 2a and 2b). We see, however, that the two lines which correspond to the frequency ω_1 in the absence of coupling (solid-line arrows) now have different frequencies:

[1] If we used higher-order expressions for the eigenvectors, we would see other lines of lower intensity appear (they disappear when $W \longrightarrow 0$).

$\omega_1 + 2\Omega$ and $\omega_1 - 2\Omega$. Similarly, the two lines corresponding to ω_2 (dashed-line arrows) now have frequencies of $\omega_2 + 2\Omega$ and $\omega_2 - 2\Omega$. The magnetic resonance spectrum is therefore now composed of two "doublets" centered at ω_1 and ω_2, the interval between the two components of each doublet being equal to 4Ω (Fig. 3b).

Thus, the dipole-dipole interaction leads to a fine structure in the magnetic resonance spectrum, for which we can give a simple physical interpretation. The magnetic moment $\mathbf{M}_1$ associated with $\mathbf{S}_1$ creates a "local field" $\mathbf{b}$ at particle (2). Since we assume $\mathbf{B}_0$ to be very large, $\mathbf{S}_1$ precesses very rapidly about Oz, so we can consider only the S_{1z} component (the local field created by the other components oscillates too rapidly to have a significant effect). The local field $\mathbf{b}$ therefore has a different direction depending on whether the spin is in the state $|+\rangle$ or $|-\rangle$, that is, depending on whether it points up or down. It follows that the total field "seen" by particle (2), which is the sum of $\mathbf{B}_0$ and $\mathbf{b}$, can take on two possible values[2]. This explains the appearance of two resonance frequencies for the spin (2). The same argument would obviously enable us to understand the origin of the doublet centered at ω_1.

2-b. Case where the two particles have equal magnetic moments

α. Zeeman levels and the magnetic resonance spectrum in the absence of the interaction

Formula (12) remains valid if we choose ω_1 and ω_2 to be equal. We shall therefore set:

$$\omega_1 = \omega_2 = \omega = -\gamma B_0 \tag{16}$$

The energy levels are shown in Figure 4a. The upper level, $|+,+\rangle$, and the lower level, $|-,-\rangle$, of energies $\hbar\omega$ and $-\hbar\omega$, are non-degenerate. On the other hand, the intermediate level, of energy 0, is two-fold degenerate: to it correspond the two eigenstates $|+,-\rangle$ and $|-,+\rangle$.

The frequencies of the magnetic resonance lines can be obtained by finding the Bohr frequencies involved in the evolution of $\langle S_{1x} + S_{2x}\rangle$ (the total magnetic moment is now proportional to the total spin $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$). We easily obtain the four transitions represented by the arrows in Figure 4a, which correspond to a single angular frequency ω. This finally yields the spectrum of Figure 5a.

β. Modifications created by the interaction

The shifts of the non-degenerate levels $|+,+\rangle$ and $|-,-\rangle$ can be obtained as they were before, and are both equal to $\hbar\Omega$ [we must replace, however, γ_1 and γ_2 by γ in expression (14) for Ω].

Since the intermediate level is two-fold degenerate, the effect of W on this level can now be obtained by diagonalizing the matrix that represents the restriction of W to the subspace $\{|+,-\rangle, |-,+\rangle\}$. The calculation of the diagonal elements is performed as above and yields:

$$\langle +,-|W|+,-\rangle = \langle -,+|W|-,+\rangle = -\hbar\Omega \tag{17}$$

[2] Actually, since $|\mathbf{B}_0| \gg |\mathbf{b}|$, it is only the component of $\mathbf{b}$ along $\mathbf{B}_0$ which is involved.

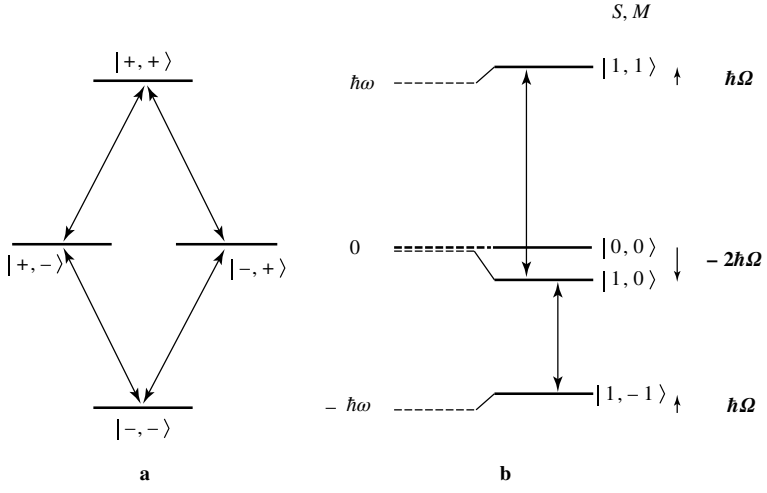


Figure 4: The two spin 1/2 particles are assumed to have the same magnetic moment and, conseqnently, the same Larmor angular frequency $\omega = -\gamma B_0$.
In the absence of a dipole-dipole interaction, we obtain three levels, one of which is two-fold degenerate (fig. a). Under the effect of the dipole-dipole interaction (fig. b), these levels undergo shifts whose approximate values (to first order in W) are indicated on the right-hand side of the figure. To zeroth-order in W, the stationary states are the eigenstates $|S, M\rangle$ of the total spin. The arrows join the levels between which $S_{1x} + S_{2x}$ has a non-zero matrix element.

As for the non-diagonal element $\langle +, -|W|-, +\rangle$, we easily see from expressions (5) and (6) that only the term T_0' contributes to it:

$$\begin{aligned}\langle +, -|W|-, +\rangle &= -\frac{\xi(r)}{4}\left(3\cos^2\theta - 1\right) \times \langle +, -|(S_{1+}S_{2-} + S_{1-}S_{2+})|-, +\rangle \\ &= -\xi(r)\frac{\hbar^2}{4}\left(3\cos^2\theta - 1\right) = -\hbar\Omega \end{aligned} \tag{18}$$

We are then led to the diagonalization of the matrix:

$$-\hbar\Omega \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{19}$$

whose eigenvalues are $-2\hbar\Omega$ and 0; they are respectively associated with the eigenvectors $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|+, -\rangle + |-, +\rangle)$ and $|\psi_2\rangle = \frac{1}{\sqrt{2}}(|+, -\rangle - |-, +\rangle)$.

Figure 4b represents the energy levels of the system of two coupled spins. The energies, to first order in W, are given by the eigenstates to zeroth order.

Note that these eigenstates are none other than the eigenstates $|S, M\rangle$ common to $\mathbf{S}^2$ and S_z, where $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$ is the total spin. Since the operator S_x commutes with $\mathbf{S}^2$, it can only couple the triplet states, that is, $|1, 0\rangle$ to $|1, 1\rangle$ and $|1, 0\rangle$ to $|1, -1\rangle$. This

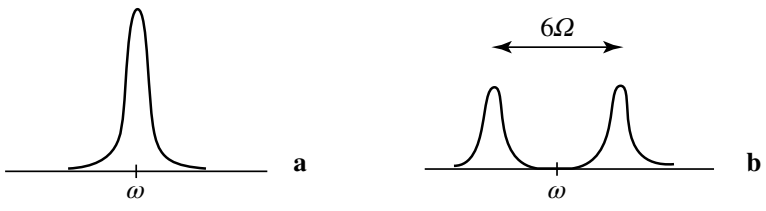


Figure 5: Shape of the magnetic resonance spectrum which can be observed for a system of two spin 1/2 particles, with the same gyromagnetic ratio, placed in a static field B_0. In the absence of a dipole-dipole interaction, we observe a single resonance (fig. a). In the presence of a dipole-dipole interaction (fig. b), the preceding line splits. The separation 6Ω between the two components of the doublet is proportional to $3\cos^2\theta - 1$, where θ is the angle between the static field B_0 and the straight line joining the two particles.

gives the two transitions represented by the arrows in Figure 4b, and to which correspond the Bohr frequencies $\omega + 3\Omega$ and $\omega - 3\Omega$. The magnetic resonance spectrum is therefore composed of a doublet centered at ω, the separation between the two components of the doublet being equal to 6Ω (Fig. 5b).

2-c. Example: the magnetic resonance spectrum of gypsum

The case studied in § 2-b above corresponds to that of two protons of a crystallization water molecule in a gypsum monocrystal ($CaSO_4$, $2H_2O$). These two protons have identical magnetic moments and can be considered to occupy fixed positions in the crystal. Moreover, they are much closer to each other than to other protons (belonging to other water molecules). Since the dipole-dipole interaction decreases very quickly when the distance increases ($1/r^3$ law), we can neglect interactions between protons belonging to other water molecules.

The magnetic resonance spectrum is indeed observed to contain a doublet[3] whose separation depends on the angle θ between the field $\mathbf{B}_0$ and the straight line joining the two protons. If we rotate the crystal with respect to the field $\mathbf{B}_0$, this angle θ varies, and the separation between the two components of the doublet changes. Thus, by studying the variations of this separation, we can determine the positions of the water molecules relative to the crystal axes.

When the sample under study is not a monocrystal, but rather a powder composed of small, randomly oriented monocrystals, θ takes on all possible values. We then observe a wide band, due to the superposition of doublets having different separations.

3. Effects of the interaction in a bound state

We shall now assume that the two particles, (1) and (2), are not fixed, but can move with respect to each other.

[3] Actually, in a gypsum monocrystal, there are two different orientations for the water molecules, and, consequently, two doublets corresponding to the two possible values of θ.

Consider, for example, the case of the hydrogen atom (a proton and an electron). When we take only the electrostatic forces into account, the ground state of this atom (in the center of mass frame) is described by the ket $|\varphi_{1,0,0}\rangle$, labeled by the quantum numbers $n = 1$, $l = 0$, $m = 0$ (*cf.* Chap. VII). The proton and the electron are spin 1/2 particles. The ground state is therefore four-fold degenerate, and a possible basis in the corresponding subspace is made up of the four vectors:

$$\{|\varphi_{1,0,0}\rangle \otimes |\varepsilon_1, \varepsilon_2\rangle\} \tag{20}$$

where ε_1, and ε_2, equal to + or −, represent respectively the eigenvalues of S_z and I_z (**S** and **I**: the electron and proton spins).

What is the effect on this ground state of the dipole-dipole interaction between **S** and **I**? The matrix elements of W are much smaller than the energy difference between the $1s$ level and the excited levels, so that it is possible to treat the effect of W by perturbation theory. To first order, it can be evaluated by diagonalizing the 4×4 matrix of elements $\langle \varphi_{1,0,0}, \varepsilon_1', \varepsilon_2' | W | \varphi_{1,0,0}, \varepsilon_1, \varepsilon_2 \rangle$. The calculation of these matrix elements, according to (5) and (6), involves angular integrals of the form:

$$\int Y_0^{0*}(\theta, \varphi)\ Y_2^q(\theta, \varphi)\ Y_0^0(\theta, \varphi)\, \mathrm{d}\Omega \tag{21}$$

which are equal to zero, according to the selection rules established in § 1-c above [in this particular case, it can be shown very simply that integral (21) is equal to zero: since Y_0^0 is a constant, expression (21) is proportional to the scalar product of Y_2^q and Y_0^0, which is equal to zero because of the spherical harmonic orthogonality relations].

The dipole-dipole interaction does not modify the energy of the ground state to first order. It enters, however, into the (hyperfine) structure of the excited levels with $l \geq 1$. We must then calculate the matrix elements $\langle \varphi_{n,l,m'}, \varepsilon_1', \varepsilon_2' | W | \varphi_{n,l,m}, \varepsilon_1, \varepsilon_2 \rangle$, that is, the integrals:

$$\int Y_l^{m'*}(\theta, \varphi)\ Y_2^q(\theta, \varphi)\ Y_l^m(\theta, \varphi)\, \mathrm{d}\Omega \tag{22}$$

which, according to (8c), become non-zero as soon as $l \geq 1$.

References and suggestions for further reading:

Evidence in nuclear magnetic resonance experiments of the magnetic dipole interactions between two spins in a rigid lattice: Abragam (14.1), Chap. IV, § II and Chap. VII, § IA; Slichter (14.2), Chap. 3; Pake (14.6).

Complement C_{XI}

Van der Waals forces

1 The electrostatic interaction Hamiltonian for two hydrogen atoms 1152

1-a Notation 1152

1-b Calculation of the electrostatic interaction energy 1153

2 Van der Waals forces between two hydrogen atoms in the $1s$ ground state 1154

2-a Existence of a $-C/R^6$ attractive potential 1154

2-b Approximate calculation of the constant C 1156

2-c Discussion 1156

3 Van der Waals forces between a hydrogen atom in the $1s$ state and a hydrogen atom in the $2p$ state 1158

3-a Energies of the stationary states of the two-atom system. Resonance effect 1158

3-b Transfer of the excitation from one atom to the other 1159

4 Interaction of a hydrogen atom in the ground state with a conducting wall 1160

The character of the forces exerted between two neutral atoms changes with the order of magnitude of the distance R separating these two atoms.

Consider, for example, two hydrogen atoms. When R is of the order of atomic dimensions (that is, of the order of the Bohr radius a_0), the electronic wave functions overlap, and the two atoms attract each other, since they tend to form an H_2 molecule. The potential energy of the system has a minimum[1] for a certain value R_e of the distance R between the atoms. The physical origin of this attraction (and therefore of the chemical bond) lies in the fact that the electrons can oscillate between the two atoms (*cf.* §§ C-2-c and C-3-d of Chapter IV). The stationary wave functions of the two electrons are no longer localized about only one of the nuclei; this lowers the energy of the ground state (*cf.* Complement G_{XI}).

At greater distances, the phenomena change completely. The electrons can no longer move from one atom to the other, since the probability amplitude of such a process decreases with the decreasing overlap of the wave functions, that is, exponentially with the distance. The preponderant effect is then the electrostatic interaction between the electric dipole moments of the two neutral atoms. This gives rise to a total energy which is attractive and which decreases, not exponentially, but with $1/R^6$. This is the origin of the *Van der Waals forces*, which we intend to study in this complement by using stationary perturbation theory (confining ourselves, for the sake of simplicity, to the case of two hydrogen atoms).

It should be clearly understood that the fundamental nature of Van der Waals forces is the same as that of the forces responsible for the chemical bond: the basic

[1] At very short distances, the repulsive forces between the nuclei always dominate.

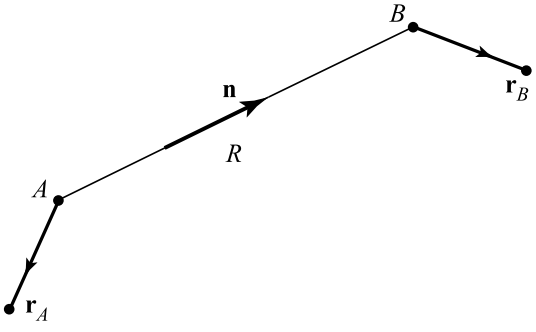


Figure 1: Relative position of the two hydrogen atoms. R is the distance between the two protons, which are situated at A and B, and $\mathbf{n}$ is the unit vector on the line joining them. $\mathbf{r}_A$ and $\mathbf{r}_B$ are the position vectors of the two electrons with respect to points A and B respectively.

Hamiltonian is electrostatic in both cases. Only the variation of the energies of the quantum stationary states of the two-atom system with respect to R allows us to define and differentiate these two types of forces.

Van der Waals forces play an important role in physical chemistry, especially when the two atoms under consideration have no valence electrons (forces between rare gas atoms, stable molecules, etc.). They are partially responsible for the differences between the behavior of a real gas and that of an ideal gas. Finally, as we have already said, these are long-range forces, and are therefore involved in the stability of colloids.

We shall begin by determining the expression for the dipole-dipole interaction Hamiltonian between two neutral hydrogen atoms (§ 1). This will enable us to study the Van der Waals forces between two atoms in the $1s$ state (§ 2), or between an atom in the $2p$ state and an atom in the $1s$ state (§ 3). Finally, we shall show (§ 4) that a hydrogen atom in the $1s$ state is attracted by its electrical mirror image in a perfectly conducting wall.

1. The electrostatic interaction Hamiltonian for two hydrogen atoms

1-a. Notation

The two protons of the two hydrogen atoms are assumed to remain motionless at points A and B (Fig. 1). We shall set:

$$\mathbf{R} = \mathbf{OB} - \mathbf{OA} \tag{1}$$

$$R = |\mathbf{R}| \tag{2}$$

$$\mathbf{n} = \frac{\mathbf{R}}{|\mathbf{R}|} \tag{3}$$

R is the distance between the two atoms, and $\mathbf{n}$ is the unit vector on the line that joins them. Let $\mathbf{r}_A$ be the position vector of the electron attached to atom (A) with respect to point A, and $\mathbf{r}_B$, the position vector of the electron attached to atom B with respect to B. We call:

$$\boldsymbol{\mathcal{D}}_A = q\mathbf{r}_A \tag{4}$$

$$\boldsymbol{\mathcal{D}}_B = q\mathbf{r}_B \tag{5}$$

the electric dipole moments of the two atoms (q is the electron charge).

We shall assume throughout this complement that:

$$R \gg |\mathbf{r}_A|, |\mathbf{r}_B| \tag{6}$$

Although they are identical, the electrons of the two atoms are well separated, and their wave functions do not overlap. It is therefore not necessary to apply the symmetrization postulate (*cf.* Chap. XIV, § D-2-b).

1-b. Calculation of the electrostatic interaction energy

Atom (A) creates at (B) an electrostatic potential U with which the charges of (B) interact. This gives rise to an interaction energy $\mathcal{W}$.

We saw in Complement $\mathrm{E_X}$ that U can be calculated in terms of R, $\mathbf{n}$ and the multipole moments of atom (A). Since (A) is neutral, the most important contribution to U is that of the electric dipole moment $\boldsymbol{\mathcal{D}}_A$. Similarly, since (B) is neutral, the most important term in $\mathcal{W}$ comes from the interaction between the dipole moment $\boldsymbol{\mathcal{D}}_B$ of (B) and the electric field $\mathbf{E} = -\boldsymbol{\nabla} U$ which is essentially created by $\boldsymbol{\mathcal{D}}_A$. This explains the name of "dipole-dipole interaction" given to the dominant term of $\mathcal{W}$. There exist, of course, smaller terms (dipole-quadrupole $\mathcal{W}_{dq}$ and $\mathcal{W}_{qd}$, quadrupole-quadrupole $\mathcal{W}_{qq}$, etc.), and $\mathcal{W}$ is written:

$$\mathcal{W} = \mathcal{W}_{dd} + \mathcal{W}_{dq} + \mathcal{W}_{qd} + \mathcal{W}_{qq} + \cdots \tag{7}$$

To calculate $\mathcal{W}_{dd}$, we shall start with the expression for the electrostatic potential created by $\boldsymbol{\mathcal{D}}_A$ at (B):

$$U(\mathbf{R}) = \frac{1}{4\pi\varepsilon_0}\frac{\boldsymbol{\mathcal{D}}_A \cdot \mathbf{R}}{R^3} \tag{8}$$

which leads to:

$$\mathbf{E} = -\boldsymbol{\nabla}_{\mathbf{R}} U = -\frac{q}{4\pi\varepsilon_0}\frac{1}{R^3}\left[\mathbf{r}_A - 3(\mathbf{r}_A \cdot \mathbf{n})\,\mathbf{n}\right] \tag{9}$$

and, consequently:

$$\mathcal{W}_{dd} = -\mathbf{E} \cdot \boldsymbol{\mathcal{D}}_{\mathbf{B}} = \frac{\mathrm{e}^2}{R^3}\left[\mathbf{r}_A \cdot \mathbf{r}_B - 3(\mathbf{r}_A \cdot \mathbf{n})(\mathbf{r}_B \cdot \mathbf{n})\right] \tag{10}$$

We have set $e^2 = q^2/4\pi\varepsilon_0$, and we have used expressions (4) and (5) for $\boldsymbol{\mathcal{D}}_A$ and $\boldsymbol{\mathcal{D}}_B$. In this complement, we shall choose the Oz axis parallel to $\mathbf{n}$, so that (10) can be written:

$$\mathcal{W}_{dd} = \frac{e^2}{R^3}(x_A x_B + y_A y_B - 2z_A z_B) \tag{11}$$

In quantum mechanics, $\mathcal{W}_{dd}$ becomes the operator W_{dd}, which can be obtained by replacing in (11) x_A, y_A, ... z_B by the corresponding observables X_A, Y_A, ... Z_B, which act in the state spaces $\mathcal{E}_A$ and $\mathcal{E}_B$ of the two hydrogen atoms[2]:

$$W_{dd} = \frac{e^2}{R^3}\left(X_A X_B + Y_A Y_B - 2Z_A Z_B\right) \tag{12}$$

2. Van der Waals forces between two hydrogen atoms in the $1s$ ground state

2-a. Existence of a $-C/R^6$ attractive potential

α. Principle of the calculation

The Hamiltonian of the system is:

$$H = H_{0A} + H_{0B} + W_{dd} \tag{13}$$

where H_{0A} and H_{0B} are the energies of atoms (A) and (B) when they are isolated.

In the absence of W_{dd}, the eigenstates of H are given by the equation:

$$(H_{0A} + H_{0B})\left|\varphi^A_{n,l,m};\, \varphi^B_{n',l'm'}\right\rangle = (E_n + E_{n'})\left|\,\varphi^A_{n,l,m};\, \varphi^B_{n',l',m'}\right\rangle \tag{14}$$

where the $|\varphi_{n,l,m}\rangle$ and the E_n were calculated in § C of Chapter VII. In particular, the ground state of $H_{0A} + H_{0B}$ is $|\varphi^A_{1,0,0}; \varphi^B_{1,0,0}\rangle$, of energy $-2E_I$. It is non-degenerate (we do not take spins into account).

The problem is to evaluate the shift in this ground state due to W_{dd} and, in particular, its R-dependence. This shift represents, so to speak, the interaction potential energy of the two atoms in the ground state.

Since W_{dd} is much smaller than H_{0A} and H_{0B}, we can calculate this effect by stationary perturbation theory.

β. First-order effect of the dipole-dipole interaction

Let us show that the first-order correction:

$$\varepsilon_1 = \left\langle \varphi^A_{1,0,0}\,;\varphi^B_{1,0,0}\left|W_{dd}\right|\varphi^A_{1,0,0};\,\varphi^B_{1,0,0}\right\rangle \tag{15}$$

is zero; the energy ε_1 involves, according to expression (12) for W_{dd}, products of the form $\left\langle\varphi^A_{1,0,0}\left|X_A\right|\varphi^A_{1,0,0}\right\rangle\left\langle\varphi^B_{1,0,0}\left|X_B\right|\varphi^B_{1,0,0}\right\rangle$ (and analogous quantities in which X_A is replaced by Y_A, Z_A and X_B by Y_B, Z_B). These products are zero since, in a stationary state of the atom, the average values of the components of the position operator are zero.

Comment:

The other terms, $W_{dq}, W_{qd}, W_{qq}\ldots$ of expansion (7) involve products of two multipole moments, one relative to (A) and the other one to (B), at least one of which is of order higher than 1. Their contributions are also zero to first order: they

[2]The translational external degrees of freedom of the two atoms are not quantized: for the sake of simplicity, we assume the two protons to be infinitely heavy and motionless. In (12), R is therefore a parameter and not an observable.

are expressed in terms of average values in the ground state of multipole operators of order greater than or equal to one, and we know (*cf.* Complement $\mathrm{E_X}$, § 2-c) that such average values are zero in an $l = 0$ state (triangle rule of Clebsch-Gordan-coefficients). Therefore we must find the second-order effect of W_{dd}, which constitutes the most important energy correction.

γ. Second-order effect of the dipole-dipole interaction

According to the results of Chapter XI, the second-order energy correction can be written:

$$\varepsilon_2 = \sum\nolimits'_{nlm,\ n'l'm'} \frac{\left|\left\langle \varphi^A_{n,l,m}\ ;\varphi^B_{n',l',m'}\left|W_{dd}\right|\varphi^A_{1,0,0};\ \varphi^B_{1,0,0}\right\rangle\right|^2}{-2E_I - E_n - E_{n'}} \tag{16}$$

where the notation $\sum'$ means that the state $\left|\varphi^A_{1,0,0}\ ;\ \varphi^B_{1,0,0}\right\rangle$ is excluded from the summation[3].

Since W_{dd} is proportional to $1/R^3$, ε_2 is proportional to $1/R^6$. Furthermore, all the energy denominators are negative, since we are starting from the ground state. Therefore, *the dipole-dipole interaction gives rise to a negative energy proportional to* $1/R^6$:

$$\varepsilon_2 = -\frac{C}{R^6} \tag{17}$$

Van der Waals forces are therefore attractive and vary with $1/R^7$.

Finally, let us calculate the expansion of the ground state to first order in W_{dd}. We find, according to formula (B-11) of Chapter XI:

$$\begin{aligned}\left|\psi_0\right\rangle &= \left|\varphi^A_{1,0,0}\ ;\ \varphi^B_{1,0,0}\right\rangle \\ &\quad + \sum\nolimits'_{nlm,\ n'l'm'} \left|\varphi^A_{n,l,m}\ ;\ \varphi^B_{n'l'm'}\right\rangle \frac{\left\langle \varphi^A_{n,l,m}\ ;\varphi^B_{n',l',m'}\left|W_{dd}\right|\varphi^A_{1,0,0};\ \varphi^B_{1,0,0}\right\rangle}{-2E_I - E_n - E_{n'}} + \ldots\end{aligned} \tag{18}$$

Comment:

The matrix elements appearing in expressions (16) and (18) above involve the quantities $\left\langle \varphi^A_{n,l,m}\left|X_A\right|\varphi^A_{1,0,0}\right\rangle \left\langle \varphi^B_{n',l',m'}\left|X_B\right|\varphi^B_{1,0,0}\right\rangle$ (and analogous quantities in which X_A and X_B are replaced by Y_A and Y_B or Z_A and Z_B), which are different from zero only if $l = 1$ and $l' = 1$. These quantites are indeed proportional to products of angular integrals

$$\left[\int Y_l^{m*}(\Omega_A)\, Y_1^q(\Omega_A)\, Y_0^0(\Omega_A)\, \mathrm{d}\Omega_A\right] \times \left[\int Y_{l'}^{m'*}(\Omega_B)\, Y_1^{q'}(\Omega_B)\, Y_0^0(\Omega_B)\, \mathrm{d}\Omega_B\right]$$

which, according to the results of Complement $\mathrm{C_X}$, are zero if $l \neq 1$ or $l' \neq 1$. We can therefore, in (16) and (18), replace l and l' by 1.

[3] This summation is performed not only over the bound states, but also over the continuous spectrum of $H_{0A} + H_{0B}$.

2-b. Approximate calculation of the constant C

According to (16) and (12), the constant C appearing in (17) is given by:

$$C = \mathrm{e}^4 \sum\nolimits'_{nlm,\ n'l'm'} \frac{\left|\left\langle \varphi^A_{n,l,m}\,;\varphi^B_{n',l',m'}\left|(\mathrm{X}_A X_B + Y_A Y_B - 2Z_A Z_B)\right|\varphi^A_{1,0,0};\,\varphi^B_{1,0,0}\right\rangle\right|^2}{2E_I + E_n + E_{n'}} \tag{19}$$

We must have $n \geqslant 2$ and $n' \geqslant 2$. For bound states, $|E_n| = E_I/n^2$ is smaller than E_I, and the error is not significant if we replace in (19) E_n and $E_{n'}$ by 0. For states in the continuous spectrum, E_n varies between 0 and $+\infty$. The matrix elements of the numerator become small, however, as soon as the size of E_n becomes appreciable, since the spatial oscillations of the wave function are then numerous in the region in which $\varphi_{1,0,0}(\mathbf{r})$ is non-zero.

To have an idea of the order of magnitude of C, we can therefore replace all the energy denominators of (19) by $2\mathrm{E}_I$. Using the closure relation and the fact that the diagonal element of W_{dd} is zero (§ 2-a-β), we then get:

$$C \simeq \frac{\mathrm{e}^4}{2E_I} \left\langle \varphi^A_{1,0,0};\,\varphi^B_{1,0,0}\right| (\mathrm{X}_A X_B + Y_A Y_B - 2Z_A Z_B)^2 \left|\varphi^A_{1,0,0};\,\varphi^B_{1,0,0}\right\rangle \tag{20}$$

This expression is simple to calculate: because of the spherical symmetry of the $1s$ state, the average values of the cross terms of the type $X_A Y_A, X_B Y_B$, ... are zero. Furthermore, and for the same reason, the various quantities:

$$\left\langle \varphi^A_{1,0,0}\left|X_A^2\right|\,\varphi^A_{1,0,0}\right\rangle, \left\langle \varphi^A_{1,0,0}\left|Y_A^2\right|\,\varphi^A_{1,0,0}\right\rangle \ldots, \left\langle \varphi^B_{1,0,0}\,\left|Z_B^2\right|\,\varphi^B_{1,0,0}\right\rangle$$

are all equal to one third of the average value of $\mathbf{R}_A^2 = X_A^2 + Y_A^2 + Z_A^2$. We finally obtain, using the expression for the wave function $\varphi_{1,0,0}(\mathbf{r})$:

$$C \simeq \frac{\mathrm{e}^4}{2E_I} \times 6 \left|\left\langle \varphi^A_{1,0,0}\right| \frac{\mathbf{R}_A^2}{3} \left|\varphi^A_{1,0,0}\right\rangle\right|^2 = 6\ e^2 a_0^5 \tag{21}$$

(where a_0 is the Bohr radius) and, consequently:

$$\varepsilon_2 \simeq -6e^2 \frac{a_0^5}{R^6} = -6\frac{e^2}{R}\left(\frac{a_0}{R}\right)^5 \tag{22}$$

The preceding calculation is valid only if $a_0 \ll R$ (no overlapping of the wave functions). Thus we see that ε_2 is of the order of the electrostatic interaction between two charges q and $-q$, multiplied by the reduction factor $(a_0/R)^5 \ll 1$.

2-c. Discussion

α. "Dynamical" interpretation of Van der Waals forces

At any given instant, the electric dipole moment (we shall say, more simply, the dipole) of each atom has an average value of zero in the ground state $\left|\varphi^A_{1,0,0}\right\rangle$ or $\left|\varphi^B_{1,0,0}\right\rangle$. This does not mean that any individual measurement of a component of this dipole will yield zero. If we make such a measurement, we generally find a non-zero value; however,

we have the same probability of finding the opposite value. The dipole of a hydrogen atom in the ground state should therefore be thought of as constantly undergoing random fluctuations.

We shall begin by neglecting the influence of one dipole on the motion of the other one. Since the two dipoles are then fluctuating randomly and independently, their mean interaction is zero: this explains the fact that W_{dd} has no first-order effect.

However, the two dipoles are not really independent. Consider the electrostatic field created by dipole (A) at (B). This field follows the fluctuations of dipole (A). The dipole it induces at (B) is therefore correlated with dipole (A), so the electrostatic field which "returns" to (A) is no longer uncorrelated with the motion of dipole (A). Thus, although the motion of dipole (A) is random, its interaction with its own field, which is "reflected" to it by (B), does not have a average value of zero. This is the physical interpretation of the second-order effect of W_{dd}.

The dynamical aspect is therefore useful for understanding the origin of Van der Waals forces. If we were to think of the two hydrogen atoms in the ground state as two spherical and "static" clouds of negative electricity (with a positive point charge at the center of each one), we would be led to a rigorously zero interaction energy.

β. Correlations between the two dipole moments

Let us show more precisely that there exists a correlation between the two dipoles.

When we take W_{dd} into account, the ground state of the system is no longer $\left|\varphi^A_{1,0,0};\varphi^B_{1,0,0}\right\rangle$, but $|\psi_0\rangle$ [*cf.* expression (18)]. As shown below, a simple calculation yields:

$$\langle\psi_0|\,X_A\,|\psi_0\rangle = ... = \langle\psi_0|\,Z_B\,|\psi_0\rangle = 0 \tag{23}$$

to first order in W_{dd}.

Consider, for example, the matrix element $\langle\psi_0|\,X_A\,|\psi_0\rangle$. The zeroth-order term, $\left\langle\varphi^A_{1,0,0};\varphi^B_{1,0,0}\,|X_A|\;\varphi^A_{1,0,0};\varphi^B_{1,0,0}\right\rangle$ is zero, since it is equal to the average value of X_A in the ground state $\left|\varphi^A_{1,0,0}\right\rangle$. To first order, the summation appearing in formula (18) must be included. Since W_{dd} contains only products of the form X_AX_B, the coefficients of the kets $\left|\varphi^A_{1,0,0};\;\varphi^B_{n',l',m'}\right\rangle$ and $\left|\varphi^A_{n,l,m};\;\varphi^B_{1,0,0}\right\rangle$ in this summation are zero. The first-order terms which could be different from zero are therefore proportional to

$$\left\langle\varphi^A_{n,l,m};\varphi^B_{n',l',m'}\,|X_A|\;\varphi^A_{1,0,0};\varphi^B_{1,0,0}\right\rangle, \quad \text{with} \;\; l\neq 0 \;\; \text{and} \;\; l'\neq 0;$$

These terms are all zero since X_A does not act on $|\varphi^B_{1,0,0}\rangle$ and $\langle\varphi^B_{n',l',m'}|\varphi^B_{1,0,0}\rangle = 0$ for $l'\neq 0$.

Thus, even in the presence of an interaction, the average values of the components of each dipole are zero. This is not surprising: in the interpretation of § 2-c-α, the dipole induced in (B) by the field of dipole (A) fluctuates randomly, like this field, and has, consequently, an average value of zero.

Let us show, on the other hand, that the two dipoles are correlated, by evaluating the average value of a product of two components, one relative to dipole (A) and the other, to dipole (B). We shall calculate $\langle\psi_0|\,(X_AX_B + Y_AY_B - 2Z_AZ_B)\,|\psi_0\rangle$, for example, which, according to (12), is nothing more than $(R^3/e^2)\,\langle\psi_0|\,W_{dd}\,|\psi_0\rangle$. Using (18),

we immediately find, taking (15) and (16) into account, that:

$$\langle\psi_0|\left(X_AX_B + Y_AY_B - 2Z_AZ_B\right)|\psi_0\rangle = 2\varepsilon_2\frac{R^3}{e^2} \neq 0 \tag{24}$$

Thus, the average values of the products X_AX_B, Y_AY_B and Z_AZ_B are not zero, as would be the products of average values $\langle X_A\rangle\langle X_B\rangle$, $\langle Y_A\rangle\langle Y_B\rangle$ and $\langle Z_A\rangle\langle Z_B\rangle$ according to (23). This proves the existence of a correlation between the two dipoles.

γ. Long-range modification of Van der Waals forces

The description of § 2-c-α above enables us to understand that the preceding calculations are no longer valid if the two atoms are too far apart. The field produced by (A) and "reflected" by (B) returns to (A) with a time lag due to the propagation $(A) \to (B) \to (A)$, and we have argued as if the interactions were instantaneous.

We can see that this propagation time can no longer be neglected when it becomes of the order of the characteristic times of the atom's evolution, that is, of the order of $2\pi/\omega_{n1}$, where $\omega_{n1} = (E_n - E_1)/\hbar$ denotes a Bohr angular frequency. In other words, the calculations performed in this complement assume that the distance R between the two atoms is much smaller than the wavelengths $2\pi c/\omega_{n1}$ of the spectrum of these atoms (about 1 000 Å).

A calculation which takes propagation effects into account gives an interaction energy which, at large distances, decreases as $1/R^7$. The $1/R^6$ law which we have found therefore applies to an intermediate range of distances, neither too large (because of the time lag) nor too small (to avoid overlapping of the wave functions).

3. Van der Waals forces between a hydrogen atom in the $1s$ state and a hydrogen atom in the $2p$ state

3-a. Energies of the stationary states of the two-atom system. Resonance effect

The first excited level of the unperturbed Hamiltonian $H_{0A} + H_{0B}$ is eight-fold degenerate. The associated eigensubspace is spanned by the eight states : $\left\{\left|\varphi^A_{1,0,0};\ \varphi^B_{2,0,0}\right\rangle;\ \left|\varphi^A_{2,0,0};\ \varphi^B_{1,0,0}\right\rangle;\ \left|\varphi^A_{1,0,0};\ \varphi^B_{2,1,m}\right\rangle,\ \text{with}\ m = -1,\ 0,\ +1;\right.$ $\left.\left|\varphi^A_{2,1,m'};\ \varphi^B_{1,0,0}\right\rangle,\ \text{with}\ m' = -1, 0, +1\right\}$, which correspond to a situation in which one of the two atoms is in the ground state, while the other one is in a state of the $n = 2$ level.

According to perturbation theory for a degenerate state, to obtain the first-order effect of W_{dd}, we must diagonalize the 8×8 matrix representing the restriction of W_{dd} to the eigensubspace. We shall show that the only non-zero matrix elements of W_{dd} are those which connect a state $\left|\varphi^A_{1,0,0};\ \varphi^B_{2,1,m}\right\rangle$ to the state $\left|\varphi^A_{2,1,m};\ \varphi^B_{1,0,0}\right\rangle$. The operators X_A, Y_A, Z_A appearing in the expression for W_{dd} are odd and can therefore couple $\left|\varphi^A_{1,0,0}\right\rangle$ only to one of the $\left|\varphi^A_{2,1,m}\right\rangle$; an analogous argument is valid for X_B, Y_B, Z_B. Finally, the dipole-dipole interaction is invariant under a rotation of the two atoms about the Oz axis which joins them; therefore W_{dd} commutes with $L_{Az} + L_{Bz}$ and can only join two states for which the sum of the eigenvalues of L_{Az} and L_{Bz} is the same.

Therefore, the preceding 8×8 matrix can be broken down into four 2×2 matrices. One of them is entirely zero (the one which concerns the $2s$ states), and the other three

are of the form:

$$\begin{pmatrix} 0 & k_m/R^3 \\ k_m/R^3 & 0 \end{pmatrix} \tag{25}$$

where we have set:

$$\left\langle \varphi^A_{1,0,0}; \varphi^B_{2,1,m} \right| W_{dd} \left| \varphi^A_{2,1,m}; \varphi^B_{1,0,0} \right\rangle = \frac{k_m}{R^3} \tag{26}$$

k_m is a calculable constant, of the order of $e^2 a_0^2$, which will not be specified here.

We can easily diagonalize matrix (25), obtaining the eigenvalues $+k_m/R^3$ and $-k_m/R^3$, associated respectively with the eigenstates:

$$\frac{1}{\sqrt{2}} \left(\left| \varphi^A_{1,0,0}; \varphi^B_{2,1,m} \right\rangle + \left| \varphi^A_{2,1,m}; \varphi^B_{1,0,0} \right\rangle \right)$$

and:

$$\frac{1}{\sqrt{2}} \left(\left| \varphi^A_{1,0,0}; \varphi^B_{2,1,m} \right\rangle - \left| \varphi^A_{2,1,m}; \varphi^B_{1,0,0} \right\rangle \right)$$

This reveals the following important results:

– The interaction energy varies as l/R^3 and not as $1/R^6$, since W_{dd} now modifies the energies to first order. The Van der Waals forces are therefore more important than they were between two hydrogen atoms in the $1s$ state (resonance effect between two different states of the total system with the same unperturbed energy).

– The sign of the interaction can be positive or negative (eigenvalues $+k_m/R^3$ and $-k_m/R^3$). There exist therefore states of the two-atom system for which there is attraction, and others for which there is repulsion.

3-b. Transfer of the excitation from one atom to the other

The two states $\left| \varphi^A_{1,0,0}; \varphi^B_{2,1,m} \right\rangle$ and $\left| \varphi^A_{2,1,m}; \varphi^B_{1,0,0} \right\rangle$ have the same unperturbed energy and are coupled by a non-diagonal perturbation. According to the general results of § C of Chapter IV (two-level system), we know that there is oscillation of the system from one level to the other with a frequency proportional to the coupling.

Therefore, if the system starts in the state $\left| \varphi^A_{1,0,0}; \varphi^B_{2,1,m} \right\rangle$ at $t = 0$, it arrives after a certain time (the larger R, the longer the time) in the state $\left| \varphi^A_{2,1,m}; \varphi^B_{1,0,0} \right\rangle$. The excitation thus passes from (B) to (A), then returns to (B), and so on.

Comment:

If the two atoms are not fixed but, for example, undergo collision, R varies over time and the passage of the excitation from one atom to the other is no longer periodic. The corresponding collisions, called resonant collisions, play an important role in the broadening of spectral lines.

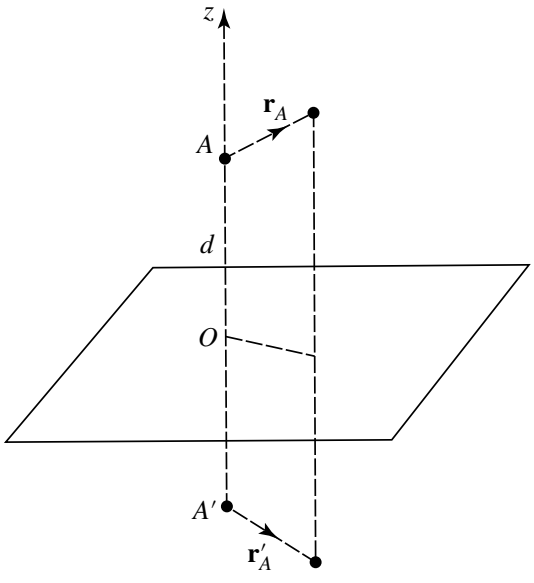


Figure 2: To calculate the interaction energy of a hydrogen atom with a perfectly conducting wall, we can assume that the electric dipole moment $q\mathbf{r}_A$ of the atom interacts with its electrical image $-q\mathbf{r}'_A$ (d is the distance between the proton A and the wall).

4. Interaction of a hydrogen atom in the ground state with a conducting wall

We shall now consider a single hydrogen atom (A) situated at a distance d from a wall which is assumed to be perfectly conducting. The Oz axis is taken along the perpendicular to the wall passing through A (Fig. 2). The distance d is assumed to be much larger than the atomic dimensions. We can therefore ignore the atomic structure of the wall, and assume that the atom interacts with its electrical image on the other side of this wall (that is, with a symmetrical atom with opposite charges). The dipole interaction energy between the atom and the wall can easily be obtained from expression (12) for W_{dd} by making the following substitutions:

$$\begin{cases} e^2 \to -e^2 \\ R \ \to 2d \\ X_B \to X'_A = X_A \\ Y_B \to Y'_A = Y_A \\ Z_B \to Z'_A = -Z_A \end{cases} \tag{27}$$

(the change of e^2 to $-e^2$ is due to the sign difference of the image charges).

Furthermore it is necessary to divide by 2 since the dipole image is fictitious,

proportional to the electric dipole of the atom[4]. We then get:

$$W = -\frac{e^2}{16d^3}(X_A^2 + Y_A^2 + 2Z_A^2) \tag{28}$$

which represents the interaction energy of the atom with the wall [W acts only on the degrees of freedom of (A)].

If the atom is in its ground state, the energy correction to first order in W is then:

$$\varepsilon'_1 = \langle \varphi_{1,0,0} | W | \varphi_{1,0,0} \rangle \tag{29}$$

Using the spherical symmetry of the 1s state, we obtain:

$$\varepsilon'_1 = -\frac{e^2}{16d^3} 4 \langle \varphi_{1,0,0} | \frac{\mathbf{R}_A^2}{3} | \varphi_{1,0,0} \rangle = -\frac{e^2 a_0^2}{4d^3} \tag{30}$$

We see that the atom is attracted by the wall: the attraction energy varies as $1/d^3$, and, therefore, the force of attraction varies as $1/d^4$.

The fact that W has an effect even to first order can easily be understood in terms of the discussion of § 2-c above. In the present case, there is a perfect correlation between the two dipoles, since they are images of each other.

References and suggestions for further reading:

Kittel (13.2), Chap. 3. p. 82; Davydov (1.20), Chap. XII. §§ 124 and 125; Langbein (12.9).

For a discussion of retardation effects, see: Power (2.11), §§ 7.5 and 8.4 (quantum electrodynamic approach); Landau and Lifshitz (7.12), Chap. XIII, § 90 (electromagnetic fluctuation approach).

See also Derjaguin's article (12.12).

[4]This 1/2 factor is easily understood if one remembers that the energy of an electrostatic system is proportional to the integral over all space of the square of the electric field. For the system of Fig. 2, the electric field is zero below the xOy plane.

Complement D_{XI}

The volume effect: the influence of the spatial extension of the nucleus on the atomic levels

1 First-order energy correction 1164
1-a Calculation of the correction 1164
1-b Discussion 1165
2 Application to some hydrogen-like systems 1166
2-a The hydrogen atom and hydrogen-like ions 1166
2-b Muonic atoms 1167

The energy levels and the stationary states of the hydrogen atom were studied in Chapter VII by assuming the proton to be a charged point particle, which creates an electrostatic $1/r$ Coulomb potential. Actually, this is not quite true. The proton is not strictly a point charge; its charge fills a volume which has a certain size (of the order of 1 fermi $= 10^{-13}$ cm). When an electron is extremely close to the center of the proton, it "sees" a potential that no longer varies as $1/r$, and which depends on the spatial charge distribution associated with the proton. This is true, furthermore, for all atoms: inside the volume of the nucleus, the electrostatic potential depends on how the charges are distributed. We thus expect the atomic energy levels, which are determined by the potential to which the electrons are subject at all points of space, to be affected by this distribution: this is what is called the "volume effect". The experimental and theoretical study of such an effect is therefore important, since it can supply information about the internal structure of nuclei.

In this complement, we shall give a simplified treatment of the volume effect of hydrogen-like atoms. To have an idea of the order of magnitude of the energy shifts it causes, we shall confine ourselves to a model in which the nucleus is represented by a sphere of radius ρ_0, in which the charge $-Zq$ is uniformly distributed. In this model, the potential created by the nucleus is (*cf.* Complement A_V, § 4-b):

$$V(r) = \begin{cases} -\dfrac{Ze^2}{r} & \text{for } r \geqslant \rho_0 \\ \dfrac{Ze^2}{2\rho_0}\left[\left(\dfrac{r}{\rho_0}\right)^2 - 3\right] & \text{for } r \leqslant \rho_0 \end{cases} \tag{1}$$

(we have set $e^2 = q^2/4\pi\varepsilon_0$). The shape of the variation of $V(r)$ with respect to r is shown in Figure 1.

The exact solution of the Schrödinger equation for an electron subject to such a potential poses a complicated problem. Therefore, we shall content ourselves with an approximate solution, based on perturbation theory. In a first approximation, we shall consider the potential to be a Coulomb potential [which amounts to setting $\rho_0 = 0$ in (1)]. The energy levels of the hydrogen atom are then the ones found in § C of Chapter VII.

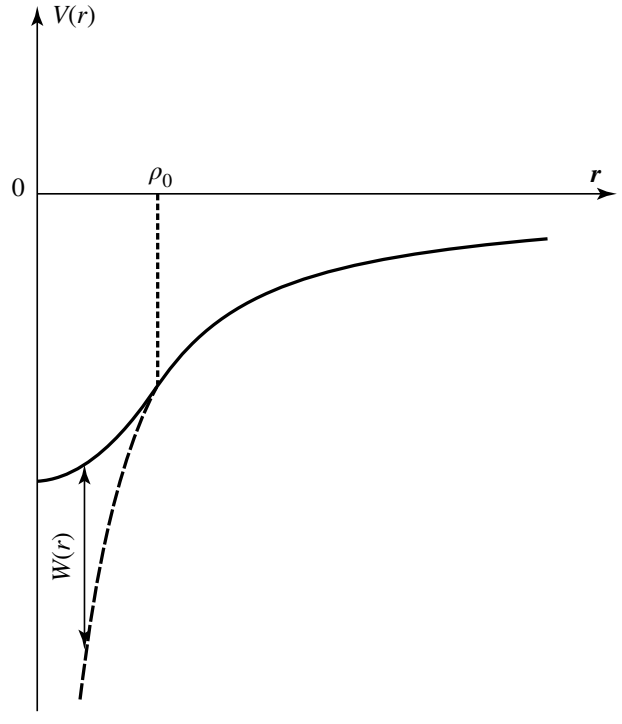


Figure 1: Variation with respect to r of the electrostatic potential $V(r)$ created by the charge distribution $-Zq$ of the nucleus, assumed to be uniformly distributed inside a sphere of radius ρ_0. For $r \leqslant \rho_0$, the potential is parabolic. For $r \geqslant \rho_0$, it is a Coulomb potential [the extension of this Coulomb potential into the $r \leqslant \rho_0$ zone is represented by the dashed line; $W(r)$ is the difference between $V(r)$ and the Coulomb potential].

We shall treat the difference $W(r)$ between the potential $V(r)$ written in (1) and the Coulomb potential as a perturbation. This difference is zero when r is greater than the radius ρ_0 of the nucleus. It is therefore reasonable that it should cause a small shift in the atomic levels (the corresponding wave functions extend over dimensions of the order of $a_0 \gg \rho_0$), which justifies a treatment by first-order perturbation theory.

1. First-order energy correction

1-a. Calculation of the correction

By definition, $W(r)$ is equal to:

$$W(r) = \begin{cases} \dfrac{Ze^2}{2\rho_0}\left[\left(\dfrac{r}{\rho_0}\right)^2 + \dfrac{2\rho_0}{r} - 3\right] & \text{if } 0 \leqslant r \leqslant \rho_0 \\ 0 & \text{if } r \geqslant \rho_0 \end{cases} \tag{2}$$

Let $|\varphi_{n,l,m}\rangle$ be the stationary states of the hydrogen-like atom in the absence of the perturbation W. To evaluate the effect of W to first order, we must calculate the matrix elements:

$$\langle\varphi_{n,l,m}|\, W\, |\varphi_{n,l',m'}\rangle = \int \mathrm{d}\Omega\, Y_l^{m*}(\Omega)\, Y_{l'}^{m'}(\Omega) \times \int_0^\infty r^2\, \mathrm{d}r R_{n,l}^*(r)\, R_{n,l'}(r)\, W(r) \tag{3}$$

In this expression, the angular integral simply gives $\delta_{ll'}\delta_{mm'}$. To simplify the radial integral, we shall make an approximation and assume[1] that:

$$\rho_0 \ll a_0 \tag{4}$$

that is, that the $r \leq \rho_0$ region, in which $W(r)$ is not zero, is much smaller than the spatial extent of the functions $R_{n,l}(r)$. When $r \leq \rho_0$, we then have:

$$R_{n,l}(r) \simeq R_{n,l}(0) \tag{5}$$

The radial integral can therefore be written:

$$I = \frac{Ze^2}{2\rho_0}\left|R_{n,l}(0)\right|^2 \int_0^{\rho_0} r^2 dr \left[\left(\frac{r}{\rho_0}\right)^2 + \frac{2\rho_0}{r} - 3\right] \tag{6}$$

which gives:

$$I = \frac{Ze^2}{10}\rho_0^2\left|R_{n,l}(0)\right|^2 \tag{7}$$

and:

$$\langle\varphi_{n,l,m}|\, W\, |\varphi_{n,l',m'}\rangle = \frac{Ze^2}{10}\rho_0^2\left|R_{n,l}(0)\right|^2 \delta_{ll'}\,\delta_{mm'} \tag{8}$$

We see that the matrix representing W in the subspace $\mathcal{E}_n$ corresponding to the nth level of the unperturbed Hamiltonian is diagonal. Therefore, the first-order energy correction associated with each state $|\varphi_{n,l,m}\rangle$ can be written simply:

$$\Delta E_{n,l} = \frac{Ze^2}{10}\rho_0^2\left|R_{n,l}(0)\right|^2 \tag{9}$$

[1] This is certainly the case for the hydrogen atom. In § 2, we shall examine condition (4) in greater detail.

This correction does not depend[2] on m. Furthermore, since $R_{n,l}(0)$ is zero unless $l = 0$ (*cf.* Chap. VII, § C-4-c), only the s states ($l = 0$ states) are shifted, by a quantity which is equal to:

$$\begin{aligned}\Delta E_{n,0} &= \frac{Ze^2}{10}\rho_0^2 \left|R_{n,0}(0)\right|^2 \\ &= \frac{2\pi Ze^2}{5}\rho_0^2 \left|\varphi_{n,0,0}(0)\right|^2 \end{aligned} \tag{10}$$

(we have used the fact that $Y_0^0 = 1/\sqrt{4\pi}$).

1-b. Discussion

$\Delta E_{n,0}$ can be written:

$$\Delta E_{n,0} = \frac{3}{10} w\mathcal{P} \tag{11}$$

where:

$$w = \frac{Ze^2}{\rho_0} \tag{12}$$

is the absolute value of the potential energy of the electron at a distance ρ_0 from the center of the nucleus, and:

$$\mathcal{P} = \frac{4}{3}\pi\rho_0^3 \left|\varphi_{n,0,0}(0)\right|^2 \tag{13}$$

is the probability of finding the electron inside the nucleus. $\mathcal{P}$ and w enter into (11) because the effect of the perturbation $W(r)$ is felt only inside the nucleus.

For the method which led us to (10) and (11) to be consistent, the correction $\Delta E_{n,0}$ must be much smaller than the energy differences between unperturbed levels. Since w is very large (an electron and a proton attract each other very strongly when they are very close), $\mathcal{P}$ must therefore be extremely small. Before taking up the more precise calculation in § 2, we shall evaluate the order of magnitude of these quantities. Let:

$$a_0(Z) = \frac{\hbar^2}{Zme^2} \tag{14}$$

be the Bohr radius when the total charge of the nucleus is $-Zq$. If n is not too high, the wave functions $\varphi_{n,0,0}(\mathbf{r})$ are practically localized inside a region of space whose volume is approximately $[[a_0(Z)]^3$. As for the nucleus, its volume is of the order of ρ_0^3, so:

$$\mathcal{P} \simeq \left[\frac{\rho_0}{a_0(Z)}\right]^3 \tag{15}$$

[2] This result could have been expected, since the perturbation W, which is invariant under rotation, is a scalar (*cf.* Complement B_{VI}, § 5-b).

Relation (11) then yields:

$$\Delta E_{n,0} \simeq \frac{Ze^2}{\rho_0}\left[\frac{\rho_0}{a_0(Z)}\right]^3 \simeq \frac{Ze^2}{a_0(Z)}\left[\frac{\rho_0}{a_0(Z)}\right]^2 \tag{16}$$

Now, $Ze^2/a_0(Z)$ is of the order of magnitude of the binding energy $E_I(Z)$ of the unperturbed atom. The relative value of the correction is therefore equal to:

$$\frac{\Delta E_{n,0}}{E_I(Z)} \simeq \left[\frac{\rho_0}{a_0(Z)}\right]^2 \tag{17}$$

If condition (4) is met, this correction will indeed be very small. We shall now calculate it more precisely in some special cases.

2. Application to some hydrogen-like systems

2-a. The hydrogen atom and hydrogen-like ions

For the ground state of the hydrogen atom, we have [*cf.* Chap. VII, relation (C-39a)]:

$$R_{1,0}(r) = 2(a_0)^{-3/2}\mathrm{e}^{-r/a_0} \tag{18}$$

[where a_0 is obtained by setting $Z = 1$ in (14)]. Formula (10) then yields:

$$\Delta E_{1,0} = \frac{2}{5}\frac{e^2}{a_0}\left(\frac{\rho_0}{a_0}\right)^2 = \frac{4}{5}E_I\left(\frac{\rho_0}{a_0}\right)^2 \tag{19}$$

Now, we know that, for hydrogen:

$$a_0 \simeq 0,53 \text{ Å} = 5.3\ 10^{-11} \text{ m} \tag{20}$$

Furthermore, the radius ρ_0 of the proton is of the order of:

$$\rho_0 \text{ (proton)} \simeq 1 \text{ F} = 10^{-15} \text{ m} \tag{21}$$

If we substitute these numerical values into (19), we obtain:

$$\Delta E_{1,0} \simeq 4.5\ 10^{-10} E_I \simeq 6\ 10^{-9} \text{ eV} \tag{22}$$

The result is therefore very small.

For a hydrogen-like ion, the nucleus has a charge of $-Zq$. We can then apply (10), which amounts to replacing e^2 in (19) by $Z\mathrm{e}^2$, and a_0 by $a_0(Z) = a_0/Z$. We obtain:

$$\Delta E_{1,0}(Z) = \frac{2}{5}\frac{Z^2e^2}{a_0}\left[\frac{\rho_0(A,Z)}{a_0} \times Z\right]^2 \tag{23}$$

where $\rho_0(A, Z)$ is the radius of the nucleus, composed of A nucleons (protons or neutrons), Z of which are protons. In practice, the number of nucleons of a nucleus is not very

different from $2\,Z$; in addition, the "nuclear density saturation" property is expressed by the approximate relation:

$$\rho_0(A, Z) \propto \mathrm{A}^{1/3} \propto \mathrm{Z}^{1/3} \tag{24}$$

The variation of the energy correction with respect to Z is then given by:

$$\Delta \mathrm{E}_{1,0}\,(\mathrm{Z}) \propto \mathrm{Z}^{14/3} \tag{25}$$

or:

$$\frac{\Delta \mathrm{E}_{1,0}(\mathrm{Z})}{\mathrm{E}_I(\mathrm{Z})} \propto \mathrm{Z}^{8/3} \tag{26}$$

$\Delta \mathrm{E}_{1,0}\,(\mathrm{Z})$ therefore varies very rapidly with Z, under the effect of several concordant factors: when Z increases, a_0 decreases and ρ_0 increases. The volume effect is therefore significantly larger for heavy hydrogen-like ions than for hydrogen.

Comment:

The volume effect also exists for all the other atoms. It is responsible for an isotopic shift of the lines of the emission spectrum. For two distinct isotopes of the same chemical element, the number Z of protons of the nucleus is the same, but the number $A - Z$ of neutrons is different; the spatial distributions of the nuclear charges are therefore not identical for the two nuclei.

Actually, for light atoms, the isotopic shift is caused principally by the nuclear finite mass effect (*cf.* Complement $\mathrm{A_{VII}}$, § 1-a-α). On the other hand, for heavy atoms (for which the reduced mass varies very little from one isotope to another), the finite mass effect is small; however, the volume effect increases with Z and becomes preponderant.

2-b. Muonic atoms

We have already discussed some simple properties of muonic atoms (*cf.* Complement $\mathrm{A_V}$, § 4 and $\mathrm{A_{VII}}$, § 2-a). In particular, we have pointed out that the Bohr radius associated with them is distinctly smaller than for ordinary atoms (this is caused by the fact that the mass of the μ^- muon is approximately equal to 207 times that of the electron). From the qualitative discussion of § 1-b, we may therefore expect an important volume effect for muonic atoms. We shall evaluate it by choosing two limiting cases: a light muonic atom (hydrogen) and a heavy one (lead).

α. *The muonic hydrogen atom*

The Bohr radius is then:

$$a_0(\mu^-, p^+) \simeq \frac{a_0}{207} \tag{27}$$

that is, of the order of 250 fermi. It therefore remains, in this case, distinctly greater than ρ_0. If we replace a_0 by $a_0/207$ in (19), we find:

$$\Delta \mathrm{E}_{1,0}(\mu^-, p^+) \simeq 1.9 \times 10^{-5}\ \mathrm{E}_I(\mu^-,\, \mathrm{p}^+) \simeq 5 \times 10^{-2}\ \mathrm{eV} \tag{28}$$

Although the volume effect is much larger than for the ordinary hydrogen atom, it still yields only a small correction to the energy levels.

β. The muonic lead atom

The Bohr radius of the muonic lead atom is [*cf.* Complement A_V, relation (25)]:

$$a_0(\mu^-, \text{Pb}) \simeq 3 \times 10^{-15} \text{ m} \tag{29}$$

The μ^- muon is now very close to the lead nucleus; it is therefore practically unaffected by the repulsion of the atomic electrons which are located at distinctly greater distances. This could lead us to believe that (10), which was proven for hydrogen-like atoms and ions, is directly applicable to this case. Actually, this is not true, since the radius of the lead nucleus is equal to:

$$\rho_0(\text{Pb}) \simeq 8.5 \text{ F} = 8.5 \times 10^{-15} \text{ m} \tag{30}$$

which is not small compared to $a_0(\mu^-, \text{Pb})$. Equation (10) would therefore lead to large corrections (several MeV), of the same order of magnitude as the energy $E_I(\mu^-, \text{Pb})$. This means that, in this case, the volume effect can no longer be treated as a perturbation (see discussion of § 4 of Complement A_V). To calculate the energy levels, it is necessary to know the potential $V(r)$ exactly and to solve the corresponding Schrödinger equation.

The muon is therefore more inside the nucleus than outside, that is, according to (1), in a region in which the potential is parabolic. In a first approximation, we could consider the potential to be parabolic everywhere (as is done in Complement A_V) and then treat as a perturbation the difference which exists for $r \geqslant \rho_0$ between the real potential and the parabolic potential. However, the extension of the wave function corresponding to such a potential is not sufficiently smaller than ρ_0 for such an approximation to lead to precise results, and the only valid method consists of solving the Schrödinger equation corresponding to the real potential.

References and suggestions for further reading:

The isotopic volume effect: Kuhn (11.1), Chap. VI, § C-3; Sobel'man (11.12), Chap. 6, § 24.

Muonic atoms (sometimes called mesic atoms): Cagnac and Pebay-Peyroula (11.2), Chap. XIX, § 7-C; De Benedetti (11.21); Wiegand (11.22); Weissenberg (16.19), § 4-2.

Complement E_{XI}

The variational method

1 Principle of the method 1169
1-a A property of the ground state of a system 1169
1-b Generalization: the Ritz theorem 1170
1-c A special case where the trial functions form a subspace . . . 1171
2 Application to a simple example 1172
2-a Exponential trial functions 1172
2-b Rational wave functions 1174
3 Discussion 1174

The perturbation theory studied in Chapter XI is not the only general approximation method applicable to conservative systems. We shall give a concise description here of another of these methods, which also has numerous applications, especially in atomic and molecular physics, nuclear physics, and solid state physics. First of all, we shall indicate, in § 1, the principle of the variational method. Then we shall use the simple example of the one-dimensional harmonic oscillator to bring out its principal features (§ 2), which we shall briefly discuss in § 3. Complements F_{XI} and G_{XI} apply the variational method to simple models which enable us to understand the behavior of electrons in a solid and the origin of the chemical bond.

1. Principle of the method

Consider an arbitrary physical system whose Hamiltonian H is time-independent. To simplify the notation, we shall assume that the entire spectrum of H is discrete and non-degenerate:

$$\mathrm{H}\,|\varphi_n\rangle = E_n\,|\varphi_n\rangle \;;n = 0,\,1,\,2,\,... \tag{1}$$

Although the Hamiltonian H is known, this is not necessarily the case for its eigenvalues E_n and the corresponding eigenstates $|\varphi_n\rangle$. The variational method is, of course, most useful in the cases in which we do not know how to diagonalize H exactly.

1-a. A property of the ground state of a system

Choose an arbitrary ket $|\psi\rangle$ of the state space of the system. The average value of the Hamiltonian H in the state $|\psi\rangle$ is such that:

$$\langle \mathrm{H}\rangle = \frac{\langle\psi|\,\mathrm{H}\,|\psi\rangle}{\langle\psi|\psi\rangle} \geqslant \mathrm{E}_0 \tag{2}$$

(where E_0 is the smallest eigenvalue of H), equality occuring if and only if $|\psi\rangle$ is an eigenvector of H with the eigenvalue E_0.

To prove inequality (2), we expand the ket $|\psi\rangle$ on the basis of eigenstates of H:

$$|\psi\rangle = \sum_n c_n |\varphi_n\rangle \tag{3}$$

We then have:

$$\langle\psi| H |\psi\rangle = \sum_n |c_n|^2 E_n \geqslant E_0 \sum_n |c_n|^2 \tag{4}$$

with, of course:

$$\langle\psi \mid \psi\rangle = \sum_n |c_n|^2 \tag{5}$$

which proves (2). For inequality (4) to become an equality, it is necessary and sufficient that all the coefficients c_n be zero, with the exception of c_0; $|\psi\rangle$ is then an eigenvector of H with the eigenvalue E_0.

This property is the basis for a method of approximate determination of E_0. We choose (in theory, arbitrarily, but in fact, by using physical criteria) a family of kets $|\psi(\alpha)\rangle$ which depend on a certain number of parameters which we symbolize by α. We calculate the average value $\langle H\rangle(\alpha)$ of the Hamiltonian H in these states, and we minimize $\langle H\rangle(\alpha)$ with respect to the parameters α. The minimal value so obtained constitutes an approximation of the ground state E_0 of the system. The kets $|\psi(\alpha)\rangle$ are called *trial kets*, and the method itself, the *variational method.*

Comment:

The preceding proof can easily be generalized to cases in which the spectrum of H is degenerate or includes a continuous part.

1-b. Generalization: the Ritz theorem

We shall show that, more generally, *the average value of the Hamiltonian H is stationary in the neighborhood of its discrete eigenvalues.*

Consider the average value of H in the state $|\psi\rangle$:

$$\langle H\rangle = \frac{\langle\psi| H |\psi\rangle}{\langle\psi \mid \psi\rangle} \tag{6}$$

as a functional of the state vector $|\psi\rangle$, and calculate its increment $\delta\langle H\rangle$ when $|\psi\rangle$ becomes $|\psi\rangle + |\delta\psi\rangle$, where $|\delta\psi\rangle$ is assumed to be infinitely small. To do so, it is useful to write (6) in the form:

$$\langle H\rangle \langle\psi \mid \psi\rangle = \langle\psi| H |\psi\rangle \tag{7}$$

and to differentiate both sides of this relation:

$$\begin{aligned} \langle\psi \mid \psi\rangle \delta\langle H\rangle + \langle H\rangle [\langle\psi \mid \delta\psi\rangle + \langle\delta\psi \mid \psi\rangle] \\ = \langle\psi| H |\delta\psi\rangle + \langle\delta\psi| H |\psi\rangle \end{aligned} \tag{8}$$

that is, since $\langle H \rangle$ is a number:

$$\langle \psi | \psi \rangle \, \delta \langle H \rangle = \langle \psi | [H - \langle H \rangle] | \delta\psi \rangle + \langle \delta\psi | [H - \langle H \rangle] | \psi \rangle \tag{9}$$

The average value $\langle H \rangle$ will be stationary if:

$$\delta \langle H \rangle = 0 \tag{10}$$

which, according to (9), means that:

$$\langle \psi | [H - \langle H \rangle] | \delta\psi \rangle + \langle \delta\psi | [H - \langle H \rangle] | \psi \rangle = 0 \tag{11}$$

We set:

$$|\varphi\rangle = [H - \langle H \rangle] | \psi \rangle \tag{12}$$

Relation (11) can then be written simply:

$$\langle \varphi \mid \delta\psi \rangle + \langle \delta\psi \mid \varphi \rangle = 0 \tag{13}$$

This last relation must be satisfied for any infinitesimal ket $|\delta\psi\rangle$. In particular, if we choose:

$$|\delta\psi\rangle = \delta\lambda \, |\varphi\rangle \tag{14}$$

(where $\delta\lambda$ is an infinitely small real number), (13) becomes:

$$2 \langle \varphi \mid \varphi \rangle \, \delta\lambda = 0 \tag{15}$$

The norm of the ket $|\varphi\rangle$ is therefore zero, and $|\varphi\rangle$ must consequently be zero. With definition (12) taken into account, this means that:

$$H \, |\psi\rangle = \langle H \rangle \, |\psi\rangle \tag{16}$$

Consequently, the average value $\langle H \rangle$ is stationary if and only if the state vector $|\psi\rangle$ to which it corresponds is an eigenvector of H, and the stationary values of $\langle H \rangle$ are the eigenvalues of the Hamiltonian.

The variational method can therefore be generalized and applied to the approximate determination of the eigenvalues of the Hamiltonian H. If the function $\langle H \rangle (\alpha)$ obtained from the trial kets $|\psi(\alpha)\rangle$ has several extrema, they give the approximate values of some of its energies E_n (*cf.* exercise 10 of Complement $\mathrm{H_{XI}}$).

1-c. A special case where the trial functions form a subspace

Assume that we choose for the trial kets the set of kets belonging to a vector subspace $\mathcal{F}$ of $\mathcal{E}$. In this case, the variational method reduces to the *resolution of the eigenvalue equation of the Hamiltonian H inside* $\mathcal{F}$, and no longer in all of $\mathcal{E}$.

To see this, we simply apply the argument of § 1-b, limiting it to the kets $|\psi\rangle$ of the subspace $\mathcal{F}$. The maxima and minima of $\langle H \rangle$, characterized by $\delta \langle H \rangle = 0$, are obtained when $|\psi\rangle$ is an eigenvector of H in $\mathcal{F}$. The corresponding eigenvalues constitute the

variational method approximation for the true eigenvalues of H in $\mathcal{E}$. They also provide upper bounds of these eigenvalues: we have seen that the lowest energy that is obtained is larger than the true energy of the ground state, but it is also possible to show (*cf.* MacDonald's article in the references of this complement) that the next lowest energy is greater than the energy of the true first excited state, etc. When the dimension of $\mathcal{F}$ is increased by one unit, one obtains a new series of energies with a new energy above all the others, which themselves decrease (or possibly remain at the same value).

We stress the fact that the restriction of the eigenvalue equation of H to a subspace $\mathcal{F}$ of the state space $\mathcal{E}$ can considerably simplify its solution. However, if $\mathcal{F}$ is badly chosen, it can also yield results which are rather far from the true eigenvalues and eigenvectors of H in $\mathcal{E}$ (*cf.* § 3). The subspace $\mathcal{F}$ must therefore be chosen so as to simplify the problem enough to make it soluble, without too greatly altering the physical reality. In certain cases, it is possible to reduce the study of a complex system to that of a two-level system (*cf.* Chap. IV), or at least, to that of a system of a limited number of levels. Another important example of this procedure is the method of the *linear combination of atomic orbitals*, widely used in molecular physics. This method (*cf.* Complement G_{XI}) essentially determines the wave functions of electrons in a molecule in the form of linear combinations of eigenfunctions associated with the various atoms which constitute the molecule, treated as if they were isolated. It therefore limits the search for the molecular states to a subspace chosen using physical criteria. Similarly, in Complement F_{XI}, we shall choose as a trial wave function for an electron in a solid a linear combination of atomic orbitals relative to the various ions which constitute this solid.

Comment:

Note that first-order perturbation theory fits into this special case of the variational method: $\mathcal{F}$ is then an eigensubspace of the unperturbed Hamiltonian H_0.

2. Application to a simple example

To illustrate the discussion of § 1 and to give an idea of the validity of the approximations obtained with the help of the variational method, we shall apply this method to the one-dimensional harmonic oscillator, whose eigenvalues and eigenstates we know (*cf.* Chap. V). We shall consider the Hamiltonian:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \tag{17}$$

and we shall solve its eigenvalue equation approximately by variational calculations.

2-a. Exponential trial functions

Since the Hamiltonian (17) is even, it can easily be shown that its ground state is necessarily represented by an even wave function. To determine the characteristics of this ground state, we shall therefore choose even trial functions. We take, for example, the one-parameter family:

$$\psi_\alpha(x) = \mathrm{e}^{-\alpha x^2} \;\; ; \;\; \alpha > 0 \tag{18}$$

The square of the norm of the ket $|\psi_\alpha\rangle$ is equal to:

$$\langle\psi_\alpha \mid \psi_\alpha\rangle = \int_{-\infty}^{+\infty} \mathrm{d}x\ \mathrm{e}^{-2\alpha x^2} \tag{19}$$

and we find:

$$\begin{aligned}\langle\psi_\alpha|\, H\, |\psi_\alpha\rangle &= \int_{-\infty}^{+\infty} \mathrm{d}x\, \mathrm{e}^{-\alpha x^2} \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2\right] \mathrm{e}^{-\alpha x^2} \\ &= \left[\frac{\hbar^2}{2m}\alpha + \frac{1}{8}m\omega^2\frac{1}{\alpha}\right] \int_{-\infty}^{+\infty} \mathrm{d}x\, \mathrm{e}^{-2\alpha x^2}\end{aligned} \tag{20}$$

so that:

$$\langle H\rangle\,(\alpha) = \frac{\hbar^2}{2m}\alpha + \frac{1}{8}m\omega^2\frac{1}{\alpha} \tag{21}$$

The derivative of the function $\langle H\rangle\,(\alpha)$ goes to zero for:

$$\alpha = \alpha_0 = \frac{1}{2}\frac{m\omega}{\hbar} \tag{22}$$

and we then have:

$$\langle H\rangle\,(\alpha_0) = \frac{1}{2}\hbar\omega \tag{23}$$

The minimum value of $\langle H\rangle\,(\alpha)$ is therefore exactly equal to the energy of the ground state of the harmonic oscillator. This result is due to the simplicity of the problem that we are studying: the wave function of the ground state happens to be precisely one of the functions of the trial family (18), the one which corresponds to value (22) of the parameter α. The variational method, in this case, gives the exact solution of the problem (this illustrates the theorem proven in § 1-a).

If we want to calculate (approximately, in theory) the first excited state E_1 of the Hamiltonian (17), we should choose trial functions which are orthogonal to the wave function of the ground state. This follows from the discussion of § 1-a, which shows that the lower bound of $\langle H\rangle$ is E_1, and no longer E_0 if the coefficient c_0 is zero. We therefore choose the trial family of odd functions:

$$\psi_\alpha(x) = x\,\mathrm{e}^{-\alpha x^2} \tag{24}$$

In this case:

$$\langle\psi_\alpha \mid \psi_\alpha\rangle = \int_{-\infty}^{+\infty} \mathrm{d}x\, x^2 \mathrm{e}^{-2\alpha x^2} \tag{25}$$

and:

$$\langle\psi_\alpha|\, H\, |\psi_\alpha\rangle = \left[\frac{\hbar^2}{2m} \times 3\alpha + \frac{1}{2}m\omega^2 \times \frac{3}{4\alpha}\right] \int_{-\infty}^{+\infty} \mathrm{d}x\, x^2\, \mathrm{e}^{-2\alpha x^2} \tag{26}$$

which yields:

$$\langle H\rangle\,(\alpha) = \frac{3\hbar^2}{2m}\alpha + \frac{3}{8}m\omega^2\frac{1}{\alpha} \tag{27}$$

This function, for the same value α_0 as above [formula (22)], presents a minimum equal to:

$$\langle H \rangle (\alpha_0) = \frac{3}{2}\hbar\omega \tag{28}$$

Here again, we find exactly the energy E_1 and the associated eigenstate because the trial family includes the correct wave function.

2-b. Rational wave functions

The calculations of § 2-a enabled us to familiarize ourselves with the variational method, but they do not really allow us to judge its effectiveness as a method of approximation, since the families chosen always included the exact wave function. Therefore, we shall now choose trial functions of a totally different type, for example[1]:

$$\psi_a(x) = \frac{1}{x^2 + a} \quad ; \quad \mathrm{a} > 0 \tag{29}$$

A simple calculation then yields:

$$\langle \psi_a \mid \psi_a \rangle = \int_{-\infty}^{+\infty} \frac{dx}{(\mathrm{x}^2 + a)^2} = \frac{\pi}{2a\sqrt{a}} \tag{30}$$

and, finally:

$$\langle H \rangle (a) = \frac{\hbar^2}{4m}\frac{1}{a} + \frac{1}{2}m\omega^2 a \tag{31}$$

The minimum value of this function is obtained for:

$$a = a_0 = \frac{1}{\sqrt{2}}\frac{\hbar}{m\omega} \tag{32}$$

and is equal to:

$$\langle H \rangle (a_0) = \frac{1}{\sqrt{2}}\hbar\omega \tag{33}$$

This minimum value is therefore equal to $\sqrt{2}$ times the exact ground state energy $\hbar\omega/2$. To measure the error committed, we can calculate the ratio of $\langle H \rangle (a_0) - \hbar\omega/2$ to the energy quantum $\hbar\omega$:

$$\frac{\langle H \rangle (a_0) - \frac{1}{2}\hbar\omega}{\hbar\omega} = \frac{\sqrt{2} - 1}{2} \simeq 20\% \tag{34}$$

3. Discussion

The example of § 2-b shows that it is easy to obtain the ground state energy of a system, without significant error, starting with arbitrarily chosen trial kets. This is one of the

[1] Our choice here is dictated by the fact that we want the necessary integrals to be analytically calculable. Of course, in most real cases, one resorts to numerical integration.

principal advantages of the variational method. Since the exact eigenvalue is a minimum of the average value $\langle H \rangle$, it is not surprising that $\langle H \rangle$ does not vary very much near this minimum.

On the other hand, as the same reasoning shows, the "approximate" state can be rather different from the true eigenstate. Thus, in the example of § 2-b, the wave function $1/(x^2 + a_0)$ [where a_0 is given by formula (32)] decreases too rapidly for small values of x and much too slowly when x becomes large. Table I gives quantitative support for this qualitative assertion. It gives, for various values of x^2, the values of the exact normalized eigenfunction:

$$\varphi_0(x) = (2\alpha_0/\pi)^{1/4} \mathrm{e}^{-\alpha_0 x^2}$$

[where α_0 was defined in (22)] and of the approximate normalized eigenfunction:

$$\sqrt{\frac{2}{\pi}} (a_0)^{3/4} \psi_{a_0}(x) = \sqrt{\frac{2}{\pi}} \frac{(a_0)^{3/4}}{x^2 + a_0} = \sqrt{\frac{2}{\pi}} \left(2\sqrt{2}\alpha_0\right)^{1/4} \frac{1}{1 + 2\sqrt{2}\alpha_0 x^2} \tag{35}$$

$x\sqrt{\alpha_0}$	$\left(\frac{2}{\pi}\right)^{1/4} \mathrm{e}^{-\alpha_0 x^2}$	$\sqrt{\frac{2}{\pi}} \frac{\left(2\sqrt{2}\right)^{1/4}}{1+2\sqrt{2}\alpha_0 x^2}$
0	0.893	1.034
1/2	0.696	0.605
1	0.329	0.270
3/2	0.094	0.140
2	0.016	0.083
5/2	0.002	0.055
3	0.000 1	0.039

Table I

It is therefore necessary to be very careful when physical properties other than the energy of the system are calculated using the approximate state obtained from the variational method. The validity of the result obtained varies enormously depending on the physical quantity under consideration. In the particular problem which we are studying here, we find, for example, that the approximate average value[2] of the operator X^2 is not very different from the exact value:

$$\frac{\langle \psi_{a_0} | X^2 | \psi_{a_0} \rangle}{\langle \psi_{a_0} | \psi_{a_0} \rangle} = \frac{1}{\sqrt{2}} \frac{\hbar}{m\omega} \tag{36}$$

which is to be compared with $\hbar/2m\omega$. On the other hand, the average value of X^4 is infinite for the wave function (35), while it is, of course, finite for the real wave function.

[2]The average value of X is automatically zero, as is correct since we have chosen even trial functions.

More generally, Table I shows that the approximation will be very poor for all properties that depend strongly on the behavior of the wave function for $x \gtrsim 2/\sqrt{\alpha_0}$.

The drawback we have just mentioned is all the more serious as it is very difficult, if not impossible, to evaluate the error in a variational calculation if we do not know the exact solution of the problem (and, of course, if we use the variational method, it is because we do not know this exact solution).

The variational method is therefore a very flexible approximation method, which can be adapted to very diverse situations and which gives great scope to physical intuition in the choice of trial kets. It gives good values for the energy rather easily, but the approximate state vectors may present certain completely unpredictable erroneous features, and we cannot check these errors. This method is particularly valuable when physical arguments give us an idea of the qualitative or semi-quantitative form of the solutions.

References and suggestions for further reading:

The Hartree-Fock method, often used in physics, is an application of the variational method. See references of Chapter XI and Complement E_{XV} of Volume III.

The variational method is of fundamental importance in molecular physics. See references of Complement G_{XI}.

For a simple presentation of the use of variational principles in physics, see Feynman II (7.2), Chap. 19.

J.K.L. MacDonald, *Physical Review* vol. 143, pages 830 à 833 (1933).

Complement F_{XI}

Energy bands of electrons in solids: a simple model

1 A first approach to the problem: qualitative discussion . . 1178

2 A more precise study using a simple model 1181

2-a Calculation of the energies and stationary states 1181

2-b Discussion 1187

A crystal is composed of atoms evenly distributed in space so as to form a three-dimensional periodic lattice. The theoretical study of the properties of a crystal, which brings into play an extremely large number of particles (nuclei and electrons), poses a problem which is so complicated that it is out of the question to treat it rigorously. We must therefore resort to approximations.

The first of these is of the same type as the Born-Oppenheimer approximation (which we encountered in § 1 of Complement A_V). It consists of considering, first of all, the positions of the nuclei as fixed, which enables us to study the stationary states of the electrons subjected to the potential created by the nuclei. The motion of the nuclei is not treated until later, using the knowledge of the electronic energies[1]. In this complement, we shall concern ourselves only with the first step of this calculation, and we shall assume the nuclei to be motionless at the nodes of the crystalline lattice.

This problem still remains extremely complicated. It is necessary to calculate the energies of a system of electrons subjected to a periodic potential and interacting with each other. We then make a second approximation: we assume that each electron, at a position $\mathbf{r}_i$, is subjected to the influence of a potential $V(\mathbf{r}_i)$ which takes into account the attraction exerted by the nuclei and the average effect of the repulsion of all the other electrons[2]. The problem is thus reduced to one involving independent particles, moving in a potential that has the periodicity of the crystalline lattice.

The physical characteristics of a crystal therefore depend, in a first approximation, on the behavior of independent electrons subjected to a periodic potential. We could be led to think that each electron remains bound to a given nucleus, as happens in isolated atoms. We shall see that, in reality, the situation is completely different. Even if an electron is initially in the neighborhood of a particular nucleus, it can move into the zone of attraction of an adjacent nucleus by the tunnel effect, then into another, and so on. Actually, the stationary states of the electrons are not localized in the neighborhood of any nucleus, but are completely delocalized: the probability density associated with them is uniformly distributed over all the nuclei[3]. Thus, the properties of an electron

[1] Recall that the study of the motion of the nuclei leads to the introduction of the normal vibrational modes of the crystal: the phonons (*cf.* Complement J_V).

[2] This approximation is of the same type as the "central field" approximation for isolated atoms (*cf.* Complement A_{XIV}, § 1).

[3] This phenomenon is analogous to the one we encountered in the study of the ammonia molecule (*cf.* Complement G_{IV}). There, since the nitrogen atom can move from one side of the plane of the hydrogen atoms to the other, by the tunnel effect, the stationary states give an equal probability of finding it in each of the two corresponding positions.

placed in a periodic potential resemble those of an electron free to move throughout the crystal more than they do those of an electron bound to a particular atom. Such a phenomenon could not exist in classical mechanics: the direction of a particle traveling through a crystal would change constantly under the influence of the potential variations (for example, upon skirting an ion). In quantum mechanics, the interference of the waves scattered by the different nuclei permit the propagation of an electron inside the crystal.

In § 1, we shall study very qualitatively how the energy levels of isolated atoms are modified when they are brought gradually closer together to form a linear chain. Then, in § 2, still confining ourselves, for simplicity, to the case of a linear chain, we shall calculate the energies and wave functions of stationary states a little more precisely. We shall perform the calculation in the "tight bonding approximation": when the electron is in one site, it can move to one of two neighboring sites via the tunnel effect. The tight bonding approximation is equivalent to assuming that the probability of its tunneling is small. We shall, in this way, establish a certain number of results (the delocalization of stationary states, the appearance of allowed and forbidden energy bands, the form of Bloch functions) which remain valid in more realistic models (three-dimensional crystals, bonds of arbitrary strength).

The "perturbation" approach that we shall adopt here constructs the stationary states of the electrons from atomic wave functions localized about the various ions. It has the advantage of showing how atomic levels change gradually to energy bands in a solid. Note, however, that the existence of energy bands can be directly established from the periodic nature of the structure in which the electron is placed (see, for example, Complement O_{III}, in which we study quantization of the energy levels in a one-dimensional periodic potential).

Finally, we stress the fact that we are concerned here only with the properties of the individual stationary states of the electrons. To construct the stationary state of a system of N electrons from these individual states, it is necessary to apply the symmetrization postulate (*cf.* Chap. XIV), since we are dealing with a system of identical particles. We shall treat this problem again in Complement C_{XIV}, when we shall describe the spectacular consequences of Pauli's exclusion principle on the physical behavior of the electrons in a solid. Many other examples of the effects of the symmetrization will be discussed in Chapters XV to XVII.

1. A first approach to the problem: qualitative discussion

Let us go back to the example of the ionized H_2^+ molecule, studied in §§ C-2-c and C-3-d of Chapter IV. Consider, therefore, two protons P_1 and P_2, whose positions are fixed, and an electron which is subject to their electrostatic attraction. This electron sees a potential $V(\mathbf{r})$, which has the form indicated in Figure 1. In terms of the distance R between P_1 and P_2 (considered as a parameter) what are the possible energies and the corresponding stationary states?

We shall begin by considering the limiting case in which $R \gg a_0$ (where a_0 is the Bohr radius of the hydrogen atom). The ground state is then two-fold degenerate: the electron can form a hydrogen atom either with P_1 or with P_2; it is practically unaffected by the attraction of the other proton, which is very far away. In other words, the coupling between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ considered in Chapter IV (localized states in the neighborhood of P_1 or P_2; *cf.* Fig. 13 of Chapter IV) is then negligible, so that $|\varphi_1\rangle$

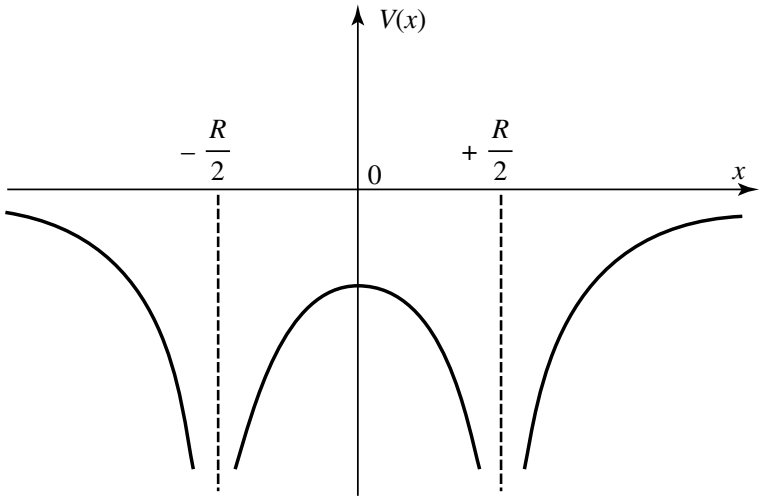


Figure 1: The potential seen by the electron as it moves along the Ox axis defined by the two protons in the ionized H_2^+ molecule. We obtain two wells separated by a barrier. If, at any instant, the electron is localized in one of the two wells, it can move into the other well via the tunnel effect.

and $|\varphi_2\rangle$ are practically stationary states.

If we now choose a value of R comparable to a_0 it is no longer possible to neglect the attraction of one or the other of the protons. If, at $t = 0$, the electron is localized in the neighborhood of one of them, and even if its energy is lower than the height of the potential barrier situated between P_1 and P_2 (*cf.* Fig. 1), it can move to the other proton by the tunnel effect. In Chapter IV we studied the effect of coupling of the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, and we showed that it produces an oscillation of the system between these two states (the dynamical aspect). We have also seen (the static aspect) that this coupling removes the degeneracy of the ground state and that the corresponding stationary states are "delocalized" (for these states, the probability of finding the electron in the neighborhood of P_1 or P_2 is the same). Figure 2 shows the form of the variation with respect to R of the possible energies of the system[4].

Two effects appear when we decrease the distance R between P_1 and P_2. On the one hand, an $R = \infty$ energy value gives rise to two distinct energies when R decreases (when the distance R is fixed at a given value R_0, the stronger the coupling between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, the greater the difference between these two energies). On the other hand, the stationary states are delocalized.

It is easy to imagine what will happen if the electron is subject to the influence, not of two, but of three identical attractive particles (protons or positive ions), arranged, for example, in a straight line at intervals of R. When R is very large, the energy levels are triply degenerate, and the stationary states of the electron can be chosen to be localized in the neighborhood of any one of the fixed particles.

If R is decreased, each energy gives rise to three generally distinct energies and, in a

[4] A detailed study of the H_2^+ ion is presented in Complement G_{XI}.

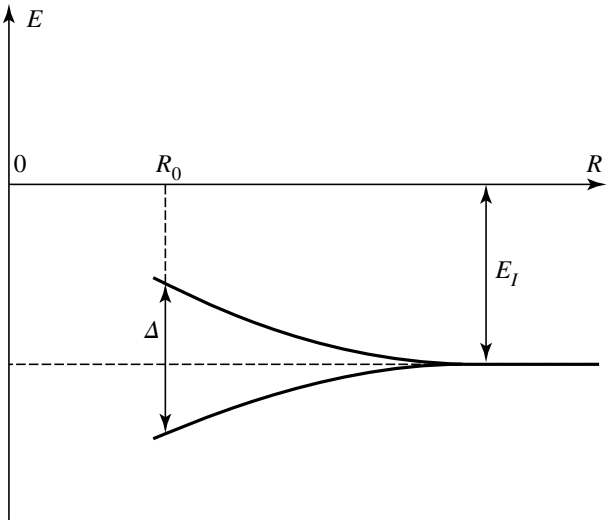


Figure 2: Variation of the energy of stationary states of the electron in terms of the distance R between the two protons of the H_2^+ ion. When R is large, there are two practically degenerate states, of energy $-E_I$. When R decreases, this degeneracy is removed. The smaller R, the greater the splitting.

stationary state, the probabilities of finding the electron in the three wells are comparable. Moreover, if, at the initial instant, the electron is localized in the right-hand well, for example, it moves into the other wells during its subsequent evolution[5].

The same ideas remain valid for a chain composed of an arbitrary number $\mathcal{N}$ of ions which attract an electron. The potential seen by the electron is then composed of $\mathcal{N}$ regularly spaced identical wells (in the limit in which $\mathcal{N} \to \infty$, it is a periodic potential). When the distance R between the ions is large, the energy levels are $\mathcal{N}$-fold degenerate. This degeneracy disappears if the ions are moved closer together: each level gives rise to distinct levels, which are distributed, as shown in Figure 3, in an energy interval of width Δ. What now happens if the value of $\mathcal{N}$ is very large? In each of the intervals Δ, the possible energies are so close that they practically form a continuum: "allowed energy bands" are thus obtained, separated by "forbidden bands". Each allowed band contains $\mathcal{N}$ levels (actually $2\mathcal{N}$ if the electron spin is taken into account). The stronger the coupling causing the electron to pass from one potential well to the next one, the greater the band width. (Consequently, we expect the lowest energy bands to be the narrowest since the tunnel effect which is responsible for this passage is less probable when the energy is smaller). The stationary states of the electron are all delocalized. The analogue here of Figure 3 of Complement M_{III} is Figure 4, which represents the energy levels and gives an idea of the spatial extension of the associated wave functions.

Finally, note that if, at $t = 0$, the electron is localized at one end of the chain, it propagates along the chain during its subsequent evolution.

[5]See exercise 8 of Complement J_{IV}.

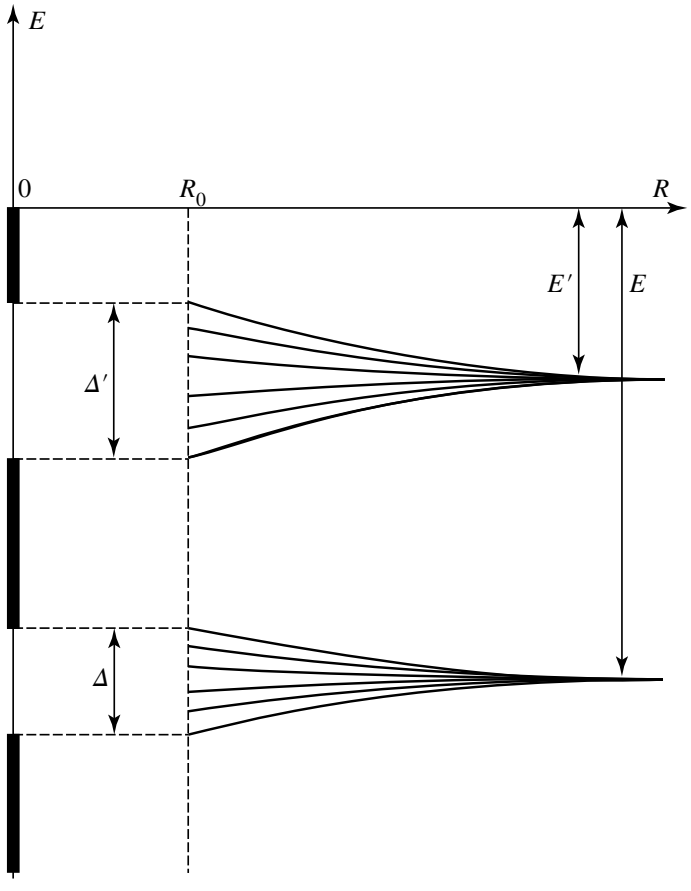


Figure 3: Energy levels of an electron subject to the action of $\mathcal{N}$ regularly spaced identical ions. When R is very large, the wave functions are localized about the various ions, and the energy levels are the atomic levels, $\mathcal{N}$-fold degenerate (the electron can form an atom with any one of the $\mathcal{N}$ ions). In the figure, two of these levels are shown, of energies $-E$ and $-E'$. When R decreases, the electron can pass from one ion to another by the tunnel effect, and the degeneracy of the levels is removed. The smaller R, the greater the splitting. For the value R_0 of R found in a crystal, each of the two original atomic levels is therefore broken down into $\mathcal{N}$ very close levels. If $\mathcal{N}$ is very large, these levels are so close that they yield energy bands, of widths Δ and Δ', separated by a forbidden band.

2. A more precise study using a simple model

2-a. Calculation of the energies and stationary states

To complete the qualitative considerations of the preceding section, we shall discuss the problem more precisely, using a simple model. We shall perform calculations analogous to those of § C of Chapter IV, but adapted to the case in which the system under consideration contains an infinite number of ions (instead of two), regularly spaced

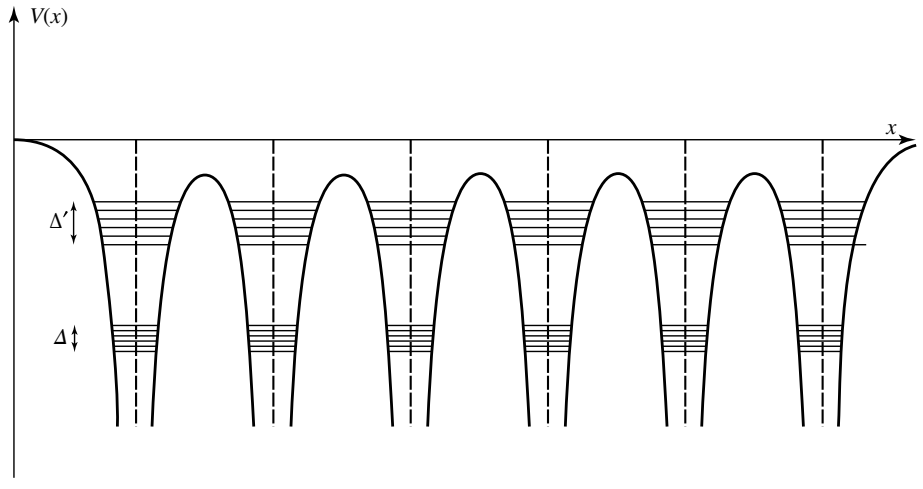


Figure 4: Energy levels for a potential composed of several regularly spaced wells. Two bands are shown in this figure, one of width Δ and the other of width Δ'. The deeper the band, the more narrow it is, since crossing the barrier by the tunnel effect is then more difficult.

in a linear chain.

α. Description of the model; simplifying hypotheses

Consider, therefore, an infinite linear chain of regularly spaced positive ions. As in Chapter IV, we shall assume that the electron, when it is bound to a given ion, has only one possible state: we shall denote by $|v_n\rangle$ the state of the electron when it forms an atom with the nth ion of the chain. For the sake of simplicity, we shall neglect the mutual overlap of the wave functions $v_n(x)$ associated with neighboring atoms, and we shall assume the $\{|v_n\rangle\}$ basis to be orthonormal:

$$\langle v_n|v_p\rangle = \delta_{np} \tag{1}$$

Moreover, we shall confine ourselves to the subspace of the state space spanned by the kets $|v_n\rangle$. It is obvious that by restricting the state space accessible to the electron in this way, we are making an approximation. This can be justified by using the variational method (*cf.* Complement E_{XI}): by diagonalizing the Hamiltonian H, not in the total space, but in the one spanned by the $|v_n\rangle$, it can be shown that we obtain a good approximation for the true energies of the electron.

We shall now write the matrix representing the Hamiltonian H in the $\{|v_n\rangle\}$ basis. Since the ions all play equivalent roles, the matrix elements $\langle v_n| H |v_n\rangle$ are necessarily all equal to the same energy E_0. In addition to these diagonal elements, H also has non-diagonal elements $\langle v_n| H |v_p\rangle$ (coupling between the various states $|v_n\rangle$, which expresses the possibility for an electron to move from one ion to another). This coupling is obviously very weak for distant ions; this is why we shall take into account only the matrix elements $\langle v_n| H |v_{n\pm1}\rangle$, which we shall choose equal to a real constant $-A$. Under these conditions,

the (infinite) matrix that represents H can be written:

$$(H) = \begin{pmatrix} \ddots & & & & & \\ & E_0 & -A & 0 & 0 & \\ & -A & E_0 & -A & 0 & \\ & 0 & -A & E_0 & -A & \\ & 0 & 0 & -A & E_0 & \\ & & & & & \ddots \end{pmatrix} \tag{2}$$

To find the possible energies and the corresponding stationary states, we must diagonalize this matrix.

β. Possible energies; the concept of an energy band

Let $|\varphi\rangle$ be an eigenvector of H; we shall write it in the form:

$$|\varphi\rangle = \sum_{q=-\infty}^{+\infty} c_q \, |v_q\rangle \tag{3}$$

Using (2), the eigenvalue equation:

$$H \, |\varphi\rangle = E \, |\varphi\rangle \tag{4}$$

projected onto $|v_q\rangle$, yields:

$$E_0 c_q - A c_{q+1} - A \, c_{q-1} = E c_q \tag{5}$$

When q takes on all positive or negative integral values, we obtain an infinite system of coupled linear equations which, in certain ways, recall the coupled equations (5) of Complement $\mathrm{J_V}$. As in that complement, we shall look for simple solutions of the form:

$$c_q = \mathrm{e}^{ikql} \tag{6}$$

where l is the distance between two adjacent ions, and k is a constant whose dimensions are those of an inverse length. We require k to belong to the "first Brillouin zone", that is, to satisfy:

$$-\frac{\pi}{l} \leqslant k < +\frac{\pi}{l} \tag{7}$$

This is always possible, because two values of k differing by $2\pi/l$ give all the coefficients c_q the same value. Substituting (6) into (5), we obtain:

$$E_0 \, \mathrm{e}^{ikql} - A \left[\mathrm{e}^{ik(q+1)l} + \mathrm{e}^{ik(q-1)l} \right] = E \, \mathrm{e}^{ikql} \tag{8}$$

that is, dividing by e^{ikql}:

$$E = E(k) = E_0 - 2A \cos kl \tag{9}$$

If this condition is satisfied, the ket $|\varphi\rangle$ given by (3) and (6) is an eigenket of H; its energy depends on the parameter k, as is indicated by (9).

Figure 5 represents the variation of E with respect to k. It shows that the possible energies are situated in the interval $[E_0 - 2A, E_0 + 2A]$. We therefore obtain an allowed energy band, whose width $4A$ is proportional to the strength of the coupling.

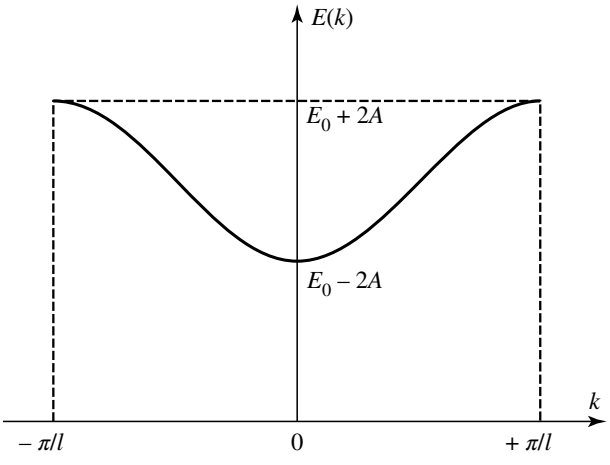


Figure 5: Possible energies of the electron in terms of the parameter k (k varies within the first Brillouin zone). An energy band therefore appears, with a width 4A which is proportional to the coupling between neighboring atoms.

γ. Stationary states; Bloch functions

Let us calculate the wave function $\varphi_k(x) = \langle x \mid \varphi_k \rangle$ associated with the stationary state $|\varphi_k\rangle$ of energy $E(k)$. Relations (3) and (6) lead to:

$$|\varphi_k\rangle = \sum_{q=-\infty}^{+\infty} \mathrm{e}^{ikql} |\upsilon_q\rangle \tag{10a}$$

that is:

$$\varphi_k(x) = \sum_{q=-\infty}^{+\infty} \mathrm{e}^{ikql} \upsilon_q(x) \tag{10b}$$

where:

$$\upsilon_q(x) = \langle x \mid \upsilon_q \rangle \tag{11}$$

is the wave function associated with the state $|\upsilon_q\rangle$. Since the state $|\upsilon_q\rangle$ can be obtained from the state $|\upsilon_0\rangle$ by a translation of ql we have:

$$\upsilon_q(x) = \upsilon_0(x - ql) \tag{12}$$

so that (10b) can be written:

$$\varphi_k(x) = \sum_{q=-\infty}^{+\infty} \mathrm{e}^{ikql} \upsilon_0(x - ql) \tag{13}$$

We now calculate $\varphi_k(x+l)$:

$$\begin{aligned}\varphi_k(x+l) &= \sum_{q=-\infty}^{+\infty} \mathrm{e}^{ikql} v_0[x-(q-1)l] \\ &= \mathrm{e}^{ikl} \sum_{q=-\infty}^{+\infty} \mathrm{e}^{ik(q-1)l} v_0[x-(q-1)l] \\ &= \mathrm{e}^{ikl}\varphi_k(x) \end{aligned} \tag{14}$$

To express this remarkable property simply, we set:

$$\varphi_k(x) = \mathrm{e}^{ikx} u_k(x) \tag{15}$$

The function $u_k(x)$ so defined then satisfies:

$$u_k(x+l) = u_k(x) \tag{16}$$

Therefore, the wave function $\varphi_k(x)$ is the product of e^{ikx} and a periodic function which has the period l of the lattice. A function of type (15) is called a *Bloch function.* Note that, if n is any integer:

$$|\varphi_k(x+nl)|^2 = |\varphi_k(x)|^2 \tag{17}$$

a result which demonstrates the delocalization of the electron: the probability density of finding the electron at any point on the x-axis is a periodic function of x.

Comment:

Expressions (15) and (16) have been proven here for a simple model. Actually, this result is more general and can be proven directly from the symmetries of the Hamiltonian H (Bloch's theorem). To show this, let us call $S(a)$ the unitary operator associated with a translation a along Ox (*cf.* Complement $\mathrm{E_{II}}$, § 3). Since the system is invariant under any translation that leaves the ion chain unchanged, we must have:

$$[H, S(l)] = 0 \tag{18a}$$

We can therefore construct a basis of eigenvectors common to the operator $S(l)$ and H. Now, equation (14) is simply the one that defines the eigenfunctions of $S(-l)$ [since this operator is unitary, its eigenvalues can always be written in the form e^{ikl}, where k satisfies condition (7); *cf.* Complement $\mathrm{C_{II}}$, § 1-c]. It is then simple to get, as before, (15) and (16) from (14).

Note that, for any a, we have, in general:

$$[H, S(a)] \neq 0 \tag{18b}$$

unlike the situation of a free particle (or one subject to the influence of a constant potential). For a free particle, since H commutes with all operators $S(a)$ (that is, with the momentum P_x; *cf.* Complement $\mathrm{E_{II}}$, § 3), the stationary wave functions are of the form:

$$w_k(x) \propto \mathrm{e}^{ikx} \tag{19}$$

This means that the function $u_k(x)$ appearing in (15) is necessarily a constant, which is a more restrictive condition. In the problem studied in this complement, the commutator of relation (18b) vanishes only for certain values of a, which implies less restrictive conditions for the wave function.

δ. Periodic boundary conditions

To each value of k in the interval $[-\pi/l, +\pi/l]$ corresponds an eigenstate $|\varphi\rangle$ of H, with the coefficients c_q appearing in expansion (3) of $|\varphi\rangle$ given by equation (6). We thus obtain an infinite continuum of stationary states. This is due to the fact that we have considered a linear chain containing an infinite number of ions. What happens when we consider a finite linear chain, of length L, composed of a large number $\mathcal{N}$ of ions?

The qualitative considerations of § 1 show that there must then be $\mathcal{N}$ levels in the band ($2\,\mathcal{N}$ if spin is taken into account). The exact determination of the corresponding $\mathcal{N}$ stationary states is a difficult problem, since it is necessary to take into account the boundary conditions at the ends of the chain. It is clear, however, that the behavior of electrons sufficiently far from the ends are little affected by the "edge effects"[6]. This is why one generally prefers, in solid state physics, to substitute new boundary conditions for the real boundary conditions; despite their artificial character, these new conditions lead to much simpler calculations, while conserving the most important properties necessary for the comprehension of the physical effects (other than the edge effects).

These new boundary conditions, called periodic boundary conditions, or "Born-Von Karman conditions" (B.V.K. conditions), require the wave function to take on the same value at both ends of the chain. We can also imagine that we are placing an infinite number of identical chains, all of length L, end to end. We then require the wave function of the electron to be periodic, with a period L. Equations (5) remain valid, as does their solution (6), but the periodicity of the wave function now implies:

$$\mathrm{e}^{ikL} = 1 \tag{20}$$

Consequently, the only possible values of k are of the form:

$$k_n = n\frac{2\pi}{L} \tag{21}$$

where n is a positive or negative integer or zero. Let us now verify that the B.V.K. conditions give the correct result for the number of stationary states contained in the band. To do so, we must calculate the number of allowed values k_n included in the first Brillouin zone defined in (7). We obtain this number by dividing the width $2\pi/l$ of this zone by the interval $2\pi/L$ between two adjacent values of k, which indeed gives us:

$$\frac{2\pi}{l}\Big/\frac{2\pi}{L} = \frac{L}{l} = N - 1 \simeq N \tag{22}$$

We should also show that the $\mathcal{N}$ stationary states obtained with the B.V.K. conditions are distributed in the allowed band with the same density[7] $\rho(E)$ as the true stationary states (associated with the real boundary conditions). As the density of states $\rho(E)$ plays a very important role in the comprehension of the physical properties of a solid (we shall discuss this point in Complement C_{XIV}), it is important for the new boundary conditions to leave it unchanged. That the B.V.K. conditions give the correct density of states will be proven in Complement C_{XIV} (§ 1-c) for the simple example of a free electron gas enclosed in a "rigid box". In this case, the true stationary states can be calculated and compared with those obtained by using the periodic boundary conditions on the walls of the box (see also § 3-a of Complement O_{III}).

[6] For a three-dimensional crystal, this amounts to establishing a distinction between "bulk effects" and "surface effects".

[7] $\rho(E)\,\mathrm{d}E$ is the number of distinct stationary states with energies included between E and $E + \mathrm{d}E$.

2-b. Discussion

Starting with a discrete non-degenerate level for an isolated atom (for example, the ground level) we have obtained a series of possible energies, grouped in an allowed band of width $4A$ for the chain of ions being considered. If we had started with another level of the atom (for example, the first excited level), we would have obtained another energy band, and so on. Each atomic level yields one energy band, as Figure 6 shows, and there appears a series of allowed bands, separated by forbidden bands.

Relation (6) shows that, for a stationary state, the probability amplitude of finding the electron in the state $|\upsilon_q\rangle$ is an oscillating function of q, whose modulus does not depend on q. This recalls the properties of phonons, the normal vibrational modes of an infinite number of coupled oscillators for which all the oscillators participate in the collective vibration with the same amplitude, but with a certain phase shift (*cf.* Complement $\mathrm{J_V}$).

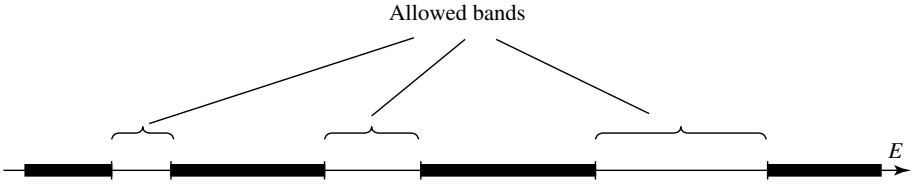


Figure 6: Allowed bands and forbidden bands on the energy axis.

How can we obtain states in which the electron is not completely delocalized? For a free electron, we saw in Chapter I that we must superpose plane waves so as to form a free "wave packet ":

$$\hat{\psi}\,(x,t) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}k\, \hat{g}\,(k)\, \mathrm{e}^{i[kx - E(k)t/\hbar]} \tag{23}$$

The maximum of this wave packet propagates at the group velocity (*cf.* Chap. I, § C):

$$\hat{V}_G = \frac{1}{\hbar} \left[\frac{\mathrm{d}E}{\mathrm{d}k}\right]_{k=k_0} = \frac{\hbar k_0}{m} \tag{24}$$

[where k_0 is the value of k for which the function $\hat{g}\,(k)$ presents a peak]. Here, we must superpose wave functions of type (15), and the corresponding ket can be written:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}k\, g(k) \mathrm{e}^{-iE(k)t/\hbar} |\varphi_k\rangle \tag{25}$$

where $g(k)$ is a function of k which has the form of a peak about $k = k_0$. We shall calculate the probability amplitude of finding the electron in the state $|\upsilon_q\rangle$. Using (10a) and (1), we can write:

$$\langle \upsilon_q \mid \psi\,(t)\rangle = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}k\, g(k)\, \mathrm{e}^{i[kql - E(k)t/\hbar]} \tag{26}$$

Replacing ql by x in this relation, we obtain a function of x:

$$\chi(x,t) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}k\, g(k)\, \mathrm{e}^{i[kx - E(k)t/\hbar]} \tag{27}$$

Only the values at the points $x = 0$, $\pm ql$, $\pm ql$, etc... of this function are really significant and yield the desired probability amplitudes.

Relation (27) is entirely analogous to (23). By applying (24), it can be shown that $\chi(x,t)$ takes on significant values only in a limited domain of the x-axis whose center moves at the velocity:

$$V_G = \frac{1}{\hbar}\left[\frac{\mathrm{d}E(k)}{\mathrm{d}k}\right]_{k=k_0} \tag{28}$$

It follows that the probability amplitude $\langle v_q \mid \psi(t)\rangle$ is large only for certain values of q: therefore, the electron is no longer delocalized, but moves in the crystal at the velocity V_G given by (28).

Equation (9) enables us to calculate this velocity explicitly:

$$V_G = \frac{2Al}{\hbar} \sin k_0 l \tag{29}$$

This function is shown in Figure 7. It is zero when $k_0 = 0$, that is, when the energy is

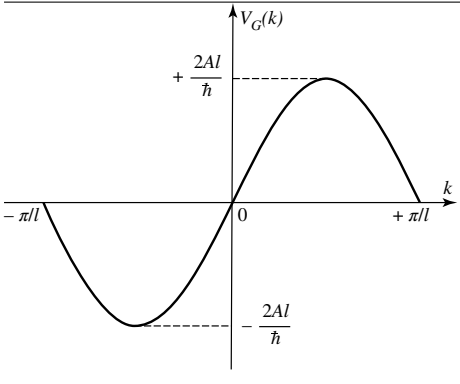


Figure 7: Group velocity of the electron as a function of the parameter k. This velocity goes to zero, not only for $k = 0$ (as for the free electron), but also for $k = \pm\pi/l$ (the edges of the first Brillouin zone).

minimal; this is also a property of the free electron. However, when k_0 takes on non-zero values, important departures from the behavior of a free electron occur. For example, as soon as $k_0 > \pi/2l$, the group velocity is no longer an increasing function of the energy. It even goes to zero when $k_0 = \pm\pi/l$ (at the borders of the first Brillouin zone). This indicates that an electron cannot move in the crystal if its energy is too close to the maximum value $E_0 + 2A$ appearing in Figure 5. The optical analogy of this situation is Bragg reflection. X rays whose wavelength is equal to the unit edge of the crystalline lattice cannot propagate in it: interference of the waves scattered by each of the ions lead to total reflection.

References and suggestions for further reading:

Feynman III (1.2), Chap. 13; Mott and Jones (13.7), Chap. II, § 4; references of section 13 of the bibliography.

Complement G_{XI}

A simple example of the chemical bond: the H_2^+ ion

1 Introduction 1189
1-a General method 1189
1-b Notation 1190
1-c Principle of the exact calculation 1192
2 The variational calculation of the energies 1192
2-a Choice of the trial kets 1192
2-b The eigenvalue equation of the Hamiltonian H in the trial ket vector subspace $\mathcal{F}$ 1194
2-c Overlap, Coulomb and resonance integrals 1195
2-d Bonding and antibonding states 1197
3 Critique of the preceding model. Possible improvements . 1201
3-a Results for small R 1201
3-b Results for large R 1202
4 Other molecular orbitals of the H_2^+ ion 1204
4-a Symmetries and quantum numbers. Spectroscopic notation . 1205
4-b Molecular orbitals constructed from the $2p$ atomic orbitals . . 1206
5 The origin of the chemical bond; the virial theorem 1210
5-a Statement of the problem 1210
5-b Some useful theorems 1211
5-c The virial theorem applied to molecules 1212
5-d Discussion 1215

1. Introduction

In this complement, we intend to show how quantum mechanics enables us to understand the existence and properties of the *chemical bond*, which is responsible for the formation of various molecules from isolated atoms. Our aim is to explain the basic nature of these phenomena and not, of course, to enter into details which could only be covered in a specialized book on molecular physics. This is why we shall study the simplest existing molecule, the H_2^+ ion, which is composed of two protons and a single electron. We have already discussed certain aspects of this problem, in Chapter IV (§ C-2-c) and in exercise 5 of Complement K_I; we shall consider it here in a more realistic and systematic fashion.

1-a. General method

When the two protons are very far from each other, the electron forms a hydrogen atom with one of them, and the other one remains isolated, in the form of an H^+ ion. If the two protons are brought closer together, the electron will be able to "jump" from

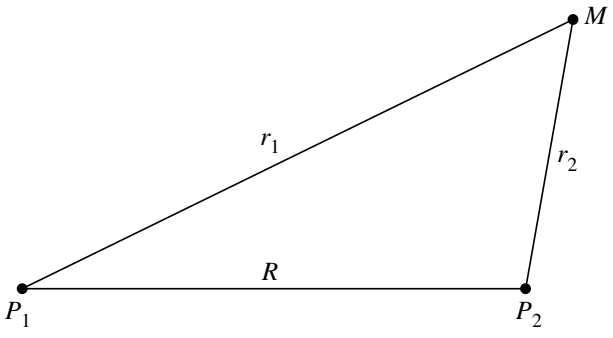


Figure 1: We call r_1 the distance between the electron (M) and proton P_1, r_2 the distance between the electron and proton P_2, and R the internuclear distance P_1P_2.

one to the other. This radically modifies the situation (*cf.* Chap. IV, § C-2). We shall therefore study the variation of the energies of the stationary states of the system with respect to the distance between the two protons. We shall see that the energy of the ground state reaches a minimum for a certain value of this distance, which explains the stability of the H_2^+ molecule.

In order to treat the problem exactly, it would be necessary to write the Hamiltonian of the three-particle system and solve its eigenvalue equation. However, it is possible to simplify this problem considerably by using the *Born-Oppenheimer approximation* (*cf.* Complement A_V, § 1-a). Since the motion of the electron in the molecule is considerably more rapid than that of the protons, the latter can be neglected in a first approximation. The problem is then reduced to the resolution of the eigenvalue equation of the Hamiltonian of the electron subject to the attraction of two protons which are assumed to be fixed. In other words, the distance R between the two protons is treated, not like a quantum mechanical variable, but like a *parameter,* on which the electronic Hamiltonian and total energy of the system depend.

In the case of the H_2^+ ion, it so happens that the equation simplified in this way is exactly soluble for all values of R. However, this is not true for other, more complex, molecules. The *variational method*, described in Complement E_{XI}, must then be used. Although we are confining ourselves here to the study of the H_2^+ ion, we shall use the variational method, since it can be generalized to the case of other molecules.

1-b. Notation

We shall call R the distance between the two protons, situated at P_1 and P_2, and r_1 and r_2 the distances of the electron to each of the two protons (Fig. 1). We shall relate these distances to a natural atomic unit, the Bohr radius a_0 (*cf.* Chap. VII, § C-2), by setting:

$$\rho = R/a_0$$
$$\rho_1 = r_1/a_0 \qquad \rho_2 = r_2/a_0 \tag{1}$$

The normalized wave function associated with the ground state $1s$ of the hydrogen

atom formed around proton P_1 can be written:

$$\varphi_1 = \frac{1}{\sqrt{\pi a_0^3}} \mathrm{e}^{-\rho_1} \tag{2}$$

Similarly, we express the energies in terms of the natural unit $E_I = e^2/2a_0$; E_I is the ionization energy of the hydrogen atom.

It will sometimes be convenient in what follows to use a system of elliptic coordinates, in which a point M of space (here, the electron) is defined by:

$$\begin{aligned} \mu &= \frac{r_1 + r_2}{R} = \frac{\rho_1 + \rho_2}{\rho} \\ \nu &= \frac{r_1 - r_2}{R} = \frac{\rho_1 - \rho_2}{\rho} \end{aligned} \tag{3}$$

and the angle φ which fixes the orientation of the MP_1P_2 plane about the P_1P_2 axis (this angle also enters into the system of polar coordinates whose Oz axis coincides with P_1P_2). If we fix μ and ν, and if φ varies between 0 and 2π, the point M describes a circle about the P_1P_2 axis. If μ (or ν) and φ are fixed, M describes an ellipse (or a hyperbola) of foci P_1 and P_2 when ν (or μ) varies. It can easily be shown that the volume element in this coordinate system is:

$$\mathrm{d}^3 r = \frac{R^3}{8}(\mu^2 - \nu^2)\ \mathrm{d}\mu\ \mathrm{d}\nu\ \mathrm{d}\varphi \tag{4}$$

To do so, we simply calculate the Jacobian J of the transformation:

$$\{x, y, z\} \Longrightarrow \{\mu, \nu, \varphi\} \tag{5}$$

We see immediately that, if P_1P_2 is chosen as the Oz axis, with the origin O in the middle of P_1P_2:

$$\begin{aligned} r_1^2 &= x^2 + y^2 + \left(z - \frac{R}{2}\right)^2 \\ r_2^2 &= x^2 + y^2 + \left(z + \frac{R}{2}\right)^2 \\ \mathrm{tg}\,\varphi &= \frac{y}{x} \end{aligned} \tag{6}$$

We can then find:

$$\begin{aligned} &\frac{\partial \mu}{\partial x} = \frac{1}{R}\left(\frac{\partial r_1}{\partial x} + \frac{\partial r_2}{\partial x}\right) = \frac{1}{R}\left(\frac{x}{r_1} + \frac{x}{r_2}\right) = \frac{\mu x}{r_1 r_2} \\ &\frac{\partial \nu}{\partial x} = \frac{1}{R}\left(\frac{\partial r_1}{\partial x} - \frac{\partial r_2}{\partial x}\right) = -\frac{\nu x}{r_1 r_2} \\ &\frac{\partial \mu}{\partial y} = \frac{\mu y}{r_1 r_2} \\ &\frac{\partial \nu}{\partial y} = -\frac{\nu y}{r_1 r_2} \\ &\frac{\partial \mu}{\partial z} = \frac{1}{R}\left[\frac{z - R/2}{r_1} + \frac{z + R/2}{r_2}\right] = \frac{\mu z + \nu R/2}{r_1 r_2} \\ &\frac{\partial \nu}{\partial z} = \frac{1}{R}\left[\frac{z - R/2}{r_1} - \frac{z + R/2}{r_2}\right] = -\frac{\nu z + \mu R/2}{r_1 r_2} \\ &\frac{\partial \varphi}{\partial x} = -\frac{y}{x^2 + y^2} \qquad \frac{\partial \varphi}{\partial y} = \frac{x}{x^2 + y^2} \qquad \frac{\partial \varphi}{\partial z} = 0 \end{aligned} \tag{7}$$

The Jacobian J can therefore be written:

$$J = -\frac{1}{(r_1 r_2)^2} \begin{vmatrix} \mu x & \mu y & \mu z + \nu R/2 \\ -\nu x & -\nu y & -\nu z - \mu R/2 \\ -y/\left(x^2+y^2\right) & x/\left(x^2+y^2\right) & 0 \end{vmatrix}$$

$$= \frac{1}{(r_1 r_2)^2} \frac{R}{2} (\mu^2 - \nu^2) \tag{8}$$

Since:

$$\mu^2 - \nu^2 = \frac{4 r_1 r_2}{R^2} \tag{9}$$

we get, finally:

$$J = \frac{8}{R^3(\mu^2 - \nu^2)} \tag{10}$$

1-c. Principle of the exact calculation

In the Born-Oppenheimer approximation, the equation to be solved in order to find the energy levels of the electron in the Coulomb field of the two fixed protons can be written:

$$\left[-\frac{\hbar^2}{2m}\Delta - \frac{e^2}{r_1} - \frac{e^2}{r_2} + \frac{e^2}{R} \right] \varphi(\mathbf{r}) = E\varphi(\mathbf{r}) \tag{11}$$

If we go into the elliptical coordinates defined in (3), we can separate the variables μ, ν and φ. Solving the equations so obtained, we find a discrete spectrum of possible energies for each value of R. We shall not perform this calculation here, but shall merely represent (the solid-line curve in Figure 2) the variation of the ground state energy with respect to R. This will enable us to compare the results we shall obtain by the variational method with the values given by the exact solution of equation (11).

2. The variational calculation of the energies

2-a. Choice of the trial kets

Assume R to be much larger than a_0. If we are concerned with values of r_1 of the order of a_0, we have, practically:

$$\frac{e^2}{r_2} \simeq \frac{e^2}{R} \qquad \text{for} \quad R,\, r_2 \gg a_0 \tag{12}$$

The Hamiltonian:

$$H = \frac{\mathbf{P^2}}{2m} - \frac{e^2}{r_1} - \frac{e^2}{r_2} + \frac{e^2}{R} \tag{13}$$

is then very close to that of a hydrogen atom centered at proton P_1. Analogous conclusions are, of course, obtained for R much larger than a_0, and r_2 of the order of a_0.

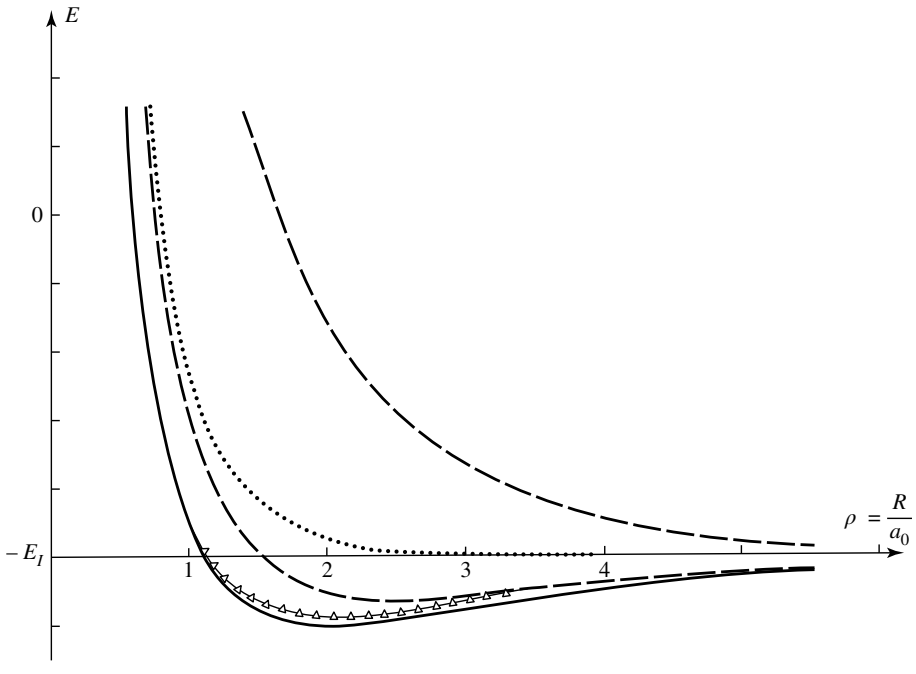


Figure 2: Variation of the energy E of the molecular ion H_2^+ with respect to the distance R between the two protons.
. solid line: the exact total energy of the ground state (the stability of the H_2^+ ion is due to the existence of a minimum in this curve).
. dotted line: the diagonal matrix element $H_{11} = H_{22}$ of the Hamiltonian H (the variation of this matrix element cannot explain the chemical bond).
. dashed line: the results of the simple variational calculation of § 2 for the bonding and antibonding states (though approximate, this calculation explains the stability of the H_2^+ ion).
. triangles: the results of the more elaborate variational calculation of § 3-a (taking atomic orbitals of adjustable radius considerably improves the accuracy, especially at small distances).

Therefore, when the two protons are very far apart, the eigenfunctions of the Hamiltonian (13) are practically the stationary wave functions of hydrogen atoms.

This is, of course, no longer true when a_0 is not negligible compared to R. We see, however, that it is convenient, for all R, to choose a family of trial kets constructed from atomic states centered at each of the two protons. This choice constitutes the application to the special case of the H_2^+ ion of a general method known as the *method of linear combination of atomic orbitals.* More precisely, we shall call $|\varphi_1\rangle$ and $|\varphi_2\rangle$ the

kets which describe the $1s$ states of the two hydrogen atoms:

$$\langle \mathbf{r} \mid \varphi_1 \rangle = \frac{1}{\sqrt{\pi a_0^3}} \, \mathrm{e}^{-\rho_1}$$
$$\langle \mathbf{r} \mid \varphi_2 \rangle = \frac{1}{\sqrt{\pi a_0^3}} \, \mathrm{e}^{-\rho_2} \tag{14}$$

We shall choose as trial kets all the kets belonging to the vector subspace $\mathcal{F}$ spanned by these two kets, that is, the set of kets $|\psi\rangle$ such that:

$$|\psi\rangle = c_1 \, |\varphi_1\rangle + c_2 \, |\varphi_2\rangle \tag{15}$$

The variational method (Complement $\mathrm{E_{XI}}$) consists of finding the stationary values of:

$$\langle H \rangle = \frac{\langle \psi | \, H \, | \psi \rangle}{\langle \psi \mid \psi \rangle} \tag{16}$$

within this subspace. Since this is a vector subspace, the average value $\langle H \rangle$ is minimal or maximal when $|\psi\rangle$ is an eigenvector of H inside this subspace $\mathcal{F}$, and the corresponding eigenvalue constitutes an approximation of a true eigenvalue of H in the total state space.

2-b. The eigenvalue equation of the Hamiltonian H in the trial ket vector subspace $\mathcal{F}$

The resolution of the eigenvalue equation of H within the subspace $\mathcal{F}$ is slightly complicated by the fact that $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are not orthogonal.

Any vector $|\psi\rangle$ of $\mathcal{F}$ is of the form (15). For it to be an eigenvector of H in $\mathcal{F}$ with the eigenvalue E, it is necessary and sufficient that:

$$\langle \varphi_i | \, H \, | \psi \rangle = E \, \langle \varphi_i \mid \psi \rangle \qquad i = 1, 2 \tag{17}$$

that is:

$$\sum_{j=1}^{2} c_j \, \langle \varphi_i | \, H \, | \varphi_j \rangle = E \sum_{j=1}^{2} c_j \, \langle \varphi_i | \varphi_j \rangle \tag{18}$$

We set:

$$S_{ij} = \langle \varphi_i \mid \varphi_j \rangle$$
$$H_{ij} = \langle \varphi_i | \, H \, | \varphi_j \rangle \tag{19}$$

We must solve a system of two linear homogeneous equations:

$$(H_{11} - ES_{11}) \, c_1 + (H_{12} - ES_{12}) \, c_2 = 0$$
$$(H_{21} - ES_{21}) \, c_1 + (H_{22} - ES_{22}) \, c_2 = 0 \tag{20}$$

This system has a non-zero solution only if:

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{21} - ES_{21} & H_{22} - ES_{22} \end{vmatrix} = 0 \tag{21}$$

The possible eigenvalues E are therefore the roots of a second-degree equation.

2-c. Overlap, Coulomb and resonance integrals

$|\varphi_1\rangle$ and $|\varphi_2\rangle$ are normalized; consequently:

$$S_{11} = S_{22} = 1 \tag{22}$$

On the other hand, $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are not orthogonal. Since the wave-functions (14) associated with these two kets are real, we have:

$$S_{12} = S_{21} = S \tag{23}$$

with:

$$S = \langle \varphi_1 \mid \varphi_2 \rangle = \int \mathrm{d}^3 r \; \varphi_1(\mathbf{r}) \, \varphi_2(\mathbf{r}) \tag{24}$$

S is called an *overlap integral*, since it receives contributions only from points of space at which the atomic wave functions φ_1 and φ_2 are both different from zero (such points exist if the two atomic orbitals partially "overlap"). A simple calculation gives:

$$S = \mathrm{e}^{-\rho} \left[1 + \rho + \frac{1}{3}\rho^2 \right] \tag{25}$$

To find this result, we can use elliptic coordinates (3), since:

$$\rho_1 = \frac{\mu + \nu}{2}\rho$$

$$\rho_2 = \frac{\mu - \nu}{2}\rho \tag{26}$$

According to expression (14) for the wave functions and the one for the volume element, (4), we must calculate:

$$\begin{aligned} S &= \frac{1}{\pi a_0^3} \int_1^{+\infty} \mathrm{d}\mu \int_{-1}^{+1} \mathrm{d}\nu \int_0^{2\pi} \mathrm{d}\varphi \; \frac{\rho^3 a_0^3}{8} \left(\mu^2 - \nu^2\right) \mathrm{e}^{-\mu\rho} \\ &= \frac{\rho^3}{2} \int_1^{+\infty} \mathrm{d}\mu \left(\mu^2 - \frac{1}{3}\right) \mathrm{e}^{-\mu\rho} \end{aligned} \tag{27}$$

which easily yields (25).

By symmetry:

$$H_{11} = H_{22} \tag{28}$$

According to expression (13) for the Hamiltonian H, we obtain:

$$H_{11} = \langle \varphi_1 | \left[\frac{\mathbf{P}^2}{2m} - \frac{e^2}{r_1} \right] |\varphi_1\rangle - \langle \varphi_1 | \frac{e^2}{r_2} |\varphi_1\rangle + \frac{e^2}{R} \langle \varphi_1 | \varphi_1 \rangle \tag{29}$$

Now, $|\varphi_1\rangle$ is a normalized eigenket of $\dfrac{\mathbf{P}^2}{2m} - \dfrac{e^2}{r_1}$. The first term of (29) is therefore equal to the energy $-E_I$ of the ground state of the hydrogen atom, and the third term is equal to e^2/R; we thus have:

$$H_{11} = -E_I + \frac{e^2}{R} - C \tag{30}$$

with:

$$C = \langle \varphi_1 | \frac{e^2}{r_2} | \varphi_1 \rangle = \int \mathrm{d}^3 r \, \frac{e^2}{r_2} \, [\varphi_1(\mathbf{r})]^2 \tag{31}$$

C is called a *Coulomb integral.* It describes (to within a change of sign) the electrostatic interaction between the proton P_2 and the charge distribution associated with the electron when it is in the $1s$ atomic state around the proton P_1. We find:

$$C = E_I \times \frac{2}{\rho} \left[1 - \mathrm{e}^{-2\rho} (1 + \rho) \right] \tag{32}$$

To find this result, we use elliptic coordinates again:

$$\begin{aligned} C &= \frac{e^2}{a_0 \rho} \frac{1}{\pi a_0^3} \frac{\rho^3 a_0^3}{8} \int \left(\mu^2 - \nu^2 \right) \mathrm{d}\mu \, \mathrm{d}\nu \, \mathrm{d}\varphi \, \frac{2}{\mu - \nu} \, \mathrm{e}^{-(\mu+\nu)\rho} \\ &= E_I \rho^2 \int_1^{+\infty} \mathrm{d}\mu \int_{-1}^{+1} \mathrm{d}\nu \, (\mu + \nu) \mathrm{e}^{-(\mu+\nu)\rho} \end{aligned} \tag{33}$$

Elementary integrations then lead to result (32).

In formula (30), C can be considered to be a modification of the repulsive energy e^2/R of the two protons: when the electron is in the state $|\varphi_1\rangle$, the corresponding charge distribution "screens" the proton P_1. Since $|\varphi_1(\mathbf{r})|^2$ is spherically symmetric about P_1, if the proton P_2 was far enough from it this charge distribution would appear to P_2 like a negative point charge e situated at its center P_1, (so that the charge of the proton P_1 would be totally cancelled). This does not actually happen unless R is much larger than a_0:

$$\lim_{R \to \infty} \left[\frac{e^2}{R} - C \right] = 0 \tag{34}$$

For finite R, the screening effect can only be partial, and we must have:

$$\frac{e^2}{R} - C > 0 \tag{35}$$

The variation of the energy $\frac{e^2}{R} - C$ with respect to R is shown in Figure 2 by the dotted line. It is clear that the variation of H_{11} (or H_{22}) with respect to R cannot explain the chemical bond, since this curve has no minimum.

Finally, let us calculate H_{12} and H_{21}. Since the wave functions $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ are real, we have:

$$H_{12} = H_{21} \tag{36}$$

Expression (13) for the Hamiltonian gives:

$$H_{12} = \langle \varphi_1 | \left[\frac{\mathbf{P}^2}{2m} - \frac{e^2}{r_2} \right] | \varphi_2 \rangle + \frac{e^2}{R} \langle \varphi_1 | \varphi_2 \rangle - \left\langle \varphi_1 \middle| \frac{e^2}{r_1} \middle| \varphi_2 \right\rangle \tag{37}$$

that is, according to definition (24) of S:

$$H_{12} = -E_I S + \frac{e^2}{R} S - A \tag{38}$$

with:

$$A = \langle \varphi_1 | \frac{e^2}{r_1} | \varphi_2 \rangle = \int \mathrm{d}^3 r \; \varphi_1(\mathbf{r}) \; \frac{e^2}{r_1} \; \varphi_2(\mathbf{r}) \tag{39}$$

We shall call A the *resonance integral*[1]. It is equal to:

$$A = E_I \times 2\mathrm{e}^{-\rho} (1 + \rho) \tag{40}$$

The use of elliptic coordinates enables us to write A in the form:

$$\begin{aligned} A &= \frac{e^2}{a_0} \frac{1}{\pi a_0^3} \frac{\rho^3 a_0^3}{8} \int \left(\mu^2 - \nu^2\right) \mathrm{d}\mu \; \mathrm{d}\nu \; \mathrm{d}\varphi \; \frac{2\mathrm{e}^{-\mu\rho}}{(\mu + \nu)\,\rho} \\ &= \rho^2 E_I \int_1^{+\infty} \mathrm{d}\mu \, 2\mu \, \mathrm{e}^{-\mu\rho} \end{aligned} \tag{41}$$

The fact that H_{12} is different from zero expresses the possibility of the electron "jumping" from the neighborhood of one of the protons to that of the other one. If, at some time, the electron is in the state $|\varphi_1\rangle$ (or $|\varphi_2\rangle$), it oscillates in time between the two sites, under the influence of H_{12}. This non-diagonal matrix element is therefore responsible for the phenomenon of *quantum resonance*, which we described qualitatively in § C-2-c of Chapter IV (hence the name of integral A).

To sum up, the parameters which are functions of R and are involved in equation (21) for the approximate energies E are:

$$\begin{aligned} &S_{11} = S_{22} = 1 \qquad S_{12} = S_{21} = S \\ &H_{11} = H_{22} = -E_I + \frac{e^2}{R} - C \\ &H_{12} = H_{21} = \left(-E_I + \frac{e^2}{R}\right) S - A \end{aligned} \tag{42}$$

where S, C and A are given by (25), (32) and (40), and are plotted in Figure 3. Note that the non-diagonal elements of determinant (21) take on significant values only if the orbitals $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ partially overlap, since the product $\varphi_1(\mathbf{r})\,\varphi_2(\mathbf{r})$ appears in definition (39) of A, as well as in that of S.

2-d. Bonding and antibonding states

α. Calculation of the approximate energies

We set:

$$\begin{aligned} E &= \varepsilon \; E_I \\ A &= \alpha \; E_I \\ C &= \gamma \; E_I \end{aligned} \tag{43}$$

[1] Certain authors call A an "exchange integral". We prefer to restrict the use of this term to another type of integral which is encountered in many-particle systems (Complement $\mathrm{B_{XIV}}$, § 2-c-β).

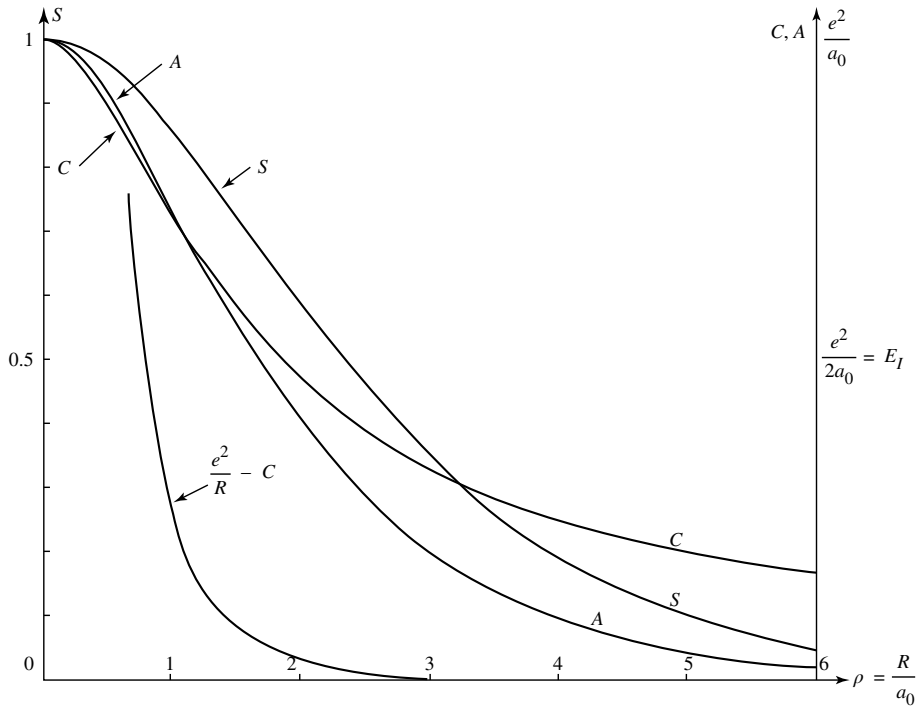


Figure 3: Variation of S (the overlap integral), C (the Coulomb integral) and A (the resonance integral) with respect to $\rho = R/a_0$. When $R \to \infty$, S and A approach zero exponentially, while C decreases only with e^2/R (the "screened" interaction $\frac{e^2}{R} - C$ of the proton P_1 with the atom centered at P_2 also decreases exponentially, however).

Equation (21) can then be written:

$$\begin{vmatrix} -1 + \frac{2}{\rho} - \gamma - \varepsilon & \left(-1 + \frac{2}{\rho}\right) S - \alpha - \varepsilon S \\ \left(-1 + \frac{2}{\rho}\right) S - \alpha - \varepsilon S & -1 + \frac{2}{\rho} - \gamma - \varepsilon \end{vmatrix} = 0 \tag{44}$$

or:

$$\left[\gamma + \varepsilon + 1 - \frac{2}{\rho}\right]^2 = \left[\alpha + \left(\varepsilon + 1 - \frac{2}{\rho}\right) S\right]^2 \tag{45}$$

This gives the following two values for ε:

$$\varepsilon_+ = -1 + \frac{2}{\rho} + \frac{\alpha - \gamma}{1 - S} \tag{46a}$$

$$\varepsilon_- = -1 + \frac{2}{\rho} - \frac{\alpha + \gamma}{1 + S} \tag{46b}$$

ε_+ and ε_- both approach -1 when ρ approaches infinity. This means that the two approximate energies $E_\pm$ approach $-E_I$, the ground state energy of an isolated hydrogen

atom, as expected (§ 2-a). Furthermore, it is convenient to choose this value as the energy origin, that is, to set:

$$\Delta E = E(\rho) - E(\infty) = E + E_I \tag{47}$$

Using (25), (32) and (40), the approximate energies ΔE_+ and ΔE_- can be written:

$$\Delta E_\pm = E_I \left\{ \frac{2}{\rho} \pm \frac{2e^{-\rho}(1+\rho) \mp \frac{2}{\rho}\left[1 - e^{-2\rho}(1+\rho)\right]}{1 \mp e^{-\rho}(1+\rho+\rho^2/3)} \right\} \tag{48}$$

The variation of $\Delta E_\pm / E_I$ with respect to ρ is shown in dashed lines in Figure 2. We see that ΔE_- has a negative minimum for a certain value of the distance R between the two protons. Although this is an approximation (*cf.* Fig. 2), it explains the existence of the chemical bond.

As we have already pointed out, the variation with respect to R of the diagonal elements H_{11} and H_{22} of determinant (21) has no minimum (dotted-line curve of Figure 2). The minimum of ΔE_- therefore is due to the non-diagonal elements H_{12} and S_{12}. This shows that the phenomenon of the chemical bond appears only if the electronic orbitals of the two atoms participating in the bond overlap sufficiently.

β. Eigenstates of H inside the subspace $\mathcal{F}$

The eigenstate corresponding to E_- is called a *bonding state*, and the one corresponding to E_+, an *antibonding state*, since E_+ always remains greater than the energy $-E_I$ of the system formed by a hydrogen atom in the ground state and an infinitely distant proton.

According to (45):

$$\gamma + \varepsilon + 1 - \frac{2}{\rho} = \pm\left[\alpha + \left(\varepsilon + 1 - \frac{2}{\rho}\right) S\right] \tag{49}$$

System (20) then gives:

$$c_1 \pm c_2 = 0 \tag{50}$$

The bonding and antibonding states are therefore symmetric and antisymmetric linear combinations of the kets $|\varphi_1\rangle$ and $|\varphi_2\rangle$. To normalize them, it must be recalled that $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are not orthogonal (their scalar product is equal to S). We therefore obtain:

$$|\psi_+\rangle = \frac{1}{\sqrt{2(1-S)}}[|\varphi_1\rangle - |\varphi_2\rangle] \tag{51a}$$

$$|\psi_-\rangle = \frac{1}{\sqrt{2(1+S)}}[|\varphi_1\rangle + |\varphi_2\rangle] \tag{51b}$$

Note that the *bonding state* $|\psi_-\rangle$, associated with E_-, *is symmetric* under exchange of $|\varphi_1\rangle$ and $|\varphi_2\rangle$, while the antibonding state is antisymmetric.

Comment:

It could have been expected that the eigenstates of H inside the subspace $\mathcal{F}$ would be symmetric and antisymmetric combinations of $|\varphi_1\rangle$ and $|\varphi_2\rangle$: for given positions of the two protons, there is symmetry with respect to the bisecting plane of P_1P_2, and H remains unchanged if the roles of the two protons are exchanged.

The bonding and antibonding states are approximate stationary states of the system under study. We pointed out in Complement E_{XI} that the variational method can give a valid approximation for the energies but gives a more debatable result for the eigenfunctions. It is instructive, however, to have an idea of the mechanism of the chemical bond, to represent graphically the wave functions associated with the bonding and antibonding states, which are often called bonding and antibonding *molecular orbitals.* To do so, we can, for example, trace the surfaces of equal $|\psi|$ (the locus of points in space for which the modulus $|\psi|$ of the wave function has a given value). If ψ is real, we indicate by a + (or −) sign the regions in which it is positive (or negative). This is what is done in Figure 4 for ψ_+ and ψ_- (the surfaces of equal $|\psi|$ are surfaces of revolution about the P_1P_2 axis, and Figure 4 only shows their cross sections in a plane containing P_1P_2). The difference between the bonding orbital and the antibonding orbital is striking. In the first one, the electronic cloud "streches out" to include both protons, while in the second one, the position probability of the electron is zero in the bisecting plane of P_1P_2.

Comment:

We can calculate the average value of the potential energy in the state $|\psi_-\rangle$, which, if we use (51b), (31) et (39), is equal to:

$$\begin{aligned}\langle V\rangle &= \langle\psi_-|\left[\frac{e^2}{R}-\frac{e^2}{r_1}-\frac{e^2}{r_2}\right]|\psi_-\rangle \\ &= \frac{e^2}{R}-\frac{1}{1+S}\left[\langle\varphi_1\left|\frac{e^2}{r_1}\right|\varphi_1\rangle+\langle\varphi_1\left|\frac{e^2}{r_1}\right|\varphi_2\rangle\right. \\ &\quad \left.+\langle\varphi_1\left|\frac{e^2}{r_2}\right|\varphi_1\rangle+\langle\varphi_1\left|\frac{e^2}{r_2}\right|\varphi_2\rangle\right] \\ &= E_I\left[\frac{2}{\rho}-\frac{1}{1+S}(2+2\alpha+\gamma)\right] \end{aligned} \tag{52}$$

Subtracting this from (46b), we obtain the kinetic energy:

$$\begin{aligned}\langle T\rangle &= \left\langle\frac{\mathbf{P}^2}{2m}\right\rangle = \langle H-V\rangle \\ &= E_I\frac{1}{1+S}(1-S+\alpha)\end{aligned} \tag{53}$$

We shall discuss later (§ 5) to what extent (52) and (53) give good approximations for the kinetic and potential energies.

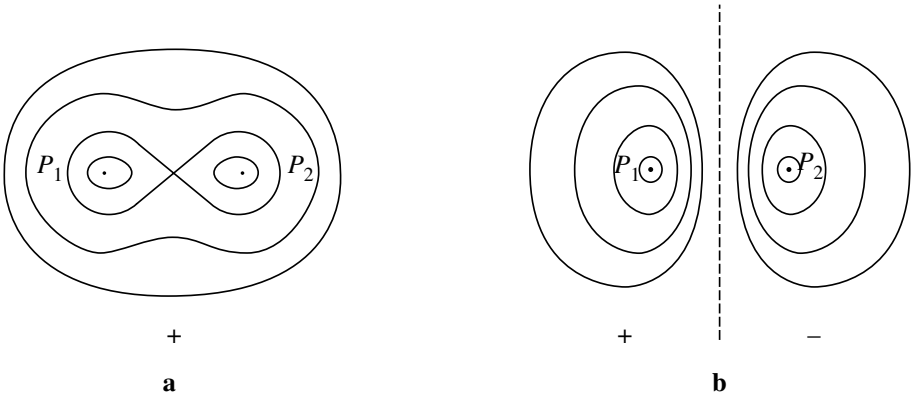


Figure 4: Schematic drawings of the bonding molecular orbital (fig. a) and the antibonding molecular orbital (fig. b) for the H_2^+ ion. We have shown the cross section in a plane containing P_1P_2 of a family of surfaces for which the modulus $|\psi|$ of the wave function has a constant given value. These are surfaces of revolution about P_1P_2 (we have shown 4 surfaces, corresponding to 4 different values of $|\psi|$). The + and − signs indicated in the figure are those of the wave function (which is real) in the corresponding regions. The dashed line is the trace of the bisecting plane of P_1P_2, which is a nodal plane for the antibonding orbit.

3. Critique of the preceding model. Possible improvements

3-a. Results for small R

What happens to the energy of the bonding state and the corresponding wave function when $R \longrightarrow 0$?

We see from Figure 3 that S, A and C approach, respectively, 1, $2E_I$ and $2E_I$ when $\rho \longrightarrow 0$. If we subtract the repulsion term e^2/R of the two protons, to obtain the electronic energy, we find:

$$E_- - \frac{e^2}{R} \underset{R\to 0}{\longrightarrow} -3E_I \tag{54}$$

In addition, since $|\varphi_1\rangle$ approaches $|\varphi_2\rangle$, $|\psi_-\rangle$ reduces to $|\varphi_1\rangle$ (the ground state $1s$ of the hydrogen atom).

This result is obviously incorrect. When $R = 0$, we have the equivalent[2] of a helium ion He^+. The electronic energy of the ground state of H_2^+ must coincide, for $R = 0$, with that of the ground state of He^+. Since the helium nucleus is a $Z = 2$ nucleus, this energy is (*cf.* Complement A_{VII}):

$$-Z^2E_I = -4E_I \tag{55}$$

and not $-3E_I$. Furthermore, the wave function $\psi_-(\mathbf{r})$ should not approach $\varphi_1(\mathbf{r}) = (\pi\alpha_0^3)^{-1/2}\,e^{-\rho_1}$, but rather $(\pi\alpha_0^3/Z^3)^{-1/2}\,e^{-Z\rho_1}$ with $Z = 2$ (the Bohr orbit is twice as small). This enables us to understand why the disagreement between the exact result and that of § 2 above becomes

[2]In addition to the two protons, the helium nucleus of course contains one or two neutrons.

important for small values of R (Fig. 2): this calculation uses atomic orbitals which are too spread out when the two protons are too close to each other.

A possible improvement therefore consists of enlarging the family of trial kets because of these physical arguments and using kets of the form:

$$|\psi\rangle = c_1|\varphi_1(Z)\rangle + c_2|\varphi_2(Z)\rangle \tag{56}$$

where $|\varphi_1(Z)\rangle$ and $|\varphi_2(Z)\rangle$ are associated with $1s$ atomic orbitals of radius a_0/Z centered at P_1 and P_2. The ground state still corresponds, for reasons of symmetry, to $c_1 = c_2$. We consider Z like a variational parameter in seeking, for each value of R, the value of Z which minimizes the energy.

The calculation can be performed completely in elliptic coordinates. We find (*cf.* Fig. 5) that the optimal value of Z decreases from $Z = 2$ for $R = 0$ to $Z = 1$ for $R \longrightarrow \infty$, as it should.

The curve obtained for ΔE_- is much closer to the exact curve (*cf.* Fig. 2). Table I gives the values of the abscissa and ordinate of the minimum of ΔE_- obtained from the various models considered in this complement. It can be seen from this table that the energies found by the variational method are always greater than the exact energy of the ground state; in addition, we see that enlarging the family of trial kets improves the results for the energy.

3-b. Results for large R

When $R \longrightarrow \infty$, we see from (48) that E_+ and E_- exponentially approach the same value $-E_I$. Actually, this limit should not be obtained so rapidly. To see this, we shall use a perturbation approach, as in Complement C_{XI}, (Van der Waals forces) or E_{XII} (the Stark effect of the hydrogen atom). Let us evaluate the perturbation of the energy of a hydrogen atom (in the $1s$ state), situated at P_2, produced by the presence of a proton P_1 situated at a distance R much greater than a_0 ($\rho \gg 1$). In the neighborhood of P_2, the proton P_1 creates an electric field $\mathbf{E}$, which varies like $1/R^2$. This field polarizes the hydrogen atom and causes an electric dipole moment $\mathbf{D}$, proportional to $\mathbf{E}$, to appear. The electronic wave function is distorted, and the barycenter of the electronic charge distribution moves closer to P_1 (Fig. 6). $\mathbf{E}$ and $\mathbf{D}$ are both proportional to $1/R^2$ and have the same sign. The electrostatic interaction between the proton P_1 and the atom situated at P_2 must therefore lower[3]. Consequently, the asymptotic behavior of ΔE_+ and ΔE_- must vary, not exponentially, but as $-c/R^4$ (where c is a positive constant) the energy by an amount which, like $-\mathbf{E}\cdot\mathbf{D}$, varies as $1/R^4$.

It is actually possible to find this result by the variational method. Instead of linearly superposing $1s$ orbitals centered at P_1 and P_2, we shall superpose hybrid orbitals χ_1, and χ_2, which are not spherically symmetric about P_1 and P_2. The hybrid orbital χ_2 is obtained, for example, by linearly superposing a $1s$ orbital and a $2p$ orbital, both centered[4] at P_2:

$$\chi_2(\mathbf{r}) = \varphi_{1s}^2(\mathbf{r}) + \sigma\,\varphi_{1p}^2(\mathbf{r}) \tag{57}$$

and has a form analogous to the one shown in Figure 6. Now, consider determinant (21). The non-diagonal elements $H_{12} = \langle\chi_1|H|\chi_2\rangle$ and $S_{12} = \langle\chi_1|\chi_2\rangle$ still approach zero exponentially when $R \longrightarrow \infty$. This is because the product $\chi_1(\mathbf{r})\chi_2(\mathbf{r})$ appears in the corresponding integrals; even though distorted, the orbitals $\chi_1(\mathbf{r})$ and $\chi_2(\mathbf{r})$ still remain localized in the neighborhoods of P_1 and P_2 respectively, and their overlap goes to zero exponentially when $R \longrightarrow \infty$. The two eigenvalues E_+ and E_- therefore both approach $H_{11} = H_{22}$ when $R \longrightarrow \infty$, since determinant (21) becomes diagonal.

Now, what does H_{22} represent? As we have seen (*cf.* § 2-c), it is the energy of a hydrogen atom placed at P_2 and perturbed by the proton P_1. The calculation of § 2 neglected

[3]More precisely, the energy is lowered by $-\frac{1}{2}\,\mathbf{E}\cdot\mathbf{D}$ (*cf.* Complement E_{XII}, § 1).

[4]The symmetry axis of the $2p$ orbital is chosen along the straight line joining the two protons.

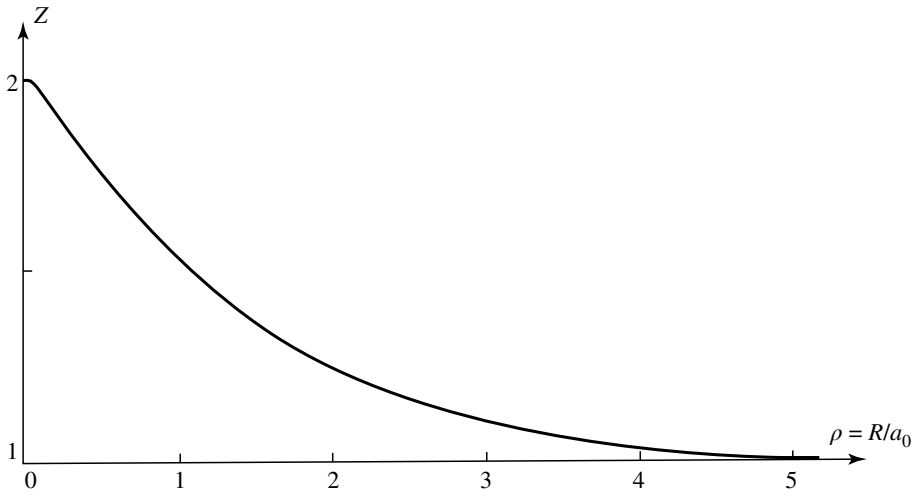


Figure 5: For each value of the internuclear distance, we have calculated the value of Z which minimizes the energy. For $R = 0$, we have the equivalent of a He^+ ion, and we indeed find $Z = 2$. For $R \gg a_0$, we have essentially an isolated hydrogen atom, which gives $Z = 1$. Between these two extremes, Z is a decreasing function of ρ. The corresponding optimal energies are represented by triangles in Figure 2.

---

any polarization of the $1s$ electronic orbital due to the effect of the electric field created by P_1, and this is why we found an energy correction decreasing exponentially when R increases. However, if, as we are doing here, we take into account the polarization of the electronic orbital, we find a correction in $-c/R^4$. The fact that, in (57), we consider only the mixing with the $2p$ orbital causes the value of c given by the variational calculation to be approximate (whereas the perturbation calculation of the polarization involves all the excited states, *cf.* Complement E_{XII}).

The two curves ΔE_+ and ΔE_- therefore do approach each other exponentially, since the difference between E_+ and E_- involves only the non-diagonal elements H_{12} and S_{12}, and their common value for large R approaches zero proportionally to $-1/R^4$ (Fig. 7).

The preceding discussion also suggests using polarized orbitals like the one in (57), not only for large R, but also for all other values of R as well. We would thus enlarge the family

---

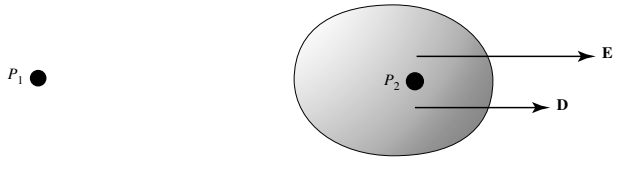


Figure 6: Under the effect of the electric field **E** *created by the proton P_1, the electronic cloud of the hydrogen atom centered at P_2 becomes distorted, and this atom acquires an electric dipole moment* **D**. *An interaction energy results which decreases with $1/R^4$ when R increases.*

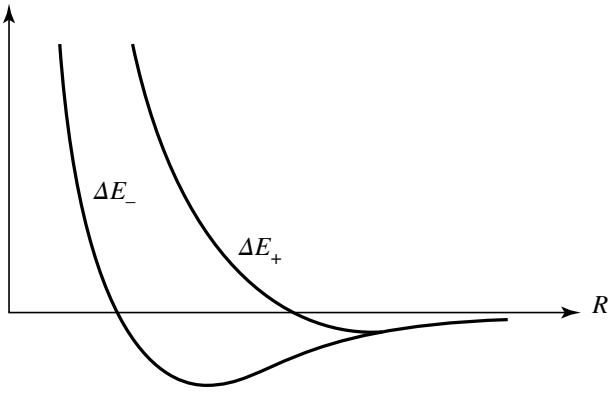


Figure 7: When $\rho \longrightarrow \infty$, the energies of the bonding and antibonding states approach each other exponentially. However, they approach their limiting value less rapidly (like $1/R^4$).

of trial kets and consequently improve the accuracy. In expression (57), we then consider σ as a variation parameter, like the parameter Z that defines the Bohr radius a_0/Z associated with the $1s$ and $2p$ orbitals. To make the method even more flexible, we choose different parameters Z and Z' for φ_{1s} and φ_{2p}. For each value of R, we then minimize the average value of H in the state $|\chi_1\rangle + |\chi_2\rangle$ (which, for reasons of symmetry, is still the ground state), and we determine the optimal values of σ, Z, Z'. The agreement with the exact solution then becomes excellent (*cf.* Table I).

4. Other molecular orbitals of the H_2^+ ion

In the preceding sections, we obtained by the variational method a bonding and an antibonding molecular orbitals. They were obtained from the ground state $1s$ of each of

	Distance d'équilibre des deux protons (abscisse du minimum de ΔE_-)	Profondeur du puits de potentiel (valeur du minimum de ΔE_-)
Méthode variationnelle du § 2 (orbitales $1s$ avec $Z = 1$)	2,50 a_0	1,76 eV
Méthode variationnelle du § **??** (orbitales $1s$ avec Z variable)	2,00 a_0	2,35 eV
Méthode variationnelle du § 3-b (orbitales hybrides avec Z, Z', σ variables)	2,00 a_0	2,73 eV
Valeurs exactes	2,00 a_0	2,79 eV

Tableau I

the two hydrogen atoms, formed about the two protons. Of course, we chose the $1s$ state because it was clear that this would be the best choice for obtaining an approximation of the ground state of a system of two protons and one electron. We can obviously envisage, with the method of linear combination of atomic orbitals (§ 2-a), using excited states of the hydrogen atom to obtain other molecular orbitals of higher energies. The main interest of these excited orbitals will be to give us an idea of the phenomena which can come into play in molecules which are more complex than the H_2^+ ion. For example, to understand the properties of a diatomic molecule containing several electrons, we can, in a first approximation, treat these electrons individually, as if they did not interact with each other. We thus determine the various possible stationary states for an isolated electron placed in the Coulomb field of the nuclei, and then place the electrons of the molecule in these states, taking the Pauli principle into account (Chap. XIV, § D-1) and filling the lowest energy states first (this procedure is analogous to the one described for many-electron atoms in Complement A_{XIV}). In this section, we shall indicate the principal properties of the excited molecular orbitals of the H_2^+ ion, while keeping in mind the possibilities of generalization to more complex molecules.

4-a. Symmetries and quantum numbers. Spectroscopic notation

(*i*) The potential V created by the two protons is symmetric with respect to revolution about the P_1P_2 axis, which we shall choose as the Oz axis. This means that V and, consequently, the Hamiltonian H of the electron, do not depend on the angular variable φ which fixes the orientation about Oz of the MP_1P_2 plane containing the Oz axis and the point M. It follows that H commutes with the component L_z of the orbital angular momentum of the electron [in the $\{|\mathbf{r}\rangle\}$ representation, L_z becomes the differential operator $\frac{\hbar}{i}\frac{\partial}{\partial\varphi}$, which commutes with any φ-independent operator]. We can then find a system of eigenstates of H that are also eigenstates of L_z, and class them according to the eigenvalues $m\hbar$ of L_z.

(*ii*) The potential V is also invariant under reflection through any plane containing P_1P_2, that is, the Oz axis. Under such a reflection, an eigenstate of L_z of eigenvalue $m\hbar$ is transformed into an eigenstate of L_z of eigenvalue $-m\hbar$ (the reflection changes the sense of revolution of the electron about Oz). Because of the invariance of V, the energy of a stationary state depends only on $|m|$.

In spectroscopic notation, we label each molecular orbital with a Greek letter indicating the value of $|m|$, as follows:

$$\begin{aligned} |m| &= 0 \leftrightarrow \sigma \\ |m| &= 1 \leftrightarrow \pi \\ |m| &= 2 \leftrightarrow \delta \end{aligned} \tag{58}$$

(note the analogy with atomic spectroscopic notation: σ, π, δ recall s, p, d). For example, since the ground state $1s$ of the hydrogen atom has a zero orbital angular momentum, the two orbitals studied in the preceding sections are σ orbitals (it can be shown that this is also true for the exact stationary wave functions, and not only for the approximate states obtained by the variational method).

This notation does not use the fact that the two protons of the H_2^+ ion have equal charges. The σ, π, δ classification of molecular orbitals therefore remains valid for a heteropolar diatomic molecule.

(*iii*) In the H_2^+ ion (and, more generally, in homopolar diatomic molecules), the potential V is invariant under reflection through the middle O of P_1P_2. We can therefore choose eigenfunctions of the Hamiltonian H in such a way that they have a definite parity with respect to the point O. For an even orbital, we add to the Greek letter which characterizes $|m|$, an index g (from the German "gerade"); this index is u ("ungerade") for odd orbitals. Thus, the bonding orbital obtained above from the $1s$ atomic states is a σ_g orbital, while the corresponding antibonding orbital is σ_u.

(*iv*) Finally, we can use the invariance of H under reflection through the bisecting plane of P_1P_2 to choose stationary wave functions which have a definite parity in this operation, that is, a parity defined with respect to the change in sign of the variable z. Functions which are odd under this reflection are labeled with an asterisk. They are necessarily zero at all points of the bisecting plane of P_1P_2, like the orbital shown in Figure 4b; these are antibonding orbitals.

Comment:

Reflection through the bisecting plane of P_1P_2 can be obtained by performing a reflection through O followed by a rotation of π about Oz. The parity (*iv*) is therefore not independent of the preceding symmetries (the "g" states will have an asterisk for odd $|m|$ and none for even $|m|$; the situation is reversed for the "u" states). However, it is convenient to consider this parity, since it enables us to determine the antibonding orbitals immediately.

4-b. Molecular orbitals constructed from the $2p$ atomic orbitals

If we start with the excited state $2s$ of the hydrogen atom, arguments analogous to those of the preceding sections will give a bonding $\sigma_g(2s)$ orbital and an antibonding $\sigma_u^*(2s)$ orbital, with forms similar to those in Figure 4. We shall therefore concern ourselves instead with molecular orbitals obtained from the excited atomic states $2p$.

α. *Orbitals constructed from* $2p_z$ *states*

We shall denote by $|\varphi^1_{2p_z}\rangle$ and $|\varphi^2_{2p_z}\rangle$ the atomic states $2p_z$ (*cf.* Complement E_{VII}, § 2-b), centered at P_1 and P_2 respectively. The form of the corresponding orbitals is shown in Figure 8 (note the choice of signs, indicated in the figure).

By a variational calculation analogous to the one in § 2, we can construct, starting with these two atomic states, two approximate eigenstates of the Hamiltonian (13). The symmetries recalled in § 4-a imply that, to within a normalization factor, these molecular states can be written:

$$|\varphi^1_{2p_z}\rangle + |\varphi^2_{2p_z}\rangle \tag{59a}$$

$$|\varphi^1_{2p_z}\rangle - |\varphi^2_{2p_z}\rangle \tag{59b}$$

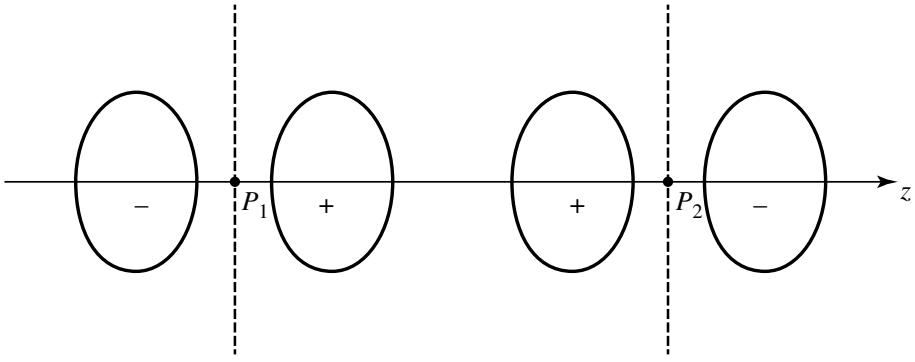


Figure 8: Schematic representation of the $2p_z$ atomic orbitals centered at P_1 and P_2 (the Oz axis is chosen along P_1P_2) and used as a basis for constructing the excited molecular orbitals $\sigma_g(2p_z)$ and $\sigma_u^*(2p_z)$ shown in Figure 9 (note the sign convention chosen).

The shape of the two molecular orbitals so obtained can easily be deduced from Figure 8; they are shown in Figure 9.

The two atomic states $2p_z$ are eigenstates of L_z with the eigenvalue zero; the same is therefore also true of the two states (59). The molecular orbital associated with (59a) is even and is written $\sigma_g(2p_z)$; the one corresponding to (59b) is odd under a reflection through O as well as under a reflection through the bisecting plane of P_1P_2, and we shall therefore denote it by $\sigma_u^*(2p_z)$.

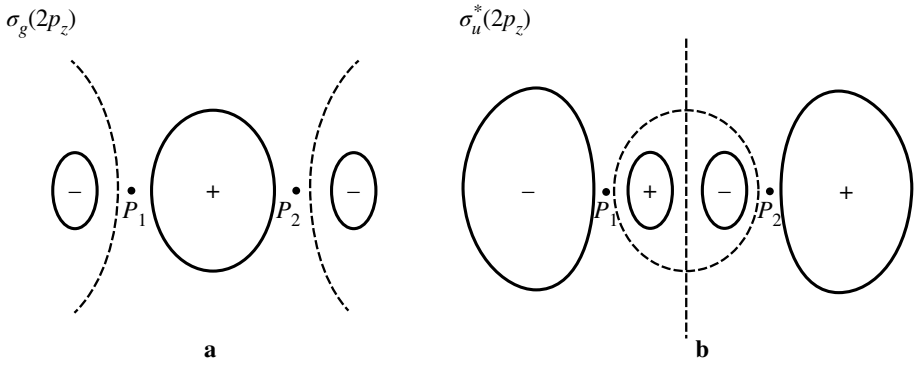


Figure 9: Schematic representation of the excited molecular orbitals: the bonding orbital $\sigma_g(2p_z)$ (fig. a) and the antibonding orbital $\sigma_u^*(2p_z)$ (fig. b). As in Figure 8, we have drawn the cross section in a plane containing P_1P_2 of a constant modulus $|\psi|$ surface. This is a surface of revolution about P_1P_2. The sign shown is that of the (real) wave function. The dashed-line curves are the cross sections in the plane of the figure of the nodal surfaces ($|\psi| = 0$).

Remark:

As we mentioned in the introduction of this complement, the major interest of the excited orbitals we study here is their application to molecules more complex than H_2^+. For such molecules, the orbitals $2s$ and $2p$ have different energies, so that it is legitimate to consider them separately. If, however, we specifically study the hydrogen molecular H_2^+ ion, the sates $2s$ and $2p$ are then degenerate, so that they are immediately mixed by the electric field of a nearby proton. In this case, there is no reason to study the orbitals $2s$ et $2p_z$ separately; it is more appropriate to introduce hybrid orbitals similar to those discussed in § 3 of Complement E_{VII}.

β. Orbitals constructed from $2p_x$ or $2p_y$ states

We shall now start with the atomic states $|\varphi^1_{2p_x}\rangle$ and $|\varphi^2_{2p_x}\rangle$, with which are associated the real wave functions (*cf.* Complement E_{VII}, § 2-b) shown in Figure 10 (note that the surfaces of equal $|\psi|$ whose cross sections in the xOz plane are given in Figure 10 are surfaces of revolution, not about Oz, but about axes parallel to Ox and passing through P_1 and P_2). Recall that the atomic orbital $2p_x$ is obtained by the linear combination of eigenstates of L_z corresponding to $m = 1$ and $m = -1$. The molecular orbitals constructed from these atomic orbitals therefore have $|m| = 1$; they are π orbitals.

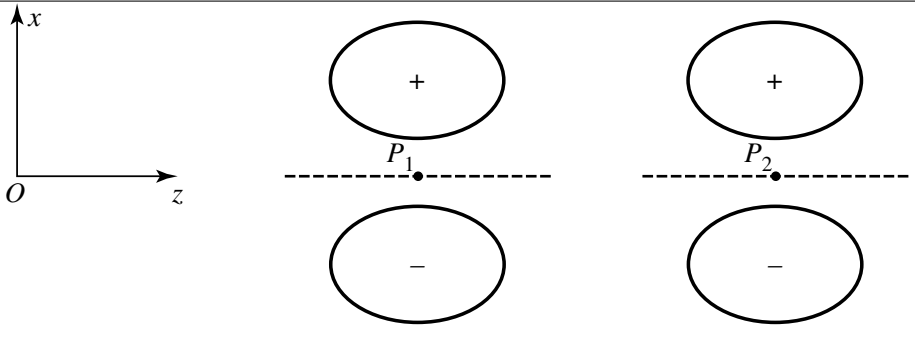


Figure 10: Schematic representation of the atomic orbitals $2p_x$ centered at P_1 and P_2 (the Oz axis is chosen along P_1P_2) and used as a basis for constructing the excited molecular orbitals $\pi_u(2p_x)$ and $\pi_g^*(2p_x)$ shown in Figure 11. For each orbital, the surface of equal $|\psi|$, whose cross section in the xOz plane is shown, is a surface of revolution, no longer about Oz, but about a straight line parallel to Ox and passing either through P_1 or P_2.

Here again, the approximate molecular states produced from the atomic states $2p_x$ are the symmetric and antisymmetric linear combinations:

$$|\varphi^1_{2p_x}\rangle + |\varphi^2_{2p_x}\rangle \tag{60a}$$

$$|\varphi^1_{2p_x}\rangle - |\varphi^2_{2p_x}\rangle \tag{60b}$$

The form of these molecular orbitals can easily be qualitatively deduced from Figure 10. The surfaces of equal $|\psi|$ are not surfaces of revolution about Oz, but are simply symmetric with respect to the xOz plane. Their cross sections in this plane are shown in

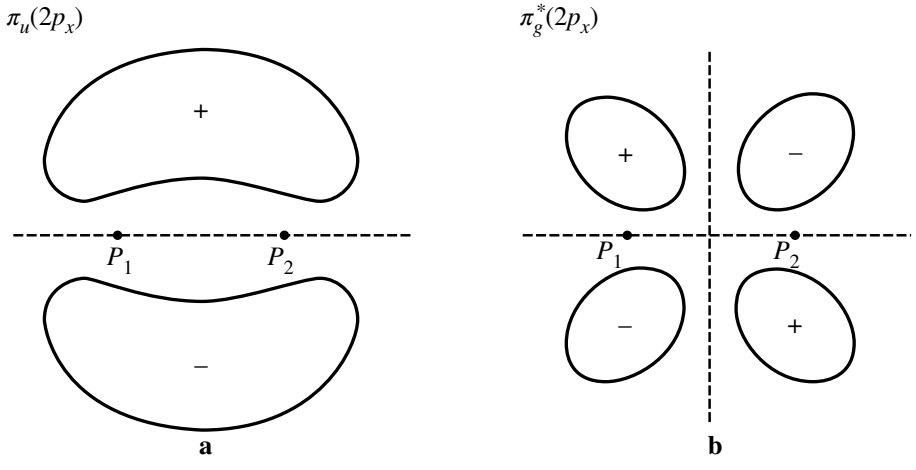


Figure 11: Schematic representation of the excited molecular orbitals: the bonding orbital $\pi_u(2p_x)$ (fig. a) and the antibonding orbital $\pi_g^*(2p_x)$ (fig. b). For each of these two orbitals, we have shown the cross section in the xOz plane of a surface on which $|\psi|$ has a given constant value. This surface is no longer a surface of revolution but is simply symmetric with respect to the xOz plane. The meaning of the signs and the dashed lines is the same as in Figures 4, 8, 9, 10.

Figure 11. We see immediately in this figure that the orbital associated with state (60a) is odd with respect to the middle O of P_1P_2 but even with respect to the bisecting plane of P_1P_2; it will therefore be denoted by $\pi_u(2p_x)$. On the other hand, the orbital corresponding to (60b) is even with respect to point O and odd with respect to the bisecting plane of P_1P_2: it is an antibonding orbital, denoted by $\pi_g^*(2p_x)$. We stress the fact that these π orbitals have *planes of symmetry,* not axes of revolution like the σ orbitals.

Of course, the molecular orbitals produced by the atomic states $2p_y$ can be deduced from the preceding ones by a rotation of $\pi/2$ about P_1P_2.

π orbitals analogous to the preceding ones are involved in the double or triple bonds of atoms such as carbon (*cf.* Complement E_{VII}, §§ 3-c and 4-c).

Comment:

We saw earlier (§ 2-d) that the energy separation of the bonding and antibonding levels is due to the overlap of the atomic wave functions. Now, for the same distance R, the overlap of the $\varphi^1_{2p_z}$ and $\varphi^2_{2p_z}$ orbitals, which point towards each other, is larger than that of $\varphi^1_{2p_x}$ and $\varphi^2_{2p_x}$, whose axes are parallel (Fig. 8 and 10). We see that the energy difference between $\sigma_g(2p_z)$ and $\sigma_u^*(2p_z)$ is larger than that between $\pi_u(2p_x)$ and $\pi_g^*(2p_x)$ [or $\pi_u(2p_y)$ and $\pi_g^*(2p_y)$]. The hierarchy of the corresponding levels is indicated in Figure 12.

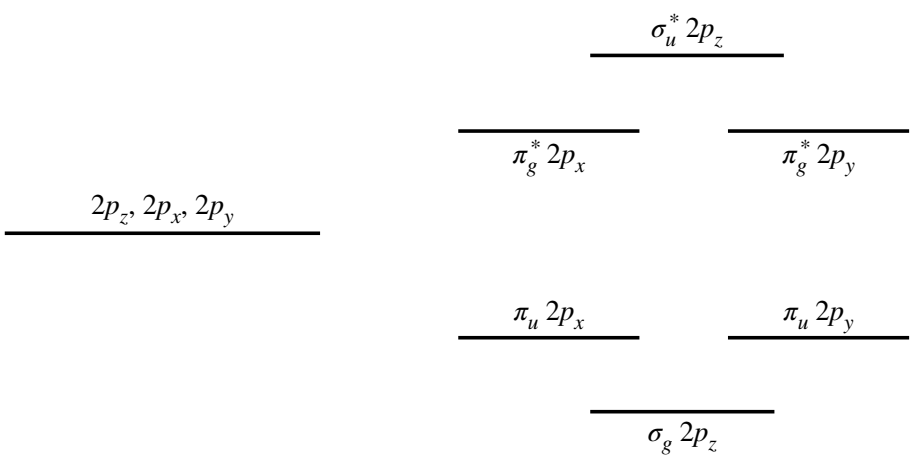


Figure 12: The energies of the various excited molecular orbitals constructed from the atomic orbitals $2p_z$, $2p_x$ and $2p_y$ centered at P_1 and P_2 (the Oz axis is chosen along P_1P_2). By symmetry, the molecular orbitals produced by the $2p_x$ atomic orbitals are degenerate with those produced by the $2p_y$ atomic orbitals. The difference between the bonding and antibonding molecular orbitals $\pi_u(2p_{x,y})$ and $\pi_g^*(2p_{x,y})$ is, however, smaller than the corresponding difference between the $\sigma_g(2p_z)$ and $\sigma_u^*(2p_z)$ molecular orbitals. This is due to the larger overlap of the two $2p_z$ atomic orbitals.

5. The origin of the chemical bond; the virial theorem

5-a. Statement of the problem

When the distance R between the protons decreases, their electrostatic repulsion e^2/R increases. The fact that the total energy $E_-(R)$ of the bonding state decreases (when R decreases from a very large value) and then passes through a minimum therefore means that the electronic energy begins by decreasing faster than e^2/R increases (of course, since this term diverges when $R \longrightarrow 0$, it is the repulsion between the protons which counts at short distances). We can then ask the following question: does the lowering of the electronic energy, which makes the chemical bond possible, arise from a lowering of the electronic potential energy or from a lowering of the kinetic energy or from both?

We have already calculated, in (52) and (53), approximate expressions for the (total) potential and kinetic energies. We might then consider studying the variation of these expressions with respect to R. Such a method, however, would have to be used with caution, since, as we have already pointed out, the eigenfunctions supplied by a variational calculation are much less precise than the energies. We shall discuss this point in greater detail in § 5-d-β below.

Actually, it is possible to answer this question rigorously, thanks to the "virial theorem", which provides exact relations between $E(R)$ and the average kinetic and potential energies. Therefore, in this section, we shall prove this theorem and discuss its physical consequences. The results obtained, furthermore, are completely general and

can be applied, not only to the molecular ion H_2^+, but also to all other molecules. Before considering the virial theorem itself, we shall begin by establishing some results which we shall need later.

5-b. Some useful theorems

α. Euler's theorem

Recall that a function $f\ (x_1, x_2, ..., x_n)$ of several variables $x_1, x_2, ...\ x_n$ is said to be homogeneous of degree s if it is multiplied by λ^s when all the variables are multiplied by λ:

$$f(\lambda x_1, \lambda x_2, \ldots \lambda x_n) = \lambda^s f(x_1, x_2, \ldots x_n) \tag{61}$$

For example, the potential of a three-dimensional harmonic oscillator:

$$V(x, y, z) = \frac{1}{2} m\omega^2 (x^2 + y^2 + z^2) \tag{62}$$

is homogeneous of degree 2. Similarly, the electrostatic interaction energy of two particles:

$$\frac{e_a e_b}{r_{ab}} = \frac{e_a e_b}{\sqrt{(x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2}} \tag{63}$$

is homogeneous of degree -1.

Euler's theorem indicates that any function f which is homogeneous of degree s satisfies the identity:

$$\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = s f(x_1, ..., x_i, ..., x_n) \tag{64}$$

To prove this, we calculate the derivatives with respect to λ of both sides of (61). The left-hand side yields:

$$\sum_i \frac{\partial f}{\partial x_i}(\lambda x_1, ..., \lambda x_n) \times \frac{\partial}{\partial \lambda}(\lambda x_i) = \sum_i x_i \frac{\partial f}{\partial x_i}(\lambda x_1, ..., \lambda x_n) \tag{65}$$

and the right-hand side yields:

$$s\lambda^{s-1} f(x_1, ..., x_n) \tag{66}$$

If we set (65) equal to (66), with $\lambda = 1$, we obtain (64).

Euler's theorem can very easily be verified in examples (62) and (63).

β. The Hellman-Feynman theorem

Let $H(\lambda)$ be a Hermitian operator which depends on a real parameter λ, and $|\psi(\lambda)\rangle$ a normalized eigenvector of $H(\lambda)$ of eigenvalue $E(\lambda)$:

$$H(\lambda)|\psi(\lambda)\rangle = E(\lambda)|\psi(\lambda)\rangle \tag{67}$$

$$\langle\psi(\lambda)|\psi(\lambda)\rangle = 1 \tag{68}$$

The Hellmann-Feynman theorem indicates that:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}E(\lambda) = \langle\psi(\lambda)|\frac{\mathrm{d}}{\mathrm{d}\lambda}H(\lambda)|\psi(\lambda)\rangle \tag{69}$$

This relation can be proven as follows. According to (67) and (68), we have:

$$E(\lambda) = \langle\psi(\lambda)|H(\lambda)|\psi(\lambda)\rangle \tag{70}$$

If we differentiate this relation with respect to λ, we obtain:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}E(\lambda) = {} & \langle\psi(\lambda)|\frac{\mathrm{d}}{\mathrm{d}\lambda}H(\lambda)|\psi(\lambda)\rangle \\ & + \left[\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\psi(\lambda)|\right]H(\lambda)|\psi(\lambda)\rangle + \langle\psi(\lambda)|H(\lambda)\left[\frac{\mathrm{d}}{\mathrm{d}\lambda}|\psi(\lambda)\rangle\right]\end{aligned} \tag{71}$$

that is, using (67) and the adjoint relation [$H(\lambda)$ is Hermitian, hence $E(\lambda)$ is real]:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}E(\lambda) = {} & \langle\psi(\lambda)|\frac{\mathrm{d}}{\mathrm{d}\lambda}H(\lambda)|\psi(\lambda)\rangle \\ & + E(\lambda)\left\{\left[\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\psi(\lambda)|\right]|\psi(\lambda)\rangle + \langle\psi(\lambda)|\left[\frac{\mathrm{d}}{\mathrm{d}\lambda}|\psi(\lambda)\rangle\right]\right\}\end{aligned} \tag{72}$$

On the right-hand side, the expression inside curly brackets is the derivative of $\langle\psi(\lambda)|\psi(\lambda)\rangle$, which is zero since $|\psi(\lambda)\rangle$ is normalized; we therefore find (69).

γ. Average value of the commutator [H, A] in an eigenstate of H

Let $|\psi\rangle$ be a normalized eigenvector of the Hermitian operator H, of eigenvalue E. For any operator A:

$$\langle\psi|[H, A]|\psi\rangle = 0 \tag{73}$$

since, as $H|\psi\rangle = E|\psi\rangle$ and $\langle\psi|H = E\langle\psi|$:

$$\langle\psi|(HA - AH)|\psi\rangle = E\langle\psi|A|\psi\rangle - E\langle\psi|A|\psi\rangle = 0 \tag{74}$$

5-c. The virial theorem applied to molecules

α. The potential energy of the system

Consider an arbitrary molecule composed of N nuclei and Q electrons. We shall denote by $\mathbf{r}_k^n (k = 1, 2, \ldots N)$ the classical positions of the nuclei, and by $\mathbf{r}_i^e$ and $\mathbf{p}_i^e (i = 1, 2, \ldots Q)$ the classical positions and momenta of the electrons. The components of these vectors will be written x_k^n, y_k^n, z_k^n, etc.

We shall use the Born-Oppenheimer approximation, considering the $\mathbf{r}_k^n$ as given classical parameters. In the quantum mechanical calculation, only the $\mathbf{r}_i^e$ and $\mathbf{p}_i^e$ become operators, $\mathbf{R}_i^e$ and $\mathbf{P}_i^e$. We must therefore solve the eigenvalue equation:

$$H(\mathbf{r}_1^n, \ldots, \mathbf{r}_N^n)|\psi(\mathbf{r}_1^n, \ldots, \mathbf{r}_N^n)\rangle = E(\mathbf{r}_1^n, \ldots, \mathbf{r}_N^n)|\psi(\mathbf{r}_1^n, \ldots, \mathbf{r}_N^n)\rangle \tag{75}$$

of a Hamiltonian H which depends on the parameters $\mathbf{r}_1^n$, ..., $\mathbf{r}_N^n$ and which acts in the state space of the electrons. The expression for H can be written:

$$H = T_e + V(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) \tag{76}$$

where T_e is the kinetic energy operator of the electrons:

$$T_e = \sum_{i=1}^{Q} \frac{1}{2m} (\mathbf{P}_i^e)^2 \tag{77}$$

and $V(\mathbf{r}_1^n, ..., \mathbf{r}_N^n)$ is the operator obtained by replacing the $\mathbf{r}_i^e$ by the operators $\mathbf{R}_i^e$ in the expression for the classical potential energy. The latter is the sum of the repulsion energy V_{ee} between the electrons, the attraction energy V_{en} between the electrons and the nuclei, and the repulsion energy V_{nn} between the nuclei, so that:

$$V(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) = V_{ee} + V_{en}(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) + V_{nn}(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) \tag{78}$$

Actually, since V_{nn} depends only on the $\mathbf{r}_k^n$ and does not involve the $\mathbf{R}_i^e$, V_{nn} is a number and not an operator acting in the state space of the electrons. The only effect of V_{nn} is therefore to shift all the energies equally, since equation (75) is equivalent to:

$$H_e(\mathbf{r}_1^n, ..., \mathbf{r}_N^n)|\psi(\mathbf{r}_1^n, ..., \mathbf{r}_N^n)\rangle = E_e(\mathbf{r}_1^n, ..., \mathbf{r}_N^n)|\psi(\mathbf{r}_1^n, ..., \mathbf{r}_N^n)\rangle \tag{79}$$

where:

$$H_e(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) = T_e + V_{ee} + V_{en}(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) = H - V_{nn}(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) \tag{80}$$

and where the electronic energy E_e is related to the total energy E by:

$$E_e(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) = E(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) - V_{nn}(\mathbf{r}_1^n, ..., \mathbf{r}_N^n) \tag{81}$$

We can apply Euler's theorem to the classical potential energy, since it is a homogeneous function of degree -1 of the *set* of electronic and nuclear coordinates. Since the operators $\mathbf{R}_i^e$ all commute with each other, we get the relation between the quantum mechanical operators:

$$\sum_{k=1}^{N} \mathbf{r}_k^n \cdot \boldsymbol{\nabla}_k^n V + \sum_{i=1}^{Q} \mathbf{R}_i^e \cdot \boldsymbol{\nabla}_i^e V = -V \tag{82}$$

where $\boldsymbol{\nabla}_k^n$ and $\boldsymbol{\nabla}_i^e$ denote the operators obtained by substitution of the $\mathbf{R}_i^e$ for the $\mathbf{r}_i^e$ in the gradients with respect to $\mathbf{r}_k^n$ and $\mathbf{r}_i^e$ in the classical expression for the potential energy. Relation (82) will serve as the foundation of our proof of the virial theorem.

β. Proof of the virial theorem

We apply (73) to the special case in which:

$$A = \sum_{i=1}^{Q} \mathbf{R}_i^e \cdot \mathbf{P}_i^e \tag{83}$$

To do so, we find the commutator of H with A :

$$\left[H, \sum_{i=1}^{Q} \mathbf{R}_i^e \cdot \mathbf{P}_i^e\right] = \sum_{i=1}^{Q} \sum_{x,y,z} \{[H,\, X_i^e] P_{xi}^e + X_i^e [H,\, P_{xi}^e]\}$$
$$= i\hbar \sum_{i=1}^{Q} \left\{ -\frac{(\mathbf{P}_i^e)^2}{m} + \mathbf{R}_i^e \cdot \boldsymbol{\nabla}_i^e V \right\} \tag{84}$$

(we have used the commutation relations of a function of the momentum with the position, or vice versa; *cf.* Complement B_{II}, § 4-c). The first term inside the curly brackets is proportional to the kinetic energy T_e. According to (82), the second term is equal to:

$$-V - \sum_{k=1}^{N} \mathbf{r}_k^n \cdot \boldsymbol{\nabla}_k^n V \tag{85}$$

Consequently, relation (73) yields:

$$2\langle T_e \rangle + \langle V \rangle + \sum_{k=1}^{N} \mathbf{r}_k^n \cdot \langle \boldsymbol{\nabla}_k^n V \rangle = 0 \tag{86}$$

that is, since the Hamiltonian H depends on the parameters $\mathbf{r}_k^n$ only through V:

$$2\langle T_e \rangle + \langle V \rangle = -\sum_{k=1}^{N} \mathbf{r}_k^n \cdot \langle \boldsymbol{\nabla}_k^n H \rangle \tag{87}$$

The components $\mathbf{r}_k^n$ here play a role analogous to that of the parameter λ in (69). Application of the Hellmann-Feynman theorem to the right-hand side of equation (87) then gives:

$$2\langle T_e \rangle + \langle V \rangle = -\sum_{k=1}^{N} \mathbf{r}_k^n \cdot \boldsymbol{\nabla}_k^n E(\mathbf{r}_1^n,\, ...,\, \mathbf{r}_k^n,\, ...,\, \mathbf{r}_N^n) \tag{88}$$

Furthermore, we obviously have:

$$\langle T_e \rangle + \langle V \rangle = E(\mathbf{r}_1^n,\, ...,\, \mathbf{r}_N^n) \tag{89}$$

We can then easily find from (88) and (89):

$$\boxed{\begin{aligned} \langle T_e \rangle &= -E - \sum_{k=1}^{N} \mathbf{r}_k^n \cdot \boldsymbol{\nabla}_k^n E \\ \langle V \rangle &= 2E + \sum_{k=1}^{N} \mathbf{r}_k^n \cdot \boldsymbol{\nabla}_k^n E \end{aligned}} \tag{90}$$

Thus, we obtain a very simple result: the virial theorem applied to molecules. It enables us to calculate the average kinetic and potential energies if we know the variation of the total energy with respect to the positions of the nuclei.

Comment:

The total electronic energy E_e and the electronic potential energy $\langle V_e \rangle$ are also related by:

$$\boxed{\langle V_e \rangle = 2E_e + \sum_{k=1}^{N} \mathbf{r}_k^n \cdot \boldsymbol{\nabla}_k^n E_e} \tag{91}$$

This relation can be proven by substituting (81) and the explicit expression for V_{nn} in terms of the $\mathbf{r}_k^n$ into the second relation of (90). However, it is simpler to note that the electronic potential energy $V_e = V_{ee} + V_{en}$, like the total potential energy V, is a homogeneous function of degree -1 of the coordinates of the system of particles. Consequently, the preceding arguments apply to H_e as well as to H, and we can simultaneously replace E by E_e and V by V_e in both relations (90).

γ. A special case: the diatomic molecule

When the number N of nuclei is equal to two, the energies depend only on the internuclear distance R. This further simplifies the expression for the virial theorem, which becomes:

$$\boxed{\begin{aligned} \langle T_e \rangle &= -E - R\frac{\mathrm{d}E}{\mathrm{d}R} \\ \langle V \rangle &= 2E + R\frac{\mathrm{d}E}{\mathrm{d}R} \end{aligned}} \tag{92}$$

Since E depends on the nuclear coordinates only through R, we have:

$$\frac{\partial E}{\partial x_k^n} = \frac{\mathrm{d}E}{\mathrm{d}R}\frac{\partial R}{\partial x_k^n} \tag{93}$$

and, consequently:

$$\sum_{k=1,2}\sum_{x,y,z} x_k^n \frac{\partial E}{\partial x_k^n} = \frac{\mathrm{d}E}{\mathrm{d}R}\sum_{k=1,2}\sum_{x,y,z} x_k^n \frac{\partial R}{\partial x_k^n} \tag{94}$$

Now, the distance R between the nuclei is a homogeneous function of degree 1 of the coordinates of the nuclei. Application of Euler's theorem to this function enables us to replace the double summation appearing on the right-hand side of (94) by R, and we finally obtain:

$$\sum_{k=1,2} \mathbf{r}_k^n \cdot \boldsymbol{\nabla}_k^n E = R\frac{\mathrm{d}E}{\mathrm{d}R} \tag{95}$$

When this result is substituted into (90), it gives relations (92).

In (92) as in (90), we can replace E by E_e and V by V_e.

5-d. Discussion

α. The chemical bond is due to a lowering of the electronic potential energy

Let E_∞ be the value of the total energy E of the system when the various nuclei are infinitely far apart. If it is possible to form a stable molecule by moving the nuclei

closer together, there must exist a certain relative arrangement of these nuclei for which the total energy E passes through a minimum $E_0 < E_\infty$. For the corresponding values of $\mathbf{r}_k^n$, we then have:

$$\boldsymbol{\nabla}_k^n E = \mathbf{0} \tag{96}$$

Relations (90) then indicate that, for this equilibrium position, the kinetic and potential energies are equal to:

$$\begin{aligned} \langle T_e \rangle_0 &= -E_0 \\ \langle V \rangle_0 &= 2E_0 \end{aligned} \tag{97}$$

Furthermore, when the nuclei are infinitely far from each other, the system is composed of a certain number of atoms or ions without mutual interactions (the energy no longer depends on the $\mathbf{r}_k^n$). For each of these subsystems, the virial theorem indicates that $\langle T_e \rangle = -E$, $\langle V \rangle = 2E$ and, for the system as a whole, we must therefore also have:

$$\begin{aligned} \langle T_e \rangle_\infty &= -E_\infty \\ \langle V \rangle_\infty &= 2E_\infty \end{aligned} \tag{98}$$

Subtracting (98) from (97) then gives:

$$\begin{aligned} \langle T_e \rangle_0 - \langle T_e \rangle_\infty &= -(E_0 - E_\infty) > 0 \\ \langle V \rangle_0 - \langle V \rangle_\infty &= 2(E_0 - E_\infty) < 0 \end{aligned} \tag{99}$$

The formation of a stable molecule is therefore always accompanied by an increase in the kinetic energy of the electrons and a decrease in the total potential energy. The electronic potential energy must, furthermore, decrease even more since the average value $\langle V_{nn} \rangle$ (the repulsion between the nuclei), which is zero at infinity, is always positive. It is therefore a lowering of the potential energy of the electrons $\langle V_{ee} + V_{en} \rangle$ that is responsible for the chemical bond. At equilibrium, this lowering must outweigh the increase in $\langle T_e \rangle$ and $\langle V_{nn} \rangle$.

β. The special case of the H_2^+ ion

(*i*) Application of the virial theorem to the approximate variational energy.

We return to the study of the variation of $\langle T_e \rangle$ and $\langle V \rangle$ for the H_2^+ ion. We shall begin by examining the predictions of the variational model of § 2, which led to the approximate expressions (52) and (53). From the second of these relations, we deduce that:

$$\Delta T_e = \langle T_e \rangle - \langle T_e \rangle_\infty = \frac{1}{1+S}(A - 2S\,E_I) \tag{100}$$

Since S is always greater than $A/2E_I$ (*cf.* Fig. 3), this calculation would tend to indicate that ΔT_e is always negative. This appears, moreover, in Figure 13, where the dashed lines represent the variations of the approximate expressions (52) and (53). In particular, we see that, according to the variational calculation, ΔT_e is negative at equilibrium ($\rho \simeq 2.5$) and ΔV is positive. These results are both incorrect, according to (99). We see here the limits of a variational calculation, which gives an acceptable value for the total energy $\langle T_e + V \rangle$, but not for $\langle T_e \rangle$ and $\langle V \rangle$ separately. These latter average values depend too strongly on the wave function.

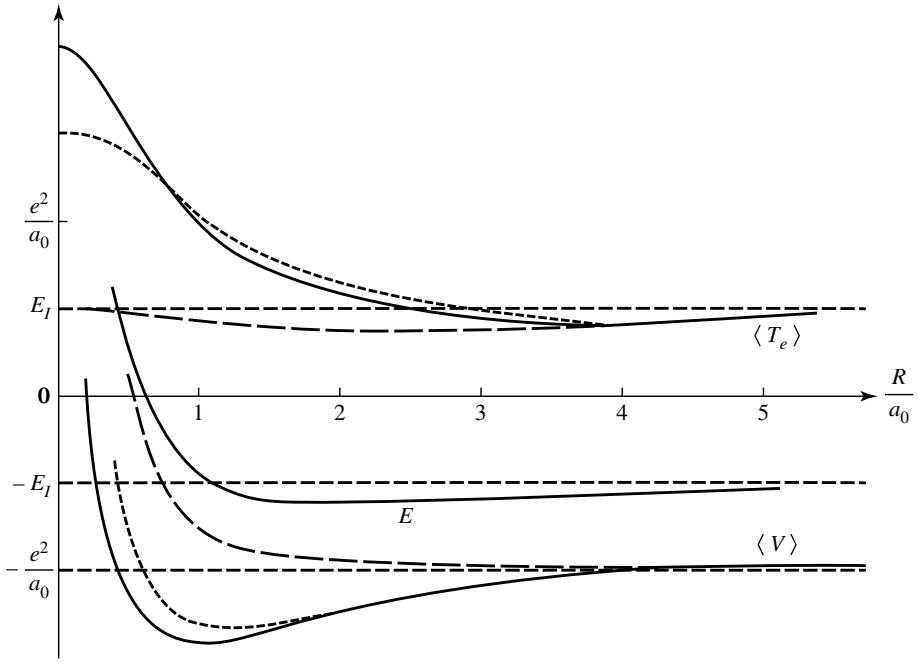


Figure 13: The electronic kinetic energy $\langle T_e \rangle$ and the potential energy $\langle V \rangle$ of the H_2^+ ion as functions of $\rho = R/a_0$ (for purposes of comparison, we have also shown the total energy $E = \langle T_e \rangle + \langle V \rangle$).
. solid lines: the exact values (the chemical bond is due to the fact that $\langle V \rangle$ decreases a little faster than $\langle T_e \rangle$ increases).
. long dashes: the average values calculated from the bonding wave function given by the simple variational method of § 2.
. short dashes: the values obtained by the application of the virial theorem to the energy given by the same variational calculation.

The virial theorem enables us, without having to resort to the rigorous calculation mentioned in § 1-c, to obtain a much better approximation for $\langle T_e \rangle$ and $\langle V \rangle$. All we need to do is apply the exact relations (92) to the energy E calculated by the variational method. We should expect an acceptable result, since the variational approximation is now used only to supply the total energy E. The values thus obtained for $\langle T_e \rangle$ and $\langle V \rangle$ are represented by short dashed lines in Figure 13. For purposes of comparison, we have shown in solid lines the exact values of $\langle T_e \rangle$ and $\langle V \rangle$ (obtained by application of the virial theorem to the solid-line curve of Figure 2). First of all, we see that for $\rho = 2.5$, the curve in short dashed lines indicates, as expected, that ΔT_e is positive and ΔV is negative. In addition, the general shape of these curves reproduces rather well that of the solid-line curves. As long as $\rho \gtrsim 1.5$, the virial theorem applied to the variational energy does give values which are very close to reality. This represents a considerable improvement over the direct calculation of the average values in the approximate states.

(*ii*) Behavior of $\langle T \rangle$ and $\langle V \rangle$

The solid-line curves of Figure 13 (the exact curves) show that $\langle T_e \rangle \longrightarrow 4E_I$ and $\langle V \rangle \longrightarrow +\infty$ when $R \longrightarrow 0$. Indeed, when $R = 0$, we have the equivalent of a He^+ ion for which the

electronic kinetic energy is $4E_I$. The divergence of $\langle V \rangle$ is due to the term $\langle V_{nn} \rangle = e^2/R$, which becomes infinite when $R \longrightarrow 0$ (the electronic potential energy $\langle V_e \rangle = \langle V \rangle - e^2/R$ remains finite and approaches $-8E_I$, which is indeed its value in the He^+ ion).

The behavior for large R deserves a more detailed discussion. We have seen above (§ 3-b) that the energy E_- of the ground state behaves, for $R \gg a_0$, like:

$$E_- \simeq -E_I - \frac{a}{R^4} \tag{101}$$

where a is a constant which is proportional to the polarizability of the hydrogen atom. By substituting this result into formulas (92), we obtain:

$$\begin{aligned} \langle T_e \rangle &\simeq E_I - \frac{3a}{R^4} \\ \langle V \rangle &\simeq -2E_I + \frac{2a}{R^4} \end{aligned} \tag{102}$$

When R decreases from a very large value, $\langle T_e \rangle$ begins by decreasing with $1/R^4$ from its asymptotic value E_I, and $\langle V \rangle$ begins by increasing from $-2E_I$. These variations then change sign (this must be so since $\langle T_e \rangle_0$ is larger than $\langle T_e \rangle_\infty$ and $\langle V \rangle_0$ is smaller than $\langle V \rangle_\infty$): as R continues to decrease (*cf.* Fig. 13), $\langle T_e \rangle$ passes through a minimum and then increases until it reaches its value $4E_I$ for $R = 0$. As for the potential energy $\langle V \rangle$, it passes through a maximum, then decreases, passes through a minimum, and then approaches infinity when $R \longrightarrow 0$. How can we interpret these variations?

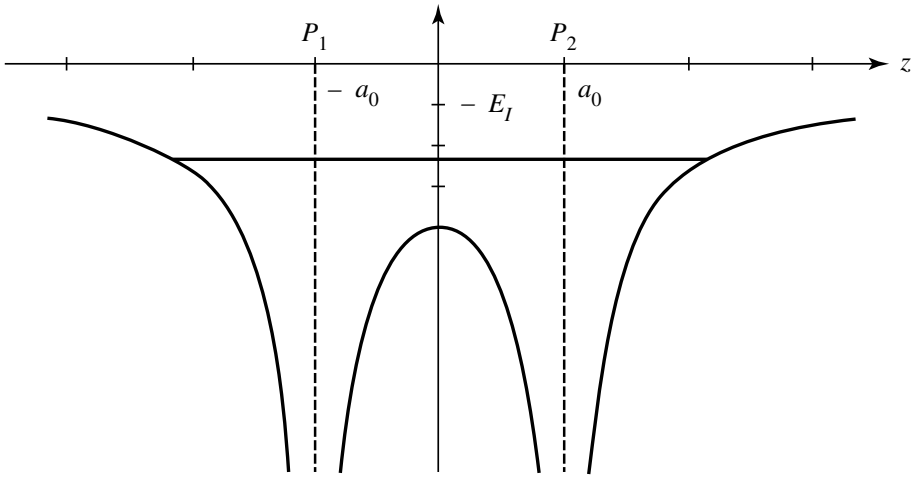


Figure 14: Variation of the potential energy V_e of the electron subjected to the simultaneous attraction of the two protons P_1 and P_2 as one moves along the line P_1P_2. In the bonding state, the wave function is concentrated in the region between P_1 and P_2, and the electron benefits simultaneously from the attraction of both protons.

As we have noted several times, the non-diagonal elements H_{12} and H_{21} of determinant (21) approach zero exponentially when $R \longrightarrow \infty$. We can therefore argue only in terms of H_{11} or H_{22} in discussing the variation of the energy of the H_2^+ ion at large internuclear distances. The problem is then reduced to the study of the perturbation of a hydrogen atom centered at P_2 by the electric field of the proton P_1. This field tends to distort the electronic orbital by

stretching it in the P_1 direction (*cf.* Fig. 6). Consequently, the wave function extends into a larger volume. According to Heisenberg's uncertainty relations, this allows the kinetic energy to decrease; this can explain the behavior of $\langle T_e \rangle$ for large R.

Arguing in terms of H_{22}, we can also explain the asymptotic behavior of $\langle V \rangle$. The discussion of § 3-b showed that, for $R \gg a_0$, the polarization of the hydrogen atom situated at P_2 makes its interaction energy $\left\langle -\frac{e^2}{r_1} + \frac{e^2}{R} \right\rangle$ with P_1 slightly negative (proportional to $-1/R^4$). If $\langle V \rangle$ is positive, it is because the potential energy $\left\langle -\frac{e^2}{r_2} \right\rangle$ of the atom at P_2 increases more rapidly, when P_1 is brought closer to P_2, than $\left\langle -\frac{e^2}{r_1} + \frac{e^2}{R} \right\rangle$ decreases. This increase in $\left\langle -\frac{e^2}{r_2} \right\rangle$ is due to the fact that the attraction of P_1 moves the electron slightly away from P_2 and carries it into regions of space in which the potential created by P_2 is less negative.

For $R \simeq R_0$ (the equilibrium position of the H_2^+ ion), the wave function of the bonding state is highly localized in the region between the two protons. The decrease in $\langle V \rangle$ (despite the increase in e^2/R) is due to the fact that the electron is in a region of space in which it benefits simultaneously from the attraction of both protons. This lowers its potential energy (*cf.* Fig. 14). This combined attraction of the two protons also leads to a decrease in the spatial extension of the electronic wave function, which is concentrated in the intermediate region. This is why, for R close to R_0, $\langle T_e \rangle$ increases when R decreases.

References and suggestions for further reading (H_2^+ ion, H_2 molecule, nature of the chemical bond, etc.):

Pauling (12.2); Pauling and Wilson (1.9), Chaps. XII and XIII; Levine (12.3), Chaps. 13 and 14; Karplus and Porter (12.1), Chap. 5, § 6; Slater (1.6), Chaps. 8 and 9; Eyring et al (12.5), Chaps. XI and XII; Coulson (12.6), Chap. IV; Wahl (12.13).

Complement H_{XI}

Exercises

1. A particle of mass m is placed in an infinite one-dimensional well of width a:

$V(x) = 0 \quad \text{for} \quad 0 \leq x \leq a$

$V(x) = +\infty \quad \text{everywhere else}$

It is subject to a perturbation W of the form:

$$W(x) = aw_0 \; \delta\left(x - \frac{a}{2}\right)$$

where w_0 is a real constant with the dimensions of an energy.

a. Calculate, to first order in w_0, the modifications induced by $W(x)$ to the energy levels of the particle.

b. Actually, the problem is exactly soluble. Setting $k = \sqrt{2mE/\hbar^2}$, show that the possible values of the energy are given by one of the two equations $\sin(ka/2) = 0$ or $\tan(ka/2) = -\hbar^2 k/maw_0$ (as in exercise 2 of Complement K_I, watch out for the discontinuity of the derivative of the wave function at $x = a/2$).

Discuss the results obtained with respect to the sign and size of w_0. In the limit $w_0 \longrightarrow 0$, show that one obtains the results of the preceding question.

2. Consider a particle of mass m placed in an infinite two-dimensional potential well of width a (*cf.* Complement G_{II}) :

$V(x, y) = 0 \quad \text{if} \quad 0 \leq x \leq a \quad \text{and} \quad 0 \leq y \leq a$

$V(x, y) = +\infty \quad \text{everywhere else}$

This particle is also subject to a perturbation W described by the potential:

$W(x, y) = w_0 \quad \text{for} \quad 0 \leq x \leq \frac{a}{2} \quad \text{and} \quad 0 \leq y \leq \frac{a}{2}$

$W(x, y) = 0 \quad \text{everywhere else.}$

a. Calculate, to first order in w_0, the perturbed energy of the ground state.

b. Same question for the first excited state. Give the corresponding wave functions to zeroth order in w_0.

3. A particle of mass m, constrained to move in the xOy plane, has a Hamiltonian:

$$H_0 = \frac{P_x^2}{2m} + \frac{P_y^2}{2m} + \frac{1}{2}m\omega^2(X^2 + Y^2)$$

(a two-dimensional harmonic oscillator, of angular frequency ω). We want to study the effect on this particle of a perturbation W given by:

$$W = \lambda_1 W_1 + \lambda_2 W_2$$

where λ_1 and λ_2 are constants, and the expressions for W_1 and W_2 are:

$$W_1 = m\omega^2 XY$$

$$W_2 = \hbar\omega \left(\frac{L_z^2}{\hbar^2} - 2 \right)$$

(L_z is the component along Oz of the orbital angular momentum of the particle).

In the perturbation calculations, consider only the corrections to first order for the energies and to zeroth order for the state vectors.

a. Indicate without calculations the eigenvalues of H_0, their degrees of degeneracy and the associated eigenvectors.

In what follows, consider only the second excited state of H_0, of energy $3\hbar\omega$ and which is three-fold degenerate.

b. Calculate the matrices representing the restrictions of W_1 and W_2 to the eigensubspace of the eigenvalue $3\hbar\omega$ of H_0.

c. Assume $\lambda_2 = 0$ and $\lambda_1 \ll 1$.

Calculate, using perturbation theory, the effect of the term $\lambda_1 W_1$ on the second excited state of H_0.

d. Compare the results obtained in *c* with the limited expansion of the exact solution, to be found with the help of the methods described in Complement H_V (normal vibrational modes of two coupled harmonic oscillators).

e. Assume $\lambda_2 \ll \lambda_1 \ll 1$. Considering the results of question *c* to be a new unperturbed situation, calculate the effect of the term $\lambda_2 W_2$.

f. Now assume that $\lambda_1 = 0$ and $\lambda_2 \ll 1$.

Using perturbation theory, find the effect of the term $\lambda_2 W_2$ on the second excited state of H_0.

g. Compare the results obtained in *f* with the exact solution, which can be found from the discussions of Complement D_{VI}.

h. Finally, assume that $\lambda_1 \ll \lambda_2 \ll 1$. Considering the results of question *f* to be a new unperturbed situation, calculate the effect of the term $\lambda_1 W_1$.

4. Consider a particle P of mass μ constrained to move in the xOy plane in a circle centered at O with fixed radius ρ (a two-dimensional rotator). The only variable of the system is the angle $\alpha = (Ox, OP)$, and the quantum state of the particle is defined by the wave function $\psi(\alpha)$ (which represents the probability amplitude of finding the particle at the point of the circle fixed by the angle α). At each point of the circle, $\psi(\alpha)$ can take on only one value, so that:

$$\psi(\alpha + 2\pi) = \psi(\alpha)$$

$\psi(\alpha)$ is normalized if:

$$\int_0^{2\pi} |\psi(\alpha)|^2 \, d\alpha = 1$$

a. Consider the operator $M = \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}\alpha}$. Is M Hermitian? Calculate the eigenvalues and normalized eigenfunctions of M. What is the physical meaning of M?

b. The kinetic energy of the particle can be written:

$$H_0 = \frac{M^2}{2\mu\rho^2}$$

Calculate the eigenvalues and eigenfunctions of H_0. Are the energies degenerate?

c. At $t = 0$, the wave function of the particle is $N\cos^2\alpha$ (where N is a normalization coefficient). Discuss the localization of the particle on the circle at a subsequent time t.

d. Assume that the particle has a charge q and that it interacts with a uniform electric field $\mathcal{E}$ parallel to Ox. We must therefore add to the Hamiltonian H_0 the perturbation:

$$W = -q\mathcal{E}\,\rho\cos\alpha$$

Calculate the new wave function of the ground state to first order in $\mathcal{E}$. Determine the proportionality coefficient χ (the linear suceptibility) between the electric dipole parallel to Ox acquired by the particle and the field $\mathcal{E}$.

e. Consider, for the ethane molecule CH_3 – CH_3, a rotation of one CH_3 group relative to the other about the straight line joining the two carbon atoms.

To a first approximation, this rotation is free, and the Hamiltonian H_0 introduced in *b* describes the rotational kinetic energy of one of the CH_3 groups relative to the other ($2\mu\rho^2$ must, however, be replaced by λI, where I is the moment of inertia of the CH_3 group with respect to the rotational axis and λ is a constant). To take account of the electrostatic interaction energy between the two CH_3 groups, we add to H_0 a term of the form:

$$W = b\cos 3\alpha$$

where b is a real constant.

Give a physical justification for the α-dependence of W. Calculate the energy and wave function of the new ground state (to first order in b for the wave function and to second order for the energy). Give a physical interpretation of the result.

5. Consider a system of angular momentum $\mathbf{J}$. We confine ourselves in this exercise to a three-dimensional subspace, spanned by the three kets $|+1\rangle$, $|0\rangle$, $|-1\rangle$, common eigenstates of $\mathbf{J}^2$ (eigenvalue $2\hbar^2$) and J_z (eigenvalues $+\hbar$, 0, $-\hbar$). The Hamiltonian H_0 of the system is:

$$H_0 = aJ_z + \frac{b}{\hbar}J_z^2$$

where a and b are two positive constants, which have the dimensions of an angular frequency.

a. What are the energy levels of the system? For what value of the ratio b/a is there degeneracy?

b. A static field $\mathbf{B}_0$ is applied in a direction $\mathbf{u}$ with polar angles θ and φ. The interaction with $\mathbf{B}_0$ of the magnetic moment of the system:

$$\mathbf{M} = \gamma \mathbf{J}$$

(γ: the gyromagnetic ratio, assumed to be negative) is described by the Hamiltonian:

$$W = \omega_0 J_u$$

where $\omega_0 = -\gamma |\mathbf{B}_0|$ is the Larmor angular frequency in the field $\mathbf{B}_0$, and J_u is the component of $\mathbf{J}$ in the $\mathbf{u}$ direction:

$$J_u = J_z \cos\theta + J_x \sin\theta \cos\varphi + J_y \sin\theta \sin\varphi$$

Write the matrix that represents W in the basis of the three eigenstates of H_0.

c. Assume that $b = a$ and that the $\mathbf{u}$ direction is parallel to Ox. We also have $\omega_0 \ll a$.

Calculate the energies and eigenstates of the system, to first order in ω_0 for the energies and to zeroth order for the eigenstates.

d. Assume that $b = 2a$ and that we again have $\omega_0 \ll a$, the direction of $\mathbf{u}$ now being arbitrary.

In the $\{|+1\rangle, |0\rangle, |-1\rangle\}$ basis, what is the expansion of the ground state $|\psi_0\rangle$ of $H_0 + W$, to first order in ω_0?

Calculate the average value $\langle \mathbf{M} \rangle$ of the magnetic moment $\mathbf{M}$ of the system in the state $|\psi_0\rangle$. Are $\langle \mathbf{M} \rangle$ and $\mathbf{B}_0$ parallel?

Show that one can write:

$$\langle M_i \rangle = \sum_j \chi_{ij} B_j$$

with $i, j = x, y, z$. Calculate the coefficients χ_{ij} (the components of the susceptibility tensor).

6. Consider a system formed by an electron spin $\mathbf{S}$ and two nuclear spins $\mathbf{I}_1$ and $\mathbf{I}_2$ ($\mathbf{S}$ is, for example, the spin of the unpaired electron of a paramagnetic diatomic molecule, and $\mathbf{I}_1$ and $\mathbf{I}_2$ are the spins of the two nuclei of this molecule).

Assume that $\mathbf{S}$, $\mathbf{I}_1$, $\mathbf{I}_2$ are all spin 1/2's. The state space of the three-spin system is spanned by the eight orthonormal kets $|\varepsilon_S, \varepsilon_1, \varepsilon_2\rangle$, common eigenvectors of S_z, I_{1z}, I_{2z}, with respective eigenvalues $\varepsilon_S \hbar/2$, $\varepsilon_1 \hbar/2$, $\varepsilon_2 \hbar/2$ (with $\varepsilon_S = \pm$, $\varepsilon_1 = \pm$, $\varepsilon_2 = \pm$). For example, the ket $|+, -, +\rangle$ corresponds to the eigenvalues $+\hbar/2$ for S_z, $-\hbar/2$ for I_{1z}, and $+\hbar/2$ for I_{2z}.

a. We begin by neglecting any coupling of the three spins. We assume, however, that they are placed in a uniform magnetic field $\mathbf{B}$ parallel to Oz. Since the gyromagnetic ratios of $\mathbf{I}_1$ and $\mathbf{I}_2$ are equal, the Hamiltonian H_0 of the system can be written:

$$H_0 = \Omega S_z + \omega I_{1z} + \omega I_{2z}$$

where Ω and ω are real, positive constants, proportional to $|\mathbf{B}|$. Assume $\Omega > 2\omega$.

What are the possible energies of the three-spin system and their degrees of degeneracy? Draw the energy diagram.

b. We now take coupling of the spins into account by adding the Hamiltonian:

$$W = a\mathbf{S} \cdot \mathbf{I}_1 + a\mathbf{S} \cdot \mathbf{I}_2$$

where a is a real, positive constant (the direct coupling of $\mathbf{I}_1$ and $\mathbf{I}_2$ is negligible).

What conditions must be satisfied by ε_S, ε_1, ε_2, ε'_S, ε'_1, ε'_2 for $a\mathbf{S} \cdot \mathbf{I}_1$ to have a non-zero matrix element between $|\varepsilon_S, \varepsilon_1, \varepsilon_2\rangle$ and $|\varepsilon'_S, \varepsilon'_1, \varepsilon'_2\rangle$? Same question for $a\mathbf{S} \cdot \mathbf{I}_2$.

c. Assume that:

$$a\hbar^2 \ll \hbar\Omega, \hbar\omega$$

so that W can be treated like a perturbation with respect to H_0. To first order in W, what are the eigenvalues of the total Hamiltonian $H = H_0 + W$? To zeroth order in H, what are the eigenstates of H? Draw the energy diagram.

d. Using the approximation of the preceding question, determine the Bohr frequencies which can appear in the evolution of $\langle S_x \rangle$ when the coupling W of the spins is taken into account.

In an E.P.R. (Electronic Paramagnetic Resonance) experiment, the frequencies of the observed resonance lines are equal to the preceding Bohr frequencies. What is the shape of the E.P.R. spectrum observed for the three-spin system? How can the coupling constant a be determined from this spectrum?

e. Now assume that the magnetic field $\mathbf{B}$ is zero, so that $\Omega = \omega = 0$. The Hamiltonian then reduces to W.

α. Let $\mathbf{I} = \mathbf{I}_1 + \mathbf{I}_2$ be the total nuclear spin. What are the eigenvalues of $\mathbf{I}^2$ and their degrees of degeneracy? Show that W has no matrix elements between eigenstates of $\mathbf{I}^2$ of different eigenvalues.

β. Let $\mathbf{J} = \mathbf{S} + \mathbf{I}$ be the total spin. What are the eigenvalues of $\mathbf{J}^2$ and their degrees of degeneracy? Determine the energy eigenvalues of the three-spin system and their degrees of degeneracy. Does the set $\{\mathbf{J}^2, J_z\}$ form a C.S.C.O.? Same question for $\{\mathbf{I}^2, \mathbf{J}^2, J_z\}$.

7. Consider a nucleus of spin $I = 3/2$, whose state space is spanned by the four vectors $|m\rangle$ ($m = +3/2, +1/2, -1/2, -3/2$), common eigenvectors of $\mathbf{I}^2$ (eigenvalue $15\hbar^2/4$) and I_z (eigenvalue $m\hbar$).

This nucleus is placed at the coordinate origin in a non-uniform electric field derived from a potential $U(x, y, z)$. The directions of the axes are chosen such that, at the origin:

$$\frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial z} = \frac{\partial^2 U}{\partial z \partial x} = 0$$

Recall that U satisfies Laplace's equation:

$$\Delta U = 0$$

We shall assume that the interaction Hamiltonian between the electric field gradient at the origin and the electric quadrupole moment of the nucleus can be written:

$$H_0 = \frac{qQ}{2I(2I-1)} \frac{1}{\hbar^2} \left[a_x I_x^2 + a_y I_y^2 + a_z I_z^2 \right]$$

where q is the electron charge, Q is a constant with the dimensions of a surface and proportional to the quadrupole moment of the nucleus, and:

$$a_x = \left(\frac{\partial^2 U}{\partial x^2} \right)_0 \quad ; \quad a_y = \left(\frac{\partial^2 U}{\partial y^2} \right)_0 \quad ; \quad a_z = \left(\frac{\partial^2 U}{\partial z^2} \right)_0$$

(the index 0 indicates that the derivatives are evaluated at the origin).

a. Show that, if U is symmetric with respect to revolution about Oz, H_0 has the form:

$$H_0 = A[3I_z^2 - I(I+1)]$$

where A is a constant to be specified. What are the eigenvalues of H_0, their degrees of degeneracy and the corresponding eigenstates?

b. Show that, in the general case, H_0 can be written:

$$H_0 = A[3I_z^2 - I(I+1)] + B(I_+^2 + I_-^2)$$

where A and B are constants, to be expressed in terms of a_x and a_y.

What is the matrix which represents H_0 in the $\{|m\rangle\}$ basis? Show that it can be broken down into two 2×2 submatrices. Determine the eigenvalues of H_0 and their degrees of degeneracy, as well as the corresponding eigenstates.

c. In addition to its quadrupole moment, the nucleus has a magnetic moment $\mathbf{M} = \gamma \mathbf{I}$ (γ: the gyromagnetic ratio). Onto the electrostatic field is superposed a magnetic field $\mathbf{B}_0$, of arbitrary direction $\mathbf{u}$. We set $\omega_0 = -\gamma |\mathbf{B}_0|$.

What term W must be added to H_0 in order to take into account the coupling between $\mathbf{M}$ and $\mathbf{B}_0$? Calculate the energies of the system to first order in B_0.

d. Assume $\mathbf{B}_0$ to be parallel to Oz and weak enough for the eigenstates found in *b* and the energies to first order in ω_0 found in *c* to be good approximations.

What are the Bohr frequencies which can appear in the evolution of $\langle I_x \rangle$? Deduce from them the shape of the nuclear magnetic resonance spectrum which can be observed with a radiofrequency field oscillating along Ox.

8. A particle of mass m is placed in an infinite one-dimensional potential well of width a:

$$\begin{cases} V(x) = 0 \quad \text{for} \quad 0 \leq x \leq a \\ V(x) = +\infty \quad \text{elsewhere} \end{cases}$$

Assume that this particle, of charge $-q$, is subject to a uniform electric field $\mathcal{E}$, with the corresponding perturbation W being:

$$W = q\mathcal{E}\left(X - \frac{a}{2}\right)$$

a. Let ε_1 and ε_2 be the corrections to first- and second-order in $\mathcal{E}$ for the ground state energy.

Show that ε_1 is zero. Give the expression for ε_2 in the form of a series, whose terms are to be calculated in terms of $q, \mathcal{E}, m, a, \hbar$ (the integrals given at the end of the exercise can be used).

b. By finding upper bounds for the terms of the series for ε_2, give an upper bound for ε_2 (*cf.* § B-2-c of Chapter XI). Similarly, give a lower bound for ε_2, obtained by retaining only the principal term of the series.

With what accuracy do the two preceding bounds enable us to bracket the exact value of the shift ΔE in the ground state to second order in $\mathcal{E}$?

c. We now want to calculate the shift ΔE by using the variational method. Choose as a trial function:

$$\psi_\lambda(x) = \sqrt{\frac{a}{2}} \sin\left(\frac{\pi x}{a}\right)\left[1 + \lambda q\mathcal{E}\left(x - \frac{a}{2}\right)\right]$$

where λ is the variational parameter. Explain this choice of trial functions.

Calculate the average energy $\langle H\rangle(\lambda)$ of the ground state to second order in $\mathcal{E}$ [assuming the expansion of $\langle H\rangle(\lambda)$ to second order in $\mathcal{E}$ to be sufficient]. Determine the optimal value of λ. Find the result ΔE_{var} given by the variational method for the shift in the ground state to second order in $\mathcal{E}$.

By comparing ΔE_{var} with the results of b, evaluate the accuracy of the variational method applied to this example.

We give the integrals:

$$\frac{2}{a}\int_0^a \left(x - \frac{a}{2}\right)\sin\left(\frac{\pi x}{a}\right)\sin\left(\frac{2n\pi x}{a}\right)\mathrm{d}x = -\frac{16na}{\pi^2}\frac{1}{(1-4n^2)^2}$$

$n = 1, 2, 3, \ldots$

$$\frac{2}{a}\int_0^a \left(x - \frac{a}{2}\right)^2 \sin^2\left(\frac{\pi x}{a}\right)\mathrm{d}x = \frac{a^2}{2}\left(\frac{1}{6} - \frac{1}{\pi^2}\right)$$

$$\frac{2}{a}\int_0^a \left(x - \frac{a}{2}\right)\sin\left(\frac{\pi x}{a}\right)\cos\left(\frac{\pi x}{a}\right)\mathrm{d}x = -\frac{a}{2\pi}$$

For all the numerical calculations, take $\pi^2 = 9.87$.

9. We want to calculate the ground state energy of the hydrogen atom by the variational method, choosing as trial functions the spherically symmetrical functions $\varphi_\alpha(\mathbf{r})$ whose r-dependcnce is given by:

$$\begin{cases} \varphi_\alpha(r) = C\left(1 - \dfrac{r}{\alpha}\right) & \text{for} \quad r \leq \alpha \\ \varphi_\alpha(r) = 0 \quad \text{for} \quad r > \alpha \end{cases}$$

C is a normalization constant and α is the variational parameter.

a. Calculate the average value of the kinetic and potential energies of the electron in the state $|\varphi_\alpha\rangle$. Express the average value of the kinetic energy in terms of $\boldsymbol{\nabla}\varphi$, so as to avoid the "delta functions" which appear in $\Delta\varphi$ (since $\boldsymbol{\nabla}\varphi$ is discontinuous).

b. Find the optimal value α_0 of α. Compare α_0 with the Bohr radius a_0.

c. Compare the approximate value obtained for the ground state energy with the exact value $-E_1$.

10. We intend to apply the variational melhod to the determination of the energies of a particle of mass m in an infinite potential well:

$V(x) = 0 \qquad -a \leq x \leq a$

$V(x) = \infty \qquad$ everywhere else

a. We begin by approximating, in the interval $[-a, +a]$, the wave function of the ground state by the simplest even polynomial which goes to zero at $x = \pm a$:

$\psi(x) = a^2 - x^2 \quad$ for $\quad -a \leq x \leq a$

$\psi(x) = 0 \qquad$ everywhere else

(a variational family reduced to a single trial function).

Calculate the average value of the Hamiltonian H in this state. Compare the result obtained with the true value.

b. Enlarge the family of trial functions by choosing an even fourth-degree polynomial which goes to zero at $x = \pm a$:

$\psi_\alpha(x) = \left(a^2 - x^2\right)\left(a^2 - \alpha x^2\right) \quad$ for $\quad -a \leq x \leq a$

$\psi_\alpha(x) = 0 \qquad$ everywhere else

(a variational family depending on the real parameter α).

(α) Show that the average value of H in the state $\psi_\alpha(x)$ is:

$$\langle H\rangle(\alpha) = \frac{\hbar^2}{2ma^2}\frac{33\alpha^2 - 42\alpha + 105}{2\alpha^2 - 12\alpha + 42}$$

(β) Show that the values of α which minimize or maximize $\langle H \rangle(\alpha)$ are given by the roots of the equation:

$$13\alpha^2 - 98\alpha + 21 = 0$$

(γ) Show that one of the roots of this equation gives, when substituted into $\langle H \rangle(\alpha)$, a value of the ground state energy that is much more precise than the one obtained in a.

(δ) What other eigenvalue is approximated when the second root of the equation obtained in b-β is used? Could this have been expected? Evaluate the precision of this determination.

c. Explain why the simplest polynomial which permits the approximation of the first excited state wave function is $x(a^2 - x^2)$.

What approximate value is then obtained for the energy of this state?

Chapter XII

An application of perturbation theory: the fine and hyperfine structure of hydrogen

A Introduction 1232

B Additional terms in the Hamiltonian 1233

B-1 The fine-structure Hamiltonian 1233

B-2 Magnetic interactions related to proton spin: the hyperfine Hamiltonian 1237

C The fine structure of the $n = 2$ level 1238

C-1 Statement of the problem 1238

C-2 Matrix representation of the fine-structure Hamiltonian W_f inside the $n = 2$ level 1239

C-3 Results: the fine structure of the $n = 2$ level 1243

D The hyperfine structure of the $n = 1$ level 1246

D-1 Statement of the problem 1246

D-2 Matrix representation of W_{hf} in the $1s$ level 1247

D-3 The hyperfine structure of the $1s$ level 1249

E The Zeeman effect of the $1s$ ground state hyperfine structure 1251

E-1 Statement of the problem 1251

E-2 The weak-field Zeeman effect 1253

E-3 The strong-field Zeeman effect 1257

E-4 The intermediate-field Zeeman effect 1261

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

A. Introduction

The most important forces inside atoms are Coulomb electrostatic forces. We took them into account in Chapter VII by choosing as the hydrogen atom Hamiltonian:

$$H_0 = \frac{\mathbf{P}^2}{2\mu} + V(R) \tag{A-1}$$

The first term represents the kinetic energy of the atom in the center of mass frame (μ is the reduced mass). The second term:

$$V(R) = \frac{q^2}{4\pi\varepsilon_0}\frac{1}{R} = -\frac{e^2}{R} \tag{A-2}$$

represents the electrostatic interaction energy between the electron and the proton (q is the electron charge). In § C of Chapter VII, we calculated in detail the eigenstates and eigenvalues of H_0.

Actually, expression (A-1) is only approximate: it does not take any relativistic effects into account. In particular, all the magnetic effects related to the electron spin are ignored. Moreover, we have not introduced the proton spin and the corresponding magnetic interactions. The error is, in reality, very small, since the hydrogen atom is a weakly relativistic system (recall that, in the Bohr model, the velocity v in the first orbit $n = 1$ satisfies $v/c = e^2/\hbar c = 1/137 \ll 1$). In addition, the magnetic moment of the proton is very small.

However, the considerable accuracy of spectroscopic experiments makes it possible to observe effects that cannot be explained in terms of the Hamiltonian (A-1). Therefore, we shall take into account the corrections we have just mentioned by writing the complete hydrogen atom Hamiltonian in the form:

$$H = H_0 + W \tag{A-3}$$

where H_0 is given by (A-1) and where W represents all the terms neglected thus far. Since W is much smaller than H_0, it is possible to calculate its effects by using the perturbation theory presented in Chapter XI. This is what we propose to do in this chapter. We shall show that W is responsible for a "fine structure", as well as for a "hyperfine structure" of the various energy levels calculated in Chapter VII. Furthermore, these structures can be measured experimentally with very great accuracy (the hyperfine structure of the $1s$ ground state of the hydrogen atom is currently known with 12 significant figures; the ratio between certain atomic frequencies has been measured with 18 digits!). We shall also consider, in this chapter and its complements, the influence of an external static magnetic or electric field on the various levels of the hydrogen atom (the Zeeman effect and the Stark effect).

This chapter actually has two goals. On the one hand, we want to use a concrete and realistic case to illustrate the general stationary perturbation theory discussed in the preceding chapter. On the other hand, this study, which bears on one of the most fundamental systems of physics (the hydrogen atom), brings out certain concepts which are basic to atomic physics. For example, § B is devoted to a thorough discussion of various relativistic and magnetic corrections. This chapter, while not indispensable for the study of the last two chapters, presents concepts fundamental to atomic physics.

B. Additional terms in the Hamiltonian

The first problem to be solved obviously consists of finding the expression for W.

B-1. The fine-structure Hamiltonian

B-1-a. The Dirac equation in the weakly relativistic domain

In Chapter IX, we mentioned that the spin appears naturally when we try to establish an equation for the electron which satisfies both the postulates of special relativity and those of quantum mechanics. Such an equation exists: it is the *Dirac equation*, which makes it possible to account for numerous phenomena (electron spin, the fine structure of hydrogen, etc.) and to predict the existence of positrons.

The most rigorous way of obtaining the expression for the relativistic corrections [appearing in the term W of (A-3)] therefore consists of first writing the Dirac equation for an electron placed in the potential $V(r)$ created by the proton (considered to be infinitely heavy and motionless at the coordinate origin). One then looks for its limiting form when the system is weakly relativistic, as is the case for the hydrogen atom. We then recognize that the description of the electron state must include a two-component spinor (*cf.* Chap. IX, § C-1). The spin operators S_x, S_y, S_z, introduced in Chapter IX then appear naturally. Finally, we obtain an expression such as (A-3) for the Hamiltonian H, in which W appears in the form of a power series expansion in v/c which we can evaluate.

It is out of the question here to study the Dirac equation, or to establish its form in the weakly relativistic domain. We shall confine ourselves to giving the first terms of the power series expansion in v/c of W and their interpretation.

$$H = m_e c^2 + \underbrace{\frac{\mathbf{P}^2}{2m_e} + V(R)}_{H_0} \underbrace{- \frac{\mathbf{P}^4}{8m_e^3 c^2}}_{W_{mv}} + \underbrace{\frac{1}{2m_e^2 c^2}\frac{1}{R}\frac{\mathrm{d}V(R)}{\mathrm{d}R}\mathbf{L}\cdot\mathbf{S}}_{W_{SO}} + \underbrace{\frac{\hbar^2}{8m_e^2 c^2}\Delta V(R)}_{W_D} + \ldots \quad \text{(B-1)}$$

We recognize in (B-1) the rest-mass energy $m_e c^2$ of the electron (the first term) and the non-relativistic Hamiltonian H_0 (the second and third terms)[1]. The following terms are called fine structure terms.

Comment:

Note that it is possible to solve the Dirac equation exactly for an electron placed in a Coulomb potential. We thus obtain the energy levels of the hydrogen atom without having to make a limited power series expansion in v/c of the eigenstates and eigenvalues of H. The "perturbation" point of view we are adopting here is, however, very useful in bringing out the form and physical meaning of the various interactions which exist inside an atom. This will later permit a generalization to the case of many-electron atoms (for which we do not know how to write the equivalent of the Dirac equation).

[1] Expression (B-1) was obtained by assuming the proton to be infinitely heavy. This is why it is the mass m_e of the electron that appears, and not, as in (A-1), the reduced mass μ of the atom. As far as H_0 is concerned, the proton finite mass effect is taken into account by replacing m_e by μ. However, we shall neglect this effect in the subsequent terms of H, which are already corrections. It would, moreover, be difficult to evaluate, since the relativistic description of a system of two interacting particles poses serious problems [it is not sufficient to replace m_e by μ in the last terms of (B-1)].

B-1-b. Interpretation of the various terms of the fine-structure Hamiltonian

α. Variation of the mass with the velocity (W_{mv} term)

(*i*) The physical origin

The physical origin of the W_{mv} term is very simple. If we start with the relativistic expression for the energy of a classical particle of rest-mass m_e and momentum $\mathbf{p}$:

$$E = c\sqrt{\mathbf{p}^2 + m_e^2 c^2} \tag{B-2}$$

and perform a limited expansion of E in powers of $|\mathbf{p}|/m_e c$, we obtain:

$$E = m_e c^2 + \frac{\mathbf{p}^2}{2m_e} - \frac{\mathbf{p}^4}{8m_e^3 c^2} + ... \tag{B-3}$$

In addition to the rest-mass energy ($m_e c^2$) and the non-relativistic kinetic energy ($\mathbf{p}^2/2m_e$), we find the term $-\mathbf{p}^4/8m_e^3 c^2$, which appears in (B-1). This term represents the first energy correction, due to the relativistic variation of the mass with the velocity.

(*ii*) Order of magnitude

To evaluate the size of this correction, we shall calculate the order of magnitude of the ratio W_{mv}/H_0:

$$\frac{W_{mv}}{H_0} \simeq \frac{\dfrac{\mathbf{p}^4}{8m_e^3 c^2}}{\dfrac{\mathbf{p}^2}{2m_e}} = \frac{\mathbf{p}^2}{4m_e^2 c^2} = \frac{1}{4}\left(\frac{v}{c}\right)^2 \simeq \alpha^2 \simeq \left(\frac{1}{137}\right)^2 \tag{B-4}$$

since we have already mentioned that, for the hydrogen atom, $v/c \simeq \alpha$. Since $H_0 \simeq 10$ eV, we see that $W_{mv} \simeq 10^{-3}$ eV.

β. Spin-orbit coupling (W_{SO} term)

(*i*) The physical origin

The electron moves at a velocity $\mathbf{v} = \mathbf{p}/m_e$ in the electrostatic field $\mathbf{E}$ created by the proton. Special relativity indicates that there then appears, in the electron frame, a magnetic field $\mathbf{B}'$ given by:

$$\mathbf{B}' = -\frac{1}{c^2}\,\mathbf{v} \times \mathbf{E} \tag{B-5}$$

to first order in v/c. Since the electron possesses an intrinsic magnetic moment $\mathbf{M}_S = q\mathbf{S}/m_e$, it interacts with this field $\mathbf{B}'$. The corresponding interaction energy can be written:

$$W' = -\mathbf{M}_S \cdot \mathbf{B}' \tag{B-6}$$

Let us express W' more explicitly. The electrostatic field $\mathbf{E}$ appearing in (B-5) is equal to $-\frac{1}{q}\frac{\mathrm{d}V(r)}{\mathrm{d}r}\frac{\mathbf{r}}{r}$, where $V(r) = -\dfrac{e^2}{r}$ is the electrostatic energy of the electron. From this, we get:

$$\mathbf{B}' = \frac{1}{qc^2}\frac{1}{r}\frac{\mathrm{d}V(r)}{\mathrm{d}r}\frac{\mathbf{p}}{m_e} \times \mathbf{r} \tag{B-7}$$

In the corresponding quantum mechanical operator, there appears:

$$\mathbf{P} \times \mathbf{R} = -\mathbf{L} \tag{B-8}$$

Finally, we obtain:

$$W' = \frac{1}{m_e^2 c^2} \frac{1}{R} \frac{\mathrm{d}V(R)}{\mathrm{d}R} \, \mathbf{L} \cdot \mathbf{S} = \frac{e^2}{m_e^2 c^2} \frac{1}{R^3} \, \mathbf{L} \cdot \mathbf{S} \tag{B-9}$$

Thus we find, to within the factor[2] 1/2, the spin-orbit term W_{SO} which appears in (B-1). This term then represents the interaction of the magnetic moment of the electron spin with the magnetic field "seen" by the electron because of its motion in the electrostatic field of the proton.

(*ii*) Order of magnitude

Since $\mathbf{L}$ and $\mathbf{S}$ are of the order of $\hbar$, we have:

$$W_{SO} \simeq \frac{e^2}{m_e^2 c^2} \frac{\hbar^2}{R^3} \tag{B-10}$$

Let us compare W_{SO} with H_0, which is of the order of e^2/R:

$$\frac{W_{SO}}{H_0} \simeq \frac{\dfrac{e^2 \hbar^2}{m_e^2 c^2 R^3}}{\dfrac{e^2}{R}} = \frac{\hbar^2}{m_e^2 c^2 R^2} \tag{B-11}$$

R is of the order of the Bohr radius, $a_0 = \hbar^2/m_e e^2$. Consequently:

$$\frac{W_{SO}}{H_0} \simeq \frac{e^4}{\hbar^2 c^2} = \alpha^2 = \left(\frac{1}{137}\right)^2 \tag{B-12}$$

γ. *The Darwin term* W_D

(*i*) The physical origin

In the Dirac equation, the interaction between the electron and the Coulomb field of the nucleus is "local"; it only depends on the value of the field at the electron position $\mathbf{r}$. However, the non-relativistic approximation (the series expansion in v/c) leads, for the two-component spinor which describes the electron state, to an equation in which the interaction between the electron and the field has become non-local. The electron is then affected by all the values taken on by the field in a domain centered at the point $\mathbf{r}$, and whose size is of the order of the Compton wavelength $\hbar/m_e c$ of the electron. This is the origin of the correction represented by the Darwin term.

To understand this more precisely, assume that the potential energy of the electron, instead of being equal to $V(\mathbf{r})$, is given by an expression of the form:

$$\int \mathrm{d}^3\rho \; f(\boldsymbol{\rho}) \, V(\mathbf{r} + \boldsymbol{\rho}) \tag{B-13}$$

[2]It can be shown that the factor 1/2 is due to the fact that the motion of the electron about the proton is not rectilinear. The electron spin therefore rotates with respect to the laboratory reference frame (Thomas precession: see Jackson (7.5) section 11-8, Omnes (16.13) chap. 4 § 2, or Bacry (10.31) Chap. 7 § 5-d).

where $f(\boldsymbol{\rho})$ is a function whose integral is equal to 1, which only depends on $|\boldsymbol{\rho}|$, and which takes on significant values only inside a volume of the order of $(\hbar/m_e c)^3$, centered at $\boldsymbol{\rho} = \mathbf{0}$.

If we neglect the variation of $V(\mathbf{r})$ over a distance of the order of $\hbar/m_e c$, we can replace $V(\mathbf{r}+\boldsymbol{\rho})$ by $V(\mathbf{r})$ in (B-13) and take $V(\mathbf{r})$ outside the integral, which is then equal to 1. (B-13) reduces, in this case, to $V(\mathbf{r})$.

A better approximation consists of replacing, in (B-13), $V(\mathbf{r}+\boldsymbol{\rho})$ by its Taylor series expansion in the neighborhood of $\boldsymbol{\rho} = \mathbf{0}$. The zeroth-order term gives $V(\mathbf{r})$. The first-order term is zero because of the spherical symmetry of $f(\boldsymbol{\rho})$. The second-order term involves the second derivatives of the potential energy $V(\mathbf{r})$ at the point $\mathbf{r}$ and quadratic functions of the components of $\boldsymbol{\rho}$, weighted by $f(\boldsymbol{\rho})$ and integrated over $\mathrm{d}^3\rho$. This leads to a result of the order of

$$(\hbar/m_e c)^2 \Delta V(\mathbf{r})$$

It is therefore easy to accept the idea that this second-order term should be the Darwin term.

(ii) Order of magnitude

Replacing $V(R)$ by $-e^2/R$, we can write the Darwin term in the form:

$$-e^2 \frac{\hbar^2}{8m_e^2 c^2}\, \Delta\left(\frac{1}{R}\right) = \frac{\pi e^2 \hbar^2}{2m_e^2 c^2}\, \delta(\mathbf{R}) \tag{B-14}$$

(we have used the expression for the Laplacian of $1/R$ given by formula (61) of Appendix II).

When we take the average value of (B-14) in an atomic state, we find a contribution equal to:

$$\frac{\pi e^2 \hbar^2}{2m_e^2 c^2}\, |\psi(\mathbf{0})|^2$$

where $\psi(\mathbf{0})$ is the value of the wave function at the origin. The Darwin term therefore affects only the s electrons, which are the only ones for which $\psi(\mathbf{0}) \neq 0$ (*cf.* Chap. VII, § C-4-c). The order of magnitude of $|\psi(\mathbf{0})|^2$ can be obtained by taking the integral of the square of the modulus of the wave function over a volume of the order of a_0^3 (where a_0 is the Bohr radius) to be equal to 1. Thus we obtain:

$$|\psi(\mathbf{0})|^2 \simeq \frac{1}{a_0^3} = \frac{m_e^3 e^6}{\hbar^6} \tag{B-15}$$

which gives the order of magnitude of the Darwin term:

$$W_D \simeq \frac{\pi e^2 \hbar^2}{2m_e^2 c^2}\, |\psi(\mathbf{0})|^2 \simeq m_e c^2 \frac{e^8}{\hbar^4 c^4} = m_e c^2 \alpha^4 \tag{B-16}$$

Since $H_0 \simeq m_e c^2 \alpha^2$, we again see that:

$$\frac{W_D}{H_0} \simeq \alpha^2 = \left(\frac{1}{137}\right)^2 \tag{B-17}$$

Thus, all the fine structure terms are about 10^4 times smaller than the non-relativistic Hamiltonian of Chapter VII.

B-2. Magnetic interactions related to proton spin: the hyperfine Hamiltonian

B-2-a. Proton spin and magnetic moment

Thus far, we have considered the proton to be a physical point of mass M_p and charge $q_p = -q$. Actually, the proton, like the electron, is a spin 1/2 particle. We shall denote by $\mathbf{I}$ the corresponding spin observable.

With the spin $\mathbf{I}$ of the proton is associated a magnetic moment $\mathbf{M}_I$. However, the gyromagnetic ratio is different from that of the electron:

$$\mathbf{M}_I = g_p \mu_n \, \mathbf{I}/\hbar \tag{B-18}$$

where μ_n is the *nuclear Bohr magneton*:

$$\mu_n = \frac{q_p \hbar}{2M_p} \tag{B-19}$$

and the factor g_p, for the proton, is equal to: $g_p \simeq 5{,}585$. Because of the presence of M_p (the proton mass) in the denominator of (B-19), μ_n is close to 2 000 times smaller than the Bohr magneton μ_B (recall that $\mu_B = q\hbar/2m_e$). Although the angular momenta of the proton and the electron are the same, nuclear magnetism, because of the mass difference, is much less important than electronic magnetism. The magnetic interactions due to the proton spin $\mathbf{I}$ are therefore very weak.

B-2-b. The magnetic hyperfine Hamiltonian W_{hf}

The electron moves, therefore, not only in the electrostatic field of the proton, but also in the magnetic field created by $\mathbf{M}_I$. When we introduce the corresponding vector potential into the Schrödinger equation[3], we find that we must add to the Hamiltonian (B-1) an additional series of terms for which the expression is (*cf.* Complement $\mathrm{A_{XII}}$):

$$W_{hf} = -\frac{\mu_0}{4\pi}\left\{\frac{q}{m_e R^3}\mathbf{L}\cdot\mathbf{M}_I + \frac{1}{R^3}\left[3(\mathbf{M}_S\cdot\mathbf{n})(\mathbf{M}_I\cdot\mathbf{n}) - \mathbf{M}_S\cdot\mathbf{M}_I\right] + \frac{8\pi}{3}\mathbf{M}_S\cdot\mathbf{M}_I\,\delta(\mathbf{R})\right\} \tag{B-20}$$

$\mathbf{M}_S$ is the spin magnetic moment of the electron, and $\mathbf{n}$ is the unit vector of the straight line joining the proton to the electron (Fig. 1).

We shall see that W_{hf} introduces energy shifts which are small compared to those created by W_f. This is why W_{hf} is called the "hyperfine structure Hamiltonian".

B-2-c. Interpretation of the various terms of W_{hf}

The first term of W_{hf} represents the interaction of the nuclear magnetic moment $\mathbf{M}_I$ with the magnetic field $(\mu_0/4\pi)q\mathbf{L}/m_e r^3$ created at the proton by the rotation of the electronic charge.

The second term represents the dipole-dipole interaction between the electronic and nuclear magnetic moments: the interaction of the magnetic moment of the electron spin with the magnetic field created by $\mathbf{M}_I$ (*cf.* Complement $\mathrm{B_{XI}}$) or vice versa.

[3]Since the hyperfine interactions are very small corrective terms, they can be found using the non-relativistic Schrödinger equation.

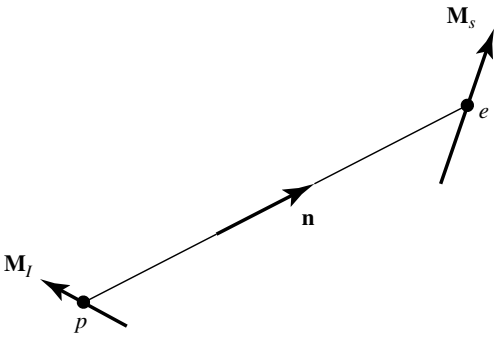


Figure 1: Relative disposition of the magnetic moments $\mathbf{M}_I$ and $\mathbf{M}_S$ of the proton and the electron; $\mathbf{n}$ is the unit vector on the line joining the two particles.

Finally, the last term, also called Fermi's "contact term", arises from the singularity at $r = 0$ of the field created by the magnetic moment of the proton. In reality, the proton is not a point. It can be shown (*cf.* Complement A_{XII}) that the magnetic field inside the proton does not have the same form as the one created outside by $\mathbf{M}_I$ (and which enters into the dipole-dipole interaction). The contact term describes the interaction of the magnetic moment of the electron spin with the magnetic field inside the proton (the "delta" function expresses the fact that this contact term exists, as its name indicates, only when the wave functions of the electron and proton overlap).

B-2-d. Orders of magnitude

It can easily be shown that the order of magnitude of the first two terms of W_{hf} is:

$$\frac{q^2\hbar^2}{m_e M_p R^3}\frac{\mu_0}{4\pi} = \frac{e^2\hbar^2}{m_e M_p c^2}\frac{1}{R^3} \tag{B-21}$$

By using (B-10), we see that these terms are about 2 000 times smaller than W_{SO}.

As for the last term of (B-20), it is also 2 000 times smaller than the Darwin term, which also contains a $\delta(\mathbf{R})$ function.

C. The fine structure of the $n = 2$ level

C-1. Statement of the problem

C-1-a. Degeneracy of the $n = 2$ level

We saw in Chapter VII that the energy of the hydrogen atom depends only on the quantum number n. The $2s$ ($n = 2$, $l = 0$) and $2p$ ($n = 2$, $l = 1$) states therefore have the same energy, equal to:

$$-\frac{E_I}{4} = -\frac{1}{8}\mu c^2\alpha^2$$

If the spins are ignored, the $2s$ subshell is composed of a single state, and the $2p$ subshell of three distinct states which differ by their eigenvalue $m_L\hbar$ of the component L_z of the orbital angular momentum $\mathbf{L}$ ($m_L = 1, 0, -1$). Because of the existence of electron and proton spins, the degeneracy of the $n = 2$ level is higher than the value calculated in

Chapter VII. The components S_z and I_z of the two spins can each take on two values: $m_S = \pm 1/2$, $m_I = \pm 1/2$. One possible orthonormal basis in the $n = 2$ level is given by the kets:

$$\left\{ \left| n = 2 \ ; \ l = 0 \ ; \ m_L = 0 \ ; \ m_S = \pm\frac{1}{2} \ ; \ m_I = \pm\frac{1}{2} \right\rangle \right\} \tag{C-1}$$

($2s$ subshell, of dimension 4) and:

$$\left\{ \left| n = 2 \ ; \ l = 1 \ ; \ m_L = -1, 0, +1 \ ; \ m_S = \pm\frac{1}{2} \ ; \ m_I = \pm\frac{1}{2} \right\rangle \right\} \tag{C-2}$$

($2p$ subshell, of dimension 12).
The $n = 2$ shell then has a total degeneracy equal to 16.

According to the results of Chapter XI (§ C), in order to calculate the effect of a perturbation W on the $n = 2$ level, it is necessary to diagonalize the 16×16 matrix representing the restriction of W to this level. The eigenvalues of this matrix are the first order corrections to the energy, and the corresponding eigenstates are the eigenstates of the Hamiltonian to zeroth order.

C-1-b. The perturbation Hamiltonian

In all of this section, we shall assume that no external field is applied to the atom. The difference W between the exact Hamiltonian H and the Hamiltonian H_0 of Chapter VII (§ C) contains fine structure terms, indicated in § B-1 above:

$$W_f = W_{mv} + W_{SO} + W_D \tag{C-3}$$

and hyperfine structure terms W_{hf}, introduced in § B-2. We thus have:

$$W = W_f + W_{hf} \tag{C-4}$$

Since W_f is close to 2 000 times larger than W_{hf} (*cf.* § B-2-d), we must obviously begin by studying the effect of W_f, before considering that of W_{hf}, on the $n = 2$ level. We shall see that the $n = 16$ degeneracy of this level is partially removed by W_f. The structure which appears in this way is called the "fine structure".

W_{hf} may then remove the remaining degeneracy of the fine structure levels and cause a "hyperfine structure" to appear inside each of these levels.

In this section (§ C), we shall confine ourselves to the study of the fine structure of the $n = 2$ level. The calculations can easily be generalized to other levels.

C-2. Matrix representation of the fine-structure Hamiltonian W_f inside the $n = 2$ level

C-2-a. General properties

The properties of W_f, as we shall see, enable us to show that the 16×16 matrix which represents it in the $n = 2$ level can be broken down into a series of square submatrices of smaller dimensions. This will considerably simplify the determination of the eigenvalues and eigenvectors of this matrix.

α. W_f does not act on the spin variables of the proton

We see from (B-1) that the fine structure terms do not depend on **I**. It follows that the proton spin can be ignored in the study of the fine structure (afterwards, we multiply by 2 all the degrees of degeneracy obtained). The dimension of the matrix to be diagonalized therefore falls from 16 to 8.

β. W_f does not connect the $2s$ and $2p$ subshells

Let us first prove that $\mathbf{L}^2$ commutes with W_f. The operator $\mathbf{L}^2$ commutes with the various components of **L**, with R ($\mathbf{L}^2$ acts only on the angular variables), with $\mathbf{P}^2$ [*cf.* formula (A-16) of Chapter VII], and with **S** ($\mathbf{L}^2$ does not act on the spin variables). $\mathbf{L}^2$ therefore commutes with W_{mv} (which is proportional to $\mathbf{P}^4$), with W_{SO} (which depends only on R, **L**, **S**), and with W_D (which depends only on R).

The $2s$ and $2p$ states are eigenstates of $\mathbf{L}^2$ with different eigenvalues (0 and $2\hbar^2$). Therefore, W_f, which commutes with $\mathbf{L}^2$, has no matrix elements between a $2s$ state and a $2p$ state. The 8×8 matrix representing W_f inside the $n = 2$ level can be broken down, consequently, into a 2×2 matrix relative to the $2s$ state and a 6×6 matrix relative to the $2p$ state:

$$(W_f)_{n=2} = \begin{array}{c} \\ 2s \\ \\ 2p \\ \\ \end{array} \begin{array}{c} \begin{array}{cc} 2s & \quad 2p \end{array} \\ \left(\begin{array}{c|c} \begin{array}{cc} \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot \end{array} & 0 \\ \hline 0 & \begin{array}{cccccc} \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{array} \end{array}\right) \end{array}$$

Comment:

The preceding property can also be considered to be a consequence of the fact that W_f is even. Under a reflection, **R** changes to $-\mathbf{R}$ ($R = |\mathbf{R}|$ remains unchanged), **P** to $-\mathbf{P}$, **L** to **L**, and **S** to **S**. It is then easy to see that W_f remains invariant. W_f therefore has no matrix elements between the $2s$ and $2p$ states, which are of opposite parity (*cf.* Complement $\mathrm{F_{II}}$).

C-2-b. Matrix representation of W_f in the $2s$ subshell

The dimension 2 of the $2s$ subspace is the result of the two possible values $m_s = \pm 1/2$ of S_z (since we are ignoring I_z for the moment).

W_{mv} and W_D do not depend on **S**. The matrices which represent these two operators in the $2s$ subspace are therefore multiples of the unit matrix, with proportionality coefficients equal, respectively, to the purely orbital matrix elements:

$$\left\langle n = 2\ ;\ l = 0\ ;\ m_L = 0 \left| -\frac{\mathbf{P}^4}{8m_e^3c^2} \right| n = 2\ ;\ l = 0\ ;\ m_L = 0 \right\rangle$$

and:

$$\left\langle n=2\ ;\ l=0\ ;\ m_L=0 \left| -\frac{\hbar^2}{8m_e^3c^2} q\,\Delta V(R) \right| n=2\ ;\ l=0\ ;\ m_L=0 \right\rangle$$

Since we know the eigenfunctions of H_0, the calculation of these matrix elements presents no theoretical difficulty. We find (*cf.* Complement B_{XII}):

$$\langle W_{mv}\rangle_{2s} = -\frac{13}{128} m_e c^2 \alpha^4 \tag{C-5}$$

$$\langle W_D\rangle_{2s} = -\frac{1}{16} m_e c^2 \alpha^4 \tag{C-6}$$

Finally, calculation of the matrix elements of W_{SO} involves "angular" matrix elements of the form $\langle l=0,\, m_L=0 \mid L_{x,y,z} \mid l=0, m_L=0\rangle$, which are zero because of the value $l=0$ of the quantum number l. Therefore:

$$\langle W_{SO}\rangle_{2s} = 0 \tag{C-7}$$

Thus, under the effect of the fine structure terms, the $2s$ subshell is shifted as a whole with respect to the position calculated in Chapter VII by an amount equal to $-5m_e c^2\alpha^4/128$.

C-2-c. Matrix representation of W_f in the $2p$ subshell

α. W_{mv} and W_D terms

The W_{mv} and W_D terms commute with the various components of $\mathbf{L}$, since $\mathbf{L}$ acts only on the angular variables and commutes with R and $\mathbf{P}^2$ (which depends on these variables only through $\mathbf{L}^2$; *cf.* chapter VII). $\mathbf{L}$ therefore commutes with W_{mv} and W_D. Consequently, W_{mv} and W_D are scalar operators with respect to the orbital variables (*cf.* Complement B_{VI}, § 5-b). Since W_{mv} and W_D do not act on the spin variables, it follows that the matrices which represent W_{mv} and W_D inside the $2p$ subspace are multiples of the unit matrix. The calculation of the proportionality coefficient is given in Complement B_{XII} and leads to:

$$\langle W_{mv}\rangle_{2p} = -\frac{7}{384} m_e c^2 \alpha^4 \tag{C-8}$$

$$\langle W_D\rangle_{2p} = 0 \tag{C-9}$$

The result (C-9) is due to the fact that W_D is proportional to $\delta(\mathbf{R})$ and can therefore have a non-zero average value only in an s state (for $l \geq 1$, the wave function is zero at the origin).

β. W_{SO} term

We must calculate the various matrix elements:

$$\left\langle n=2\ ;\ l=1\ ;\ s=\frac{1}{2}\ ;\ m'_L\ ;\ m'_S \mid \xi(R)\,\mathbf{L}\cdot\mathbf{S} \mid n=2\ ;\ l=1\ ;\ s=\frac{1}{2}\ ;\ m_L\ ;\ m_S \right\rangle \tag{C-10}$$

with:

$$\xi(R) = \frac{e^2}{2m_e^2c^2}\frac{1}{R^3} \tag{C-11}$$

If we use the $\{|\mathbf{r}\rangle\}$ representation, we can separate the radial part of matrix element (C-10) from the angular and spin parts. Thus we obtain:

$$\xi_{2p}\left\langle l=1\ ;\ s=\frac{1}{2}\ ;\ m'_L;m'_S \mid \mathbf{L}\cdot\mathbf{S} \mid l=1\ ;\ s=\frac{1}{2}\ ;\ m_L\ ;\ m_S\right\rangle \tag{C-12}$$

where ξ_{2p} is a number, equal to the radial integral:

$$\xi_{2p} = \frac{e^2}{2m_e^2c^2}\int_0^\infty \frac{1}{r^3}\left|R_{21}(r)\right|^2\ r^2\,\mathrm{d}r \tag{C-13}$$

Since we know the radial function $R_{21}(r)$ of the $2p$ state, we can calculate ξ_{2p}. We find (*cf.* Complement $\mathrm{B_{XII}}$):

$$\xi_{2p} = \frac{1}{48\hbar^2}m_ec^2\alpha^4 \tag{C-14}$$

The radial variables have therefore disappeared. According to (C-12), the problem is reduced to the diagonalization of the operator $\xi_{2p}\,\mathbf{L}\cdot\mathbf{S}$, which acts only on the angular and spin variables.

To represent the operator $\xi_{2p}\,\mathbf{L}\cdot\mathbf{S}$ by a matrix, several different bases can be chosen:

- first of all, the basis:

$$\left\{\left|l=1;s=\frac{1}{2};m_L;m_S\right\rangle\right\} \tag{C-15}$$

which we have used thus far and which is constructed from common eigenstates of $\mathbf{L}^2$, $\mathbf{S}^2$, L_z, S_z;

- or, introducing the total angular momentum:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{C-16}$$

the basis:

$$\left\{\left|l=1;s=\frac{1}{2};J;m_J\right\rangle\right\} \tag{C-17}$$

constructed from the eigenstates common to $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$, J_z. According to the results of chapter X, since $l=1$ and $s=1/2$, J can take on two values: $J=1+1/2=3/2$ and $J=1-1/2=1/2$. Furthermore, we know how to go from one basis to the other, thanks to the Clebsch-Gordan coefficients [formulas (36) of Complement $\mathrm{A_X}$].

We shall now show that the second basis (C-17) is better adapted than the first one to the problem which interests us here, since $\xi_{2p}\,\mathbf{L}\cdot\mathbf{S}$ is diagonal in the basis (C-17). To see this, we square both sides of (C-16). We find ($\mathbf{L}$ and $\mathbf{S}$ commute):

$$\mathbf{J}^2 = (\mathbf{L}+\mathbf{S})^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\,\mathbf{L}\cdot\mathbf{S} \tag{C-18}$$

which gives:

$$\xi_{2p}\,\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\xi_{2p}\left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2\right) \tag{C-19}$$

Each of the basis vectors (C-17) is an eigenstate of $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$; we thus have:

$$\xi_{2p}\,\mathbf{L}\cdot\mathbf{S}\left|l=1;s=\frac{1}{2};J;m_J\right\rangle = \frac{1}{2}\xi_{2p}\hbar^2\left[J(J+1)-2-\frac{3}{4}\right]\left|l=1;s=\frac{1}{2};J;m_J\right\rangle \tag{C-20}$$

We see from (C-20) that the eigenvalues of $\xi_{2p}\,\mathbf{L}\cdot\mathbf{S}$ depend only on J and not on m_J; they are equal to:

$$\frac{1}{2}\xi_{2p}\left[\frac{3}{4}-2-\frac{3}{4}\right]\hbar^2 = -\xi_{2p}\,\hbar^2 = -\frac{1}{48}m_e c^2\alpha^4 \tag{C-21}$$

for $J = 1/2$, and:

$$\frac{1}{2}\xi_{2p}\left[\frac{15}{4}-2-\frac{3}{4}\right]\hbar^2 = +\frac{1}{2}\xi_{2p}\,\hbar^2 = \frac{1}{96}m_e c^2\alpha^4 \tag{C-22}$$

for $J = 3/2$.

The six-fold degeneracy of the $2p$ level is therefore partially removed by W_{SO}. We obtain a four-fold degenerate level corresponding to $J = 3/2$, and a two-fold degenerate level corresponding to $J = 1/2$. The $(2J + 1)$-fold degeneracy of each J state is an essential degeneracy related to the rotation invariance of W_f.

Comments:

(*i*) In the $2s$ subspace ($l = 0$, $s = 1/2$), J can take on a single value, $J = 0+1/2 = 1/2$.

(*ii*) In the $2p$ subspace, W_{mv} and W_D are represented by multiples of the unit matrix. This property remains valid in any basis since the unit matrix is invariant under a change of basis. The choice of basis (C-17), required by the W_{SO} term, is therefore also adapted to the W_{mv} and W_D terms.

C-3. Results: the fine structure of the $n = 2$ level

C-3-a. Spectroscopic notation

In addition to the quantum numbers n, l (and s), the preceding discussion introduced the quantum number J on which the energy correction due to the spin-orbit coupling term depends.

For the $2s$ level, $J = 1/2$; for the $2p$ level, $J = 1/2$ or $J = 3/2$. The level associated with a set of values, n, l, J is generally denoted by adding an index J to the symbol representing the (n, l) subshell in spectroscopic notation (*cf.* Chap. VII, § C-4-b):

$$n\,l_J \tag{C-23}$$

where l stands for the letter s for $l = 0$, p for $l = 1$, d for $l = 2$, f for $l = 3$... Thus, the $n = 2$ level of the hydrogen atom gives rise to the $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$ levels.

C-3-b. Positions of the $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$ levels

By regrouping the results of § 2, we can now calculate the positions of the $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$ levels with respect to the "unperturbed" energy of the $n = 2$ level calculated in Chapter VII and equal to $-\mu c^2\alpha^2/8$.

According to the results of § 2.b, the $2s_{1/2}$ level is lowered by a quantity equal to:

$$-\frac{5}{128}m_e c^2\alpha^4 \tag{C-24}$$

According to the results of § 2.c, the $2p_{1/2}$ level is lowered by a quantity equal to:

$$\left(-\frac{7}{384}-\frac{1}{48}\right)m_e c^2\alpha^4 = -\frac{5}{128}m_e c^2\alpha^4 \tag{C-25}$$

Thus we see that the $2s_{1/2}$ and $2p_{1/2}$ levels have the same energy. According to the theory presented here, this degeneracy must be considered to be *accidental*, as opposed to the essential $(2J+1)$-fold degeneracy of each J level.

Finally, the $2p_{3/2}$ level is lowered by a quantity:

$$\left(-\frac{7}{384}+\frac{1}{96}\right)m_e c^2\alpha^4 = -\frac{1}{128}m_e c^2\alpha^4 \tag{C-26}$$

The preceding results are shown in Figure 2.

Comments:

(*i*) Only the spin-orbit coupling is responsible for the separation between the $2p_{1/2}$ and $2p_{3/2}$ levels, since W_{mv} and W_D shift the entire $2p$ level as a whole.

(*ii*) The hydrogen atom can go from the $2p$ state to the $1s$ state by emitting a Lyman α photon ($\lambda = 1\,216$ Å). The material presented in this chapter shows that, because of the spin-orbit coupling, the Lyman α line actually contains two neighboring lines[4], $2p_{1/2} \to 1s_{1/2}$ and $2p_{3/2} \to 1s_{1/2}$, separated by an energy difference equal to:

$$\frac{4}{128}m_e c^2\alpha^4 = \frac{1}{32}m_e c^2\alpha^4$$

When they are observed with a sufficient resolution, the lines of the hydrogen spectrum therefore present a "fine structure".

(*iii*) We see in Figure 2 that the two levels with the same J have the same energy. This result is not merely true to first order in W_f: it remains valid to all orders. The exact solution of the Dirac equation gives, for the energy of a level characterized by the quantum numbers n, l, s, J, the value:

$$E_{n,J} = m_e c^2\left[1+\alpha^2\left(n-J-\frac{1}{2}+\sqrt{(J+1/2)^2-\alpha^2}\right)^{-2}\right]^{-1/2} \tag{C-27}$$

[4] In the ground state, $l = 0$ and $s = 1/2$, so J can take on a single value $J = 1/2$. W_f therefore does not remove the degeneracy of the $1s$ state, and there is only one fine structure level, the $1s_{1/2}$ level. This is a special case, since the ground state is the only one for which l is necessarily zero. This is why we have chosen here to study the excited $n = 2$ level.

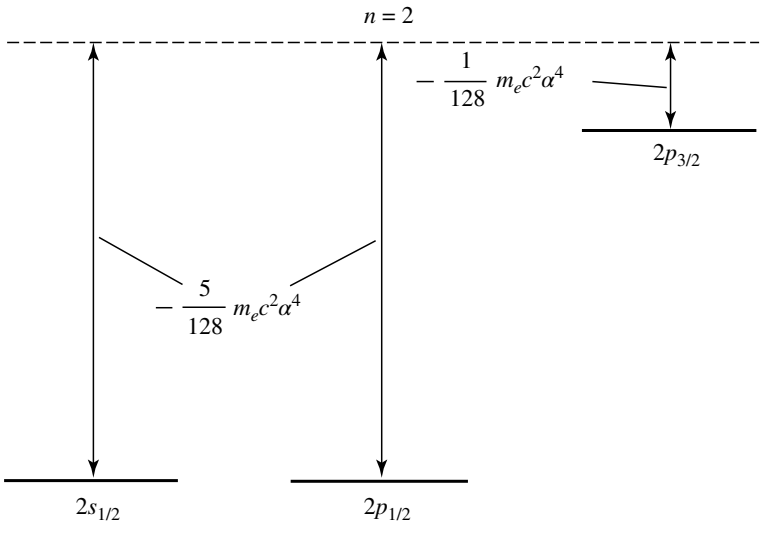


Figure 2: Fine structure of the $n = 2$ level of the hydrogen atom. Under the effect of the fine structure Hamiltonian W_f, the $n = 2$ level splits into three fine structure levels, written $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$. We have indicated the algebraic values of the shifts, calculated to first order in W_f. The shifts are the same for the $2s_{1/2}$ and $2p_{1/2}$ levels (a result which remains valid, moreover, to all orders in W_f). When we take into account the quantum mechanical nature of the electromagnetic field, we find that the degeneracy between the $2s_{1/2}$ and $2p_{1/2}$ levels is removed (the Lamb shift; see Figure 4).

We see that the energy depends only on n and J, and not on l.

If we make a limited expansion of formula (C-27) in powers of α, we obtain:

$$E_{n,J} = m_e c^2 - \frac{1}{2} m_e c^2 \alpha^2 \frac{1}{n^2} - \frac{m_e c^2}{2n^4} \left(\frac{n}{J + 1/2)} - \frac{3}{4} \right) \alpha^4 + ... \qquad \text{(C-28)}$$

The first term is the rest-mass-energy of the electron. The second term follows from the theory of Chapter VII. The third term gives the correction to first order in W_f calculated in this chapter.

(*iv*) Even in the absence of an external field and incident photons, a fluctuating electromagnetic field must be considered to exist in space (*cf.* Complement $\mathrm{K_V}$, § 3-d-δ). This phenomenon is related to the quantum mechanical nature of the electromagnetic field, which we have not taken into consideration here. The coupling of the atom with these fluctuations of the electromagnetic field removes the degeneracy between the $2s_{1/2}$ and $2p_{1/2}$ levels. The $2s_{1/2}$ level is raised with respect to the $2p_{1/2}$ level by a quantity called the "Lamb shift", which is of the order of 1 060 MHz (Fig. 4, page 1250).

The theoretical and experimental study of this phenomenon, which was discovered in 1949, has been the object of a great deal of research, leading to the development of modern quantum electrodynamics.

D. The hyperfine structure of the $n = 1$ level

It would now seem logical to study the effect of W_{hf} inside the fine structure levels $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$, in order to see if the interactions related to the proton spin $\mathbf{I}$ cause a hyperfine structure to appear in each of these levels. However, since W_f does not remove the degeneracy of the ground state $1s$, it is simpler to study the effect of W_{hf} on this state. The results obtained in this special case can easily be generalized to the $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$ levels.

D-1. Statement of the problem

D-1-a. The degeneracy of the $1s$ level

For the $1s$ level, there is no orbital degeneracy ($l = 0$). On the other hand, the S_z and I_z components of $\mathbf{S}$ and $\mathbf{I}$ can still take on two values: $m_S = \pm 1/2$ and $m_I = \pm 1/2$. The degeneracy of the $1s$ level is therefore equal to 4, and a possible basis in this level is given by the vectors:

$$\left\{ \left| n = 1; l = 0; m_L = 0; m_S = \pm\frac{1}{2}; m_I = \pm\frac{1}{2} \right\rangle \right\} \tag{D-1}$$

D-1-b. The $1s$ level has no fine structure

We shall show that the W_f term does not remove the degeneracy of the $1s$ level.

The W_{mv} and W_D terms do not act on m_S and m_I, and are represented in the $1s$ subspace by multiples of the unit matrix. We find (*cf.* Complement $\mathrm{B_{XII}}$):

$$\langle W_{mv} \rangle_{1s} = -\frac{5}{8} m_e c^2 \alpha^4 \tag{D-2}$$

$$\langle W_D \rangle_{1s} = \frac{1}{2} m_e c^2 \alpha^4 \tag{D-3}$$

Finally, calculation of the matrix elements of the W_{SO} term involves the "angular" matrix elements $\langle l = 0, m_L = 0 \mid L_{x,y,z} \mid l = 0, m_L = 0 \rangle$, which are obviously zero ($l = 0$); therefore:

$$\langle W_{SO} \rangle_{1s} = 0 \tag{D-4}$$

In conclusion, W_f merely shifts the $1s$ level as a whole by a quantity equal to:

$$\left(-\frac{5}{8} + \frac{1}{2} \right) m_e c^2 \alpha^4 = -\frac{1}{8} m_e c^2 \alpha^4 \tag{D-5}$$

without splitting the level. This result could have been foreseen: since $l = 0$ and $s = 1/2$, J can take on a single value, $J = 1/2$, and the $1s$ level therefore gives rise to only one fine structure level, $2s_{1/2}$.

Since the Hamiltonian W_f does not split the $1s$ level, we can now consider the effect of the W_{hf} term. To do so, we must first calculate the matrix which represents W_{hf} in the $1s$ level.

D-2. Matrix representation of W_{hf} in the $1s$ level

D-2-a. Terms other than the contact term

Let us show that the first two terms of W_{hf} [formula (B-20)] make no contribution.

Calculation of the contribution from the first term, $-\frac{\mu_0}{4\pi}\frac{q}{m_e R^3}\mathbf{L}\cdot\mathbf{M}_I$, leads to the "angular" matrix elements $\langle l=0; m_L=0 \mid \mathbf{L} \mid l=0, m_L=0\rangle$, which are obviously zero ($l=0$).

Similarly, it can be shown (*cf.* Complement $\mathrm{B_{XI}}$, § 3) that the matrix elements of the second term (the dipole-dipole interaction) are zero because of the spherical symmetry of the $1s$ state.

D-2-b. The contact term

The matrix elements of the last term of (B-20), that is, of the contact term, are of the form:

$$\left\langle n=1; l=0; m_L=0; m'_S; m'_I \right| -\frac{2\mu_0}{3}\,\mathbf{M}_S\cdot\mathbf{M}_I\;\delta(\mathbf{R})\left| n=1; l=0; m_L=0; m_S; m_I \right\rangle \tag{D-6}$$

If we go into the $\{|\mathbf{r}\rangle\}$ representation, we can separate the orbital and spin parts of this matrix element and put it in the form:

$$\mathcal{A}\,\langle m'_S; m'_I \mid \mathbf{I}\cdot\mathbf{S} \mid m_S; m_I\rangle \tag{D-7}$$

where $\mathcal{A}$ is a number given by:

$$\begin{aligned}\mathcal{A} &= \frac{q^2}{3\varepsilon_0 c^2}\frac{g_p}{m_e M_p}\,\langle n=1; l=0; m_L=0 \mid \delta(\mathbf{R}) \mid n=1; l=0; m_L=0\rangle \\ &= \frac{q^2}{3\varepsilon_0 c^2}\frac{g_p}{m_e M_p}\frac{1}{4\pi}\,|R_{10}(0)|^2 \\ &= \frac{4}{3}g_p\frac{m_e}{M_p}m_e c^2\alpha^4\left(1+\frac{m_e}{M_p}\right)^{-3}\frac{1}{\hbar^2}\end{aligned} \tag{D-8}$$

We have used the expressions relating $\mathbf{M}_S$ and $\mathbf{M}_I$ to $\mathbf{S}$ and $\mathbf{I}$ [*cf.* B-18], as well as the expression for the radial function $R_{10}(r)$ given in § C-4-c of Chapter VII[5].

The orbital variables have therefore completely disappeared, and we are left with a problem of two spin 1/2's, $\mathbf{I}$ and $\mathbf{S}$, coupled by an interaction of the form:

$$\mathcal{A}\,\mathbf{I}\cdot\mathbf{S} \tag{D-9}$$

where $\mathcal{A}$ is a constant.

[5]The factor $(1+m_e/M_p)^{-3}$ in (D-8) arises from the fact that it is the reduced mass μ which enters into $R_{10}(0)$. It so happens that, for the contact term, it is correct to take the nuclear finite mass effect into account in this way.

D-2-c. Eigenstates and eigenvalues of the contact term

To represent the operator $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$, we have thus far considered only the basis:

$$\left\{\left|s=\frac{1}{2};I=\frac{1}{2};m_S;m_I\right\rangle\right\} \tag{D-10}$$

formed by the eigenvectors common to $\mathbf{S}^2$, $\mathbf{I}^2$, S_z, I_z. We can also, by introducing the total angular momentum[6]:

$$\mathbf{F}=\mathbf{S}+\mathbf{I} \tag{D-11}$$

use the basis:

$$\left\{\left|s=\frac{1}{2};I=\frac{1}{2};F;m_F\right\rangle\right\} \tag{D-12}$$

formed by the eigenstates common to $\mathbf{S}^2$, $\mathbf{I}^2$, $\mathbf{F}^2$, F_z. Since $s=I=1/2$, F can take on only the two values $F=0$ and $F=1$. We can easily pass from one basis to the other by means of (B-22) and (B-23) of Chapter X.

The $\{|F,m_F\rangle\}$ basis is better adapted than the $\{|m_S,m_I\rangle\}$ basis to the study of the operator $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$, as this operator is represented in the $\{|F,m_F\rangle\}$ basis by a diagonal matrix (for the sake of simplicity, we do not explicitly write $s=1/2$ and $I=1/2$). This is true, since we obtain, from (D-11):

$$\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}=\frac{\mathcal{A}}{2}\left(\mathbf{F}^2-\mathbf{I}^2-\mathbf{S}^2\right) \tag{D-13}$$

It follows that the states $|F,m_F\rangle$ are eigenstates of $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$:

$$\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}|F,m_F\rangle=\frac{\mathcal{A}\hbar^2}{2}\left[F(F+1)-I(I+1)-S(S+1)\right]|F,m_F\rangle \tag{D-14}$$

We see from (D-14) that the eigenvalues depend only on F, and not on m_F. They are equal to:

$$\frac{\mathcal{A}\hbar^2}{2}\left[2-\frac{3}{4}-\frac{3}{4}\right]=\frac{\mathcal{A}\hbar^2}{4} \tag{D-15}$$

for $F=1$, and:

$$\frac{\mathcal{A}\hbar^2}{2}\left[0-\frac{3}{4}-\frac{3}{4}\right]=-\frac{3\mathcal{A}\hbar^2}{4} \tag{D-16}$$

for $F=0$.

[6]The total angular momentum is actually $\mathbf{F}=\mathbf{L}+\mathbf{S}+\mathbf{I}$, that is, $\mathbf{F}=\mathbf{J}+\mathbf{I}$. However, for the ground state, the orbital angular momentum is zero, so $\mathbf{F}$ reduces to (D-11).

The four-fold degeneracy of the $1s$ level is therefore partially removed by W_{hf}. We obtain a three-fold degenerate $F = 1$ level and a non-degenerate $F = 0$ level. The $(2F + 1)$-fold degeneracy of the $F = 1$ level is essential and is related to the invariance of W_{hf} under a rotation of the total system.

D-3. The hyperfine structure of the $1s$ level

D-3-a. Positions of the levels

Under the effect of W_f, the energy of the $1s$ level is lowered by a quantity $m_e c^2 \alpha^4/8$ with respect to the value $-\mu c^2 \alpha^2/2$ calculated in Chapter VII. W_{hf} then splits the $1s_{1/2}$ level into two hyperfine levels, separated by an energy $\mathcal{A}\hbar^2$ (Fig. 3). $\mathcal{A}\hbar^2$ is often called the "hyperfine structure of the ground state".

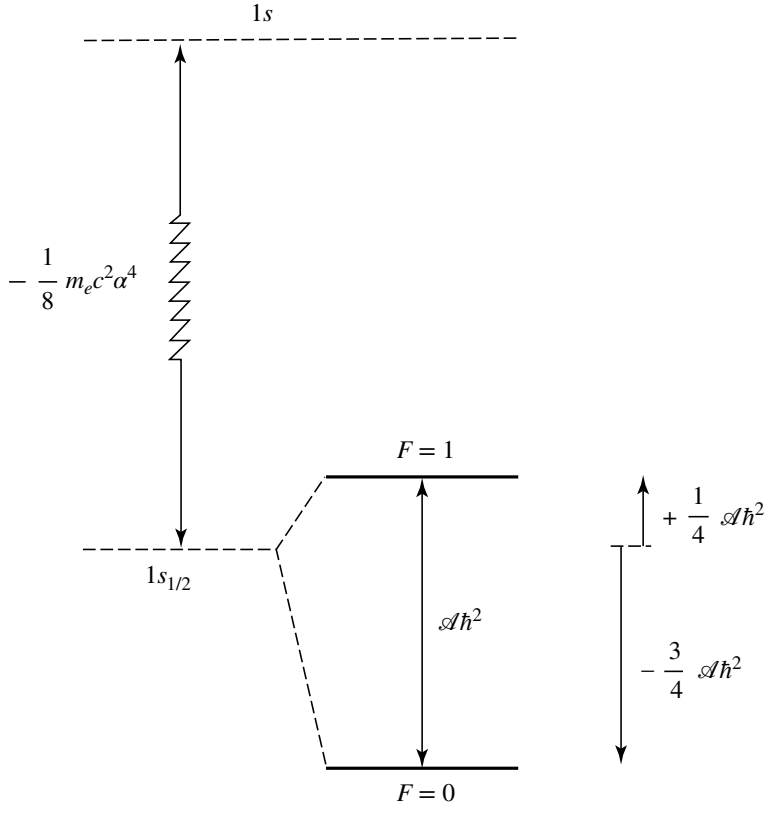


Figure 3: The hyperfine structure of the $n = 1$ level of the hydrogen atom. Under the effect of W_{hf}, the $n = 1$ level undergoes a global shift equal to $-m_e c^2 \alpha^4/8$; J can take on a single value, $J = 1/2$. When the hyperfine coupling W_{hf} is taken into account, the $1s_{1/2}$ level splits into two hyperfine levels, $F = 1$ and $F = 0$. The hyperfine transition $F = 1 \leftrightarrow F = 0$ (the 21 cm line studied in radioastronomy) has a frequency which is known experimentally to twelve significant figures (thanks to the hydrogen maser).

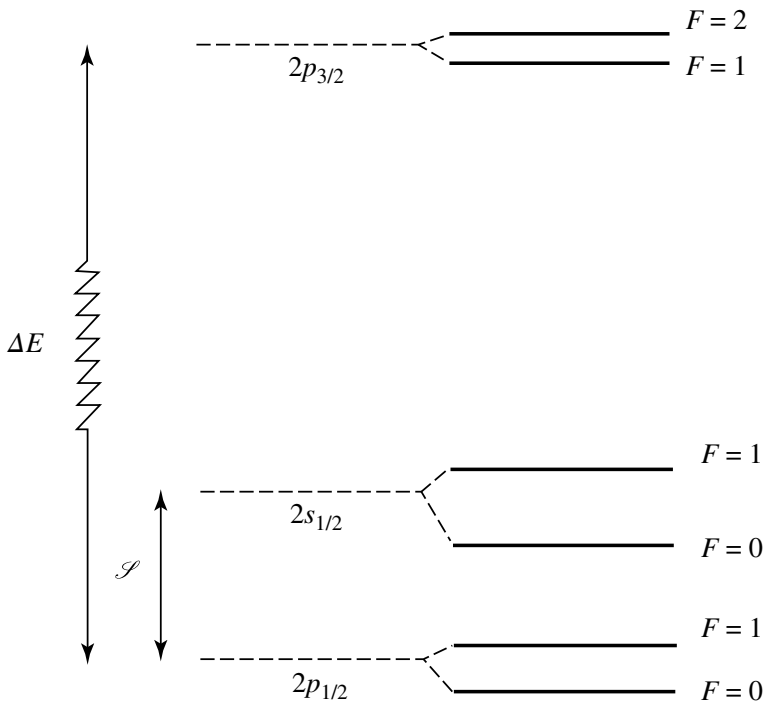


Figure 4: The hyperfine structure of the $n = 2$ level of the hydrogen atom. The separation $\mathscr{S}$ between the two levels $2s_{1/2}$ and $2p_{1/2}$ is the Lamb shift, which is about ten times smaller than the fine structure splitting ΔE separating the two levels $2p_{1/2}$ and $2p_{3/2}$ ($\mathscr{S} \simeq 1\,057{,}8$ MHz: $\Delta E \simeq 10\,969{,}1$ MHz). When the hyperfine coupling W_{hf} is taken into account, each level splits into two hyperfine sublevels (the corresponding value of the quantum number F is indicated on the right-hand side of the figure). The hyperfine splittings are equal to 23.7 MHz for the $2p_{3/2}$ level, 177.56 MHz for the $2s_{1/2}$ level and 59.19 MHz for the $2p_{1/2}$ level (for the sake of clarity, the figure is not drawn to scale).

Comment:

It could be found, similarly, that W_{hf} splits each of the fine structure levels $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$ into a series of hyperfine levels, corresponding to all the values of F separated by one unit and included between $J + I$ and $|J - I|$. For the $2s_{1/2}$ and $2p_{1/2}$ levels, we have $J = 1/2$. Therefore, F takes on the two values $F = 1$ and $F = 0$. For the $2p_{3/2}$ level, $J = 3/2$, and, consequently, we have $F = 2$ and $F = 1$ (*cf.* Fig. 4).

D-3-b. Importance of the hyperfine structure of the $1s$ level

The hyperfine structure of the ground state of the hydrogen atom is currently the physical quantity which is known experimentally to the highest number of significant

figures. Expressed in Hz, it is equal to[7]:

$$\frac{\mathcal{A}\hbar}{2\pi} = 1\,420\,405\,751,767 \pm 0{,}001 \text{ Hz} \tag{D-17}$$

Such a high degree of experimental accuracy was made possible by the development of the "hydrogen maser" in 1963. The principle of such a device is, very schematically, the following: hydrogen atoms, previously sorted (by a magnetic selection of the Stern-Gerlach type) so as to choose those in the upper hyperfine level $F = 1$, are stored in a glass cell (the arrangement is similar to the one shown in Figure 6 of Complement $\mathrm{F_{IV}}$). This constitutes an amplifying medium for the hyperfine frequency $[E(F = 1) - E(F = 0)]/h$. If the cell is placed in a cavity tuned to the hyperfine frequency, and if the losses of the cavity are small enough for the gain to be greater than the losses, the system becomes unstable and can oscillate: we obtain an "atomic oscillator" (a maser). The frequency of the oscillator is very stable and of great spectral purity. Its measurement gives directly the value of the hyperfine splitting, expressed in Hz.

Note, finally, that hydrogen atoms in interstellar space are detected in radioastronomy by the radiation they emit spontaneously when they fall from the $F = 1$ hyperfine level to the $F = 0$ hyperfine level of the ground state (this transition corresponds to a wave length of 21 cm). Most of the information we possess about interstellar hydrogen clouds is supplied by the study of this 21 cm line.

E. The Zeeman effect of the $1s$ ground state hyperfine structure

E-1. Statement of the problem

E-1-a. The Zeeman Hamiltonian W_Z

We now assume the atom to be placed in a static uniform magnetic field $\mathbf{B}_0$ parallel to Oz. This field interacts with the various magnetic moments present in the atom: the orbital and spin magnetic moments of the electron, $\mathbf{M}_L = \dfrac{q}{2m_e}\mathbf{L}$ and $\mathbf{M}_S = \dfrac{q}{m_e}\mathbf{S}$, and the magnetic moment of the nucleus, $\mathbf{M}_I = -\dfrac{qg_p}{2M_p}\mathbf{I}$ [*cf.* expression (B-18)].

The Zeeman Hamiltonian W_Z which describes the interaction energy of the atom with the field $\mathbf{B}_0$ can then be written:

$$\begin{aligned} W_Z &= -\mathbf{B}_0 \cdot (\mathbf{M}_L + \mathbf{M}_S + \mathbf{M}_I) \\ &= \omega_0 (L_z + 2S_z) + \omega_n I_z \end{aligned} \tag{E-1}$$

where ω_0 (the Larmor angular frequency in the field $\mathbf{B}_0$) and ω_n are defined by:

$$\begin{cases} \omega_0 = -\dfrac{q}{2m_e} B_0 & \text{(E-2)} \\ \omega_n = \dfrac{q}{2M_p} g_p B_0 & \text{(E-3)} \end{cases}$$

Since $M_p \gg m_e$, we clearly have:

$$|\omega_0| \gg |\omega_n| \tag{E-4}$$

[7] The calculations presented in this chapter are obviously completely incapable of predicting all these significant figures. Moreover, even the most advanced theories cannot, at the present time, explain more than the first five or six figures of (D-17).

Comment:

Rigorously, W_Z contains another term, which is quadratic in B_0 (the diamagnetic term). This term does not act on the electronic and nuclear spin variables and merely shifts the $1s$ level as a whole, without modifying its Zeeman diagram, which we shall study later. Moreover, it is much smaller than (E-1). Recall that a detailed study of the effect of the diamagnetic term is presented in Complement D_{VII}.

E-1-b. The perturbation "seen" by the $1s$ level

In this section, we propose to study the effect of W_Z on the $1s$ ground state of the hydrogen atom (the case of the $n = 2$ level is slightly more complicated since, in a zero magnetic field, this level possesses both a fine and a hyperfine structure, while the $n = 1$ level has only a hyperfine structure; the principle of the calculation is nevertheless the same). Even with the strongest magnetic fields that can be produced in the laboratory, W_Z is much smaller than the distance between the $1s$ level and the other levels; consequently, its effect can be treated by perturbation theory.

The effect of a magnetic field on an atomic energy level is called the "Zeeman effect". When B_0 is plotted on the x-axis and the energies of the various sublevels it creates are plotted on the y-axis, a *Zeeman diagram* is obtained.

If B_0 is sufficiently strong, the Zeeman Hamiltonian W_Z can be of the same order of magnitude as the hyperfine Hamiltonian[8] W_{hf}, or even larger. On the other hand, if B_0 is very weak, $W_Z \ll W_{hf}$. Therefore, in general it is not possible to establish the relative importance of W_Z and W_{hf}. To obtain the energies of the various sublevels, $(W_Z + W_{hf})$ must be diagonalized inside the $n = 1$ level.

We showed in § D-2 that the restriction of W_{hf} to the $n = 1$ level could be put in the form $\mathcal{A}\mathbf{I} \cdot \mathbf{S}$. Using expression (E-1) for W_Z, we see that we must also calculate matrix elements of the form:

$$\langle n = 1; l = 0; m_L = 0; m'_S; m'_I | \, \omega_0(L_z + 2S_z) + \omega_n I_z \, | n = 1; l = 0; m_L = 0; m_S; m_I \rangle \tag{E-5}$$

The contribution of $\omega_0 L_z$ is zero, since l and m_L are zero. Since $2\omega_0 S_z + \omega_n I_z$ acts only on the spin variables, we can, for this term, separate the orbital part of the matrix element:

$$\langle n = 1; l = 0; m_L = 0 \mid n = 1; l = 0; m_L = 0 \rangle = 1 \tag{E-6}$$

from the spin part.

In conclusion, therefore, we must, ignoring the quantum numbers n, l, m_L, diagonalize the operator:

$$\mathcal{A}\mathbf{I} \cdot \mathbf{S} + 2\omega_0 S_z + \omega_n I_z \tag{E-7}$$

which acts only on the spin degrees of freedom. To do so, we can use either the $\{|m_S, m_I\rangle\}$ basis or the $\{|F, m_F\rangle\}$ basis.

According to (E-4), the last term of (E-7) is much smaller than the second one. To simplify the discussion, we shall neglect the term $w_n l_z$ from now on (it would be possible,

[8] Recall that W_f shifts the $1s$ level as a whole: it therefore also shifts the Zeeman diagram as a whole.

however, to take it into account[9]). The perturbation "seen" by the $1s$ level can therefore be written, finally:

$$\mathcal{A}\,\mathbf{I}\cdot\mathbf{S} + 2\omega_0 S_z \tag{E-8}$$

E-1-c. Different domains of field strength

By varying B_0, we can continuously modify the magnitude of the Zeeman term $2\omega_0 S_z$. We shall consider three different field strengths, determined by the respective orders of magnitude of the hyperfine term and the Zeeman term:

(i) $\hbar\omega_0 \ll \mathcal{A}\hbar^2$: weak fields

(ii) $\hbar\omega_0 \gg \mathcal{A}\hbar^2$: strong fields

(iii) $\hbar\omega_0 \simeq \mathcal{A}\hbar^2$: intermediate fields

We shall later see that it is possible to diagonalize operator (E-8) exactly. However, in order to give a particularly simple example of perturbation theory, we shall use a slightly different method in cases (i) and (ii). In case (i), we shall treat $2\omega_0 S_z$ like a perturbation with respect to $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$. On the other hand, in case (ii), we shall treat $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$ like a perturbation with respect to $2\omega_0 S_z$. The exact diagonalization of the set of two operators, indispensable in case (iii), will allow us to check the preceding results.

E-2. The weak-field Zeeman effect

The eigenstates and eigenvalues of $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$ have already been determined (§ D-2). We therefore obtain two different levels: the three-fold degenerate level,

$$\{|F=1;\, m_F=-1,0,+1\rangle\},$$

of energy $\mathcal{A}\hbar^2/4$, and the non-degenerate level, $\{|F=0; m_F=0\rangle\}$, of energy $-3\mathcal{A}\hbar^2/4$. Since we are treating $2\omega_0 S_z$ like a perturbation with respect to $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$, we must now separately diagonalize the two matrices representing $2\omega_0 S_z$ in the two levels, $F=1$ and $F=0$, corresponding to two distinct eigenvalues of $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$.

E-2-a. Matrix representation of S_z in the $\{|F, m_F\rangle\}$ basis

Since we shall need it later, we shall begin by writing the matrix which represents S_z in the $\{|F, m_F\rangle\}$ basis (for the problem which concerns us here, it would suffice to write the two submatrices corresponding to the $F=1$ and $F=0$ subspaces).

By using formulas (B-22) and (B-23) of Chapter X, we easily obtain:

$$\begin{cases} S_z\,|F=1;\, m_F=1\rangle &= \dfrac{\hbar}{2}\,|F=1;\, m_F=1\rangle \\ S_z\,|F=1;\, m_F=0\rangle &= \dfrac{\hbar}{2}\,|F=0;\, m_F=0\rangle \\ S_z\,|F=1;\, m_F=-1\rangle &= -\dfrac{\hbar}{2}\,|F=1;\, m_F=-1\rangle \\ S_z\,|F=0;\, m_F=0\rangle &= \dfrac{\hbar}{2}\,|F=1;\, m_F=0\rangle \end{cases} \tag{E-9}$$

[9] This is what we do in Complement $\mathrm{C_{XII}}$, in which we study the hydrogen-like systems (muonium, positronium) for which it is not possible to neglect the magnetic moment of one of the two particles.

which gives the following expression for the matrix representing S_z in the $\{|F, m_F\rangle\}$ basis (the basis vectors are arranged in the order $|1,1\rangle$, $|1,0\rangle$, $|1,-1\rangle$, $|0,0\rangle$):

$$(S_z) = \frac{\hbar}{2}\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ \hline 0 & 1 & 0 & 0 \end{array}\right) \tag{E-10}$$

Comment:

It is instructive to compare the preceding matrix with the one which represents F_z in the same basis:

$$(F_z) = \hbar\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right) \tag{E-11}$$

We see, first of all, that the two matrices are not proportional: the (F_z) matrix is diagonal, while the (S_z) one is not.

However, if we confine ourselves to the restrictions of the two matrices in the $F = 1$ subspace [limited by the darker line in expressions (E-10) and (E-11)], we see that they are proportional. Denoting by P_1 the projector onto the $F = 1$ subspace (*cf.* Complement B_{II}), we have:

$$P_1 S_z P_1 = \frac{1}{2} P_1 F_z P_1 \tag{E-12}$$

It would be simple to show that the same relation exists between S_x and F_x on the one hand, and S_y and F_y, on the other.

We have thus found a special case of the Wigner-Eckart theorem (Complement D_X), according to which, in a given eigensubspace of the total angular momentum, all the matrices which represent vector operators are proportional. It is clear from this example that this proportionality exists only for the restrictions of operators to a given eigensubspace of the total angular momentum, and not for the operators themselves.

Moreover, the proportionality coefficient 1/2 which appears in (E-12) can be obtained immediately from the projection theorem. According to formula (30) of Complement E_X, this coefficient is equal to:

$$\frac{\langle \mathbf{S}\cdot\mathbf{F}\rangle_{F=1}}{\langle \mathbf{F}^2\rangle_{F=1}} = \frac{F(F+1)+s(s+1)-I(I+1)}{2F(F+1)} \tag{E-13}$$

Since $s = I = 1/2$, (E-13) is indeed equal to 1/2.

E-2-b. Weak-field eigenstates and eigenvalues

According to the results of § a, the matrix which represents $2\omega_0 S_z$ in the $F = 1$ level can be written:

$$\left(\begin{array}{ccc} \hbar\omega_0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar\omega_0 \end{array}\right) \tag{E-14}$$

In the $F = 0$ level, this matrix reduces to a number, equal to 0.

Since these two matrices are diagonal, we can immediately find the weak-field eigenstates (to zeroth order in ω_0) and the eigenvalues (to first order in ω_0):

$$\begin{array}{ll} \textit{Eigenstates} & \textit{Eigenvalues} \\ |F=1;\, m_F=1\rangle & \Leftrightarrow \dfrac{\mathcal{A}\hbar^2}{4} + \hbar\omega_0 \\ |F=1;\, m_F=0\rangle & \Leftrightarrow \dfrac{\mathcal{A}\hbar^2}{4} + 0 \\ |F=1;\, m_F=-1\rangle & \Leftrightarrow \dfrac{\mathcal{A}\hbar^2}{4} - \hbar\omega_0 \\ |F=0;\, m_F=0\rangle & \Leftrightarrow -3\dfrac{\mathcal{A}\hbar^2}{4} + 0 \end{array} \tag{E-15}$$

In Figure 5, we have plotted $\hbar\omega_0$ on the x-axis and the energies of the four Zeeman sublevels on the y-axis (Zeeman diagram). In a zero field, we have the two hyperfine levels, $F = 1$ and $F = 0$. When the field B_0 is turned on, the $|F = 0,\, m_F = 0\rangle$ sublevel, which is not degenerate, starts horizontally; as for the $F = 1$ level, its three-fold degeneracy is completely removed: three equidistant sublevels are obtained, varying linearly with $\hbar\omega_0$ with slopes of $+1$, 0, -1 respectively.

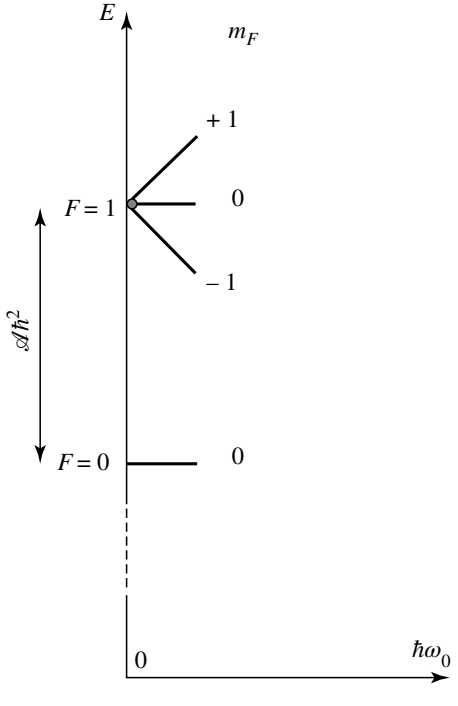


Figure 5: The weak-field Zeeman diagram of the $1s$ ground state of the hydrogen atom. The hyperfine $F = 1$ level splits into three equidistant levels, each of which corresponds to a well-defined value of the quantum number m_F. The $F = 0$ level does not undergo any shift to first order in ω_0.

The preceding treatment is valid as long as the difference $\hbar\omega_0$ between two adjacent Zeeman sublevels of the $F = 1$ level remains much smaller than the zero-field difference between the $F = 1$ and $F = 0$ levels (the hyperfine structure).

Comment:

The Wigner-Eckart theorem, mentioned above, makes it possible to show that, in a given level F of the total angular momentum, the Zeeman Hamiltonian $\omega_0(L_z + 2S_z)$ is represented by a matrix proportional to F_z. Thus, we can write, denoting the projector onto the F level by P_F:

$$P_F\left[\omega_0(L_z + 2S_z)\right]P_F = g_F\omega_0 P_F F_z P_F \tag{E-16}$$

g_F is called the *Landé factor* of the F state. In the case which concerns us here, $g_{F=1} = 1$.

E-2-c. The Bohr frequencies involved in the evolution of $\langle F\rangle$ and $\langle S\rangle$; comparison with the vector model of the atom

In this section, we shall determine the different Bohr frequencies which appear in the evolution of $\langle\mathbf{F}\rangle$ and $\langle\mathbf{S}\rangle$, and show that certain aspects of the results obtained recall those found by using the vector model of the atom (*cf.* Complement $\mathrm{F_X}$).

First of all, we shall briefly review the predictions of the vector model of the atom (in which the various angular momenta are treated like classical vectors) as far as the hyperfine coupling between $\mathbf{I}$ and $\mathbf{S}$ is concerned. In a zero field, $\mathbf{F} = \mathbf{I} + \mathbf{S}$ is a constant of the motion. $\mathbf{I}$ and $\mathbf{S}$ precess about their resultant $\mathbf{F}$ with an angular velocity proportional to the coupling constant $\mathcal{A}$ between $\mathbf{I}$ and $\mathbf{S}$. If the system is, in addition, placed in a weak static field $\mathbf{B}_0$ parallel to Oz, onto the rapid precessional motion of $\mathbf{I}$ and $\mathbf{S}$ about $\mathbf{F}$ is superposed a slow precessional motion of $\mathbf{F}$ about Oz (Larmor precession; Fig. 6).

F_z is therefore a constant of the motion, while S_z has a static part (the projection onto Oz of the component of $\mathbf{S}$ parallel to $\mathbf{F}$), and a part which is modulated by the hyperfine precession frequency (the projection onto Oz of the component of $\mathbf{S}$ perpendicular to $\mathbf{F}$, which precesses about $\mathbf{F}$).

Let us compare these semi-classical results with those of the quantum theory presented earlier in this section. To do so, we must consider the time evolution of the average values $\langle F_z\rangle$ et $\langle S_z\rangle$. According to the discussion of § D-2-d of Chapter III, the average value $\langle G\rangle(t)$ of a physical quantity G contains a series of components which oscillate at the various Bohr frequencies $(E - E')/h$ of the system. Also, a given Bohr frequency appears in $\langle G\rangle(t)$ only if the matrix element of G between the states corresponding to the two energies is different from zero. In the problem which concerns us here, the eigenstates of the weak-field Hamiltonian are the $|F, m_F\rangle$ states. Now consider the two matrices (E-10) and (E-11) which represent S_z and F_z in this basis. Since F_z has only diagonal matrix elements, no Bohr frequency different from zero can appear in $\langle F_z\rangle(t)$: $\langle F_z\rangle$ is therefore static. On the other hand, S_z has, not only diagonal matrix elements (with which is associated a static component of $\langle S_z\rangle$), but also a non-diagonal element between the $|F = 1; m_F = 0\rangle$ and $|F = 0; m_F = 0\rangle$ states, whose energy difference is $\mathcal{A}\hbar^2$, according to Table (E-15) (or Figure 5). It follows that $\langle S_z\rangle$ has, in addition to a static component, a component modulated at the angular frequency $\mathcal{A}\hbar$. This result recalls the one obtained using the vector model of the atom[10].

[10] A parallel could also be established between the evolution of $\langle F_x\rangle$, $\langle S_x\rangle$, $\langle F_y\rangle$, $\langle S_y\rangle$, and that of the projections of the vectors $\mathbf{F}$ and $\mathbf{S}$ of Figure 6 onto Ox and Oy. However, the motion of $\langle\mathbf{F}\rangle$ and $\langle\mathbf{S}\rangle$ does not coincide perfectly with that of the classical angular momenta. In particular, the modulus of $\langle\mathbf{S}\rangle$ is not necessarily constant (in quantum mechanics, $\left\langle\mathbf{S}^2\right\rangle \neq \langle\mathbf{S}\rangle^2$); see discussion of Complement $\mathrm{F_X}$.

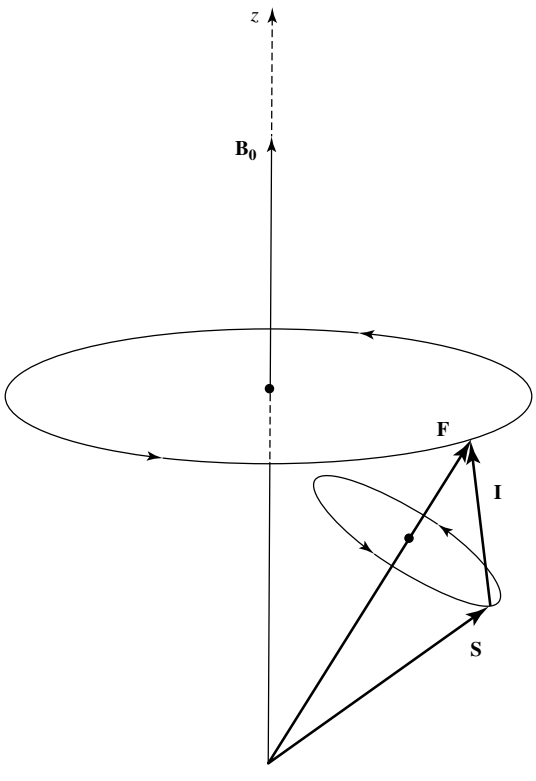


Figure 6: The motion of **S**, **I** *and* **F** *in the vector model of the atom.* **S**, **I** *precess rapidly about* **F** *under the effect of the hyperfine coupling. In a weak field,* **F** *slowly precesses about* $\mathbf{B}_0$ *(Larmor precession).*

Comment:

A relation can be established between perturbation theory and the vector model of the atom. The influence of a weak field B_0 on the $F = 1$ and $F = 0$ levels can be obtained by retaining in the Zeeman Hamiltonian $2\omega_0 S_z$ only the matrix elements in the $F = 1$ and $F = 0$ levels, "forgetting" the matrix element of S_z between $|F = 1; m_F = 0\rangle$ and $|F = 0; m_F = 0\rangle$. Proceeding in this way, we also "forget" the modulated component of $\langle S_z \rangle$, which is proportional to this matrix element. We therefore keep only the component of $\langle \mathbf{S} \rangle$ parallel to $\langle \mathbf{F} \rangle$.

Now, this is precisely what we do in the vector model of the atom when we want to evaluate the interaction energy with the field $\mathbf{B}_0$. In a weak field, **F** does precess about $\mathbf{B}_0$ much more slowly than **S** does about **F**. The interaction of $\mathbf{B}_0$ with the component of **S** perpendicular to **F** therefore has no effect, on the average; only the projection of **S** onto **F** counts. This is how, for example, the Landé factor is calculated.

E-3. The strong-field Zeeman effect

We must now start by diagonalizing the Zeeman term.

E-3-a. Eigenstates and eigenvalues of the Zeeman term

This term is diagonal in the $\{|m_S, m_I\rangle\}$ basis:

$$2\omega_0 S_z \,|m_S, m_I\rangle = 2m_S\,\hbar\omega_0\,|m_S, m_I\rangle \tag{E-17}$$

Since $m_S = \pm 1/2$, the eigenvalues are equal to $\pm\hbar\omega_0$. Each of them is therefore two-fold degenerate, because of the two possible values of m_I. We therefore have[11] :

$$\begin{aligned} 2\omega_0 S_z|+,\pm\rangle &= +\hbar\omega_0|+,\pm\rangle \\ 2\omega_0 S_z|-,\pm\rangle &= -\hbar\omega_0|-,\pm\rangle \end{aligned} \tag{E-18}$$

E-3-b. Effects of the hyperfine term considered as a perturbation

The corrections to first order in $\mathcal{A}$ can be obtained by diagonalizing the restrictions of the operator $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$ to the two subspaces $\{|+,\pm\rangle\}$ and $\{|-,\pm\rangle\}$ corresponding to the two different eigenvalues of $2\omega_0 S_z$.

First of all, notice that, in each of these two subspaces, the two basis vectors $|+,+\rangle$ and $|+,-\rangle$ (or $|-,+\rangle$ and $|-,-\rangle$) are also eigenvectors of F_z, but do not correspond to the same value of $m_F = m_S + m_I$. Since the operator $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S} = \frac{\mathcal{A}}{2}\left(F^2 - I^2 - S^2\right)$ commutes with F_z, it has no matrix elements between the two states $|+,+\rangle$ and $|+,-\rangle$, or $|-,+\rangle$ and $|-,-\rangle$. The two matrices representing $\mathcal{A}\,\mathbf{I}\cdot\mathbf{S}$ in the two subspaces $\{|+,\pm\rangle\}$ and $\{|-,\pm\rangle\}$ are then diagonal, and their eigenvalues are simply the diagonal elements $\langle m_S; m_I \mid \mathcal{A}\,\mathbf{I}\cdot\mathbf{S} \mid m_S; m_I\rangle$ which can also be written, using the relation:

$$\mathbf{I}\cdot\mathbf{S} = I_z S_z + \frac{1}{2}\left(I_+ S_- + I_- S_+\right) \tag{E-19}$$

in the form:

$$\begin{aligned} &\langle m_S, m_I \mid \mathcal{A}\,\mathbf{I}\cdot\mathbf{S} \mid m_S, m_I\rangle \\ &\qquad = \langle m_S, m_I \mid \mathcal{A} I_z S_z \mid m_S, m_I\rangle = \mathcal{A}\hbar^2 m_S m_I \end{aligned} \tag{E-20}$$

Finally, in a strong field, the eigenstates (to zeroth order in $\mathcal{A}$) and the eigenvalues (to first order in $\mathcal{A}$) are:

$$\begin{array}{lll} \textit{Eigenstates} & & \textit{Eigenvalues} \\ |+,+\rangle & \Leftrightarrow & \hbar\omega_0 + \dfrac{\mathcal{A}\hbar^2}{4} \\ |+,-\rangle & \Leftrightarrow & \hbar\omega_0 - \dfrac{\mathcal{A}\hbar^2}{4} \\ |-,+\rangle & \Leftrightarrow & -\hbar\omega_0 - \dfrac{\mathcal{A}\hbar^2}{4} \\ |-,-\rangle & \Leftrightarrow & -\hbar\omega_0 + \dfrac{\mathcal{A}\hbar^2}{4} \end{array} \tag{E-21}$$

[11] To simplify the notation, we shall often write $|\varepsilon_S, \varepsilon_I\rangle$ instead of $|m_S, m_I\rangle$, where ε_S and ε_I are equal to $+$ or $-$, depending on the signs of m_S and m_I.

In Figure 7, the solid-line curves on the right-hand side (for $\hbar\omega_0 \gg \mathcal{A}\hbar^2$) represent the strong-field energy levels: we obtain two parallel straight lines of slope $+1$, separated by an energy $\mathcal{A}\hbar^2/2$, and two parallel straight lines of slope -1, also separated by $\mathcal{A}\hbar^2/2$. The perturbation treatments presented in this section and the preceding one therefore give the strong-field asymptotes and the tangents at the origin of the energy levels.

Comment:

The strong-field splitting $\mathcal{A}\hbar^2/2$ of the two states, $|+,+\rangle$ and $|+,-\rangle$ or $|-,+\rangle$ and $|-,-\rangle$, can be interpreted in the following way. We have seen that only the term

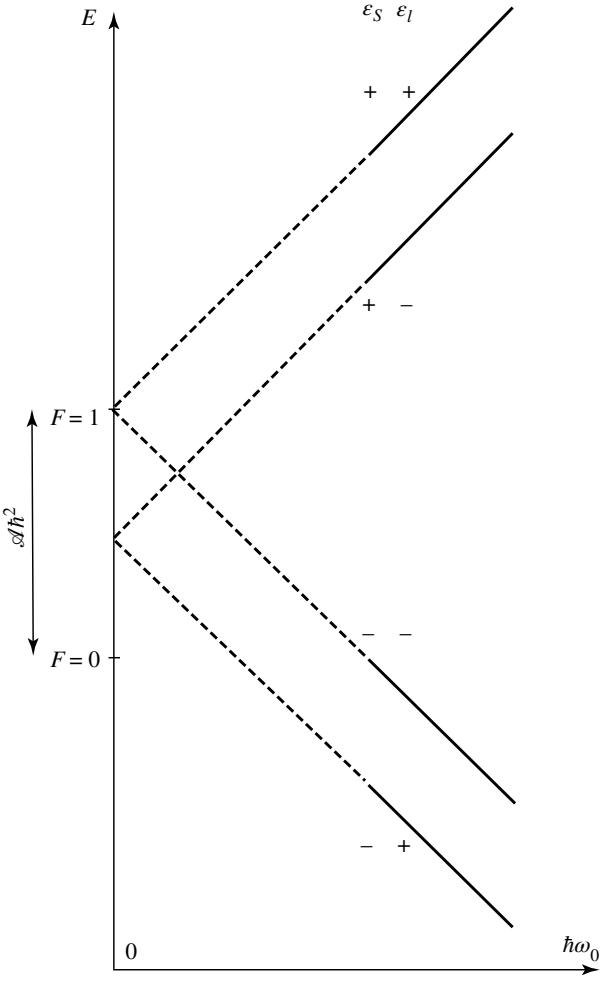


Figure 7: The strong-field Zeeman diagram of the $1s$ *ground state of the hydrogen atom. For each orientation of the electronic spin ($\varepsilon_S = +$ or $\varepsilon_S = -$), we obtain two parallel straight lines separated by an energy $\mathcal{A}\hbar^2/2$, each one corresponding to a different orientation of the proton spin ($\varepsilon_I = +$ or $\varepsilon_I = -$).*

$I_z S_z$ of expression (E-19) for $\mathbf{I} \cdot \mathbf{S}$ is involved in a strong field, when the hyperfine coupling is treated like a perturbation of the Zeeman term. The total strong-field Hamiltonian (E-8) can therefore be written:

$$2\omega_0 S_z + \mathcal{A} I_z S_z = 2\left(\omega_0 + \frac{\mathcal{A}}{2} I_z\right) S_z \tag{E-22}$$

It is as if the electronic spin "saw", in addition to the external field $\mathbf{B}_0$, a smaller "internal field", arising from the hyperfine coupling between $\mathbf{I}$ and $\mathbf{S}$ and having two possible values, depending on whether the nuclear spin points up or down. This field adds to or substracts from $\mathbf{B}_0$ and is responsible for the energy difference between $|+,+\rangle$ and $|+,-\rangle$ or between $|-,+\rangle$ and $|-,-\rangle$.

E-3-c. The Bohr frequencies involved in the evolution of $\langle S_z \rangle$

In a strong field, the Zeeman coupling of $\mathbf{S}$ with $\mathbf{B}_0$ is more important than the hyperfine coupling of $\mathbf{S}$ with $\mathbf{I}$. If we start by neglecting this hyperfine coupling, the vector model of the atom predicts that $\mathbf{S}$ will precess (very rapidly since $|\mathbf{B}_0|$ is large) about the Oz direction of $\mathbf{B}_0$ ($\mathbf{I}$ remains motionless, since we have assumed ω_n to be negligible).

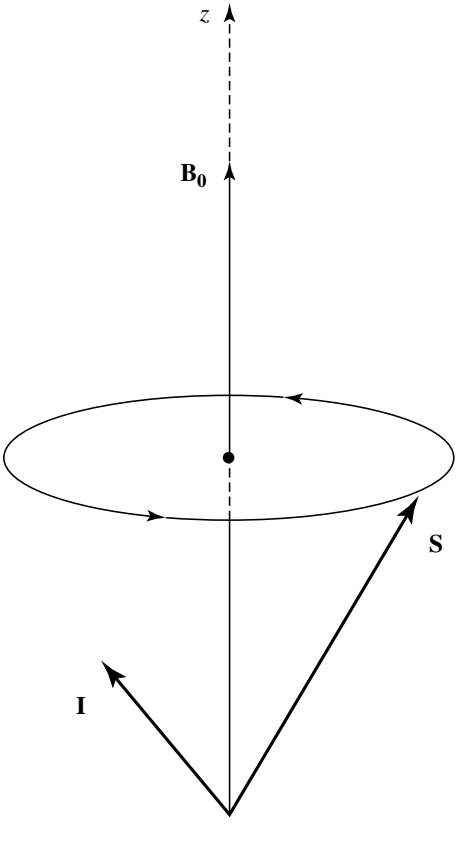


Figure 8: The motion of $\mathbf{S}$ *in the vector model of the atom. In a strong field,* $\mathbf{S}$ *precesses rapidly about* $\mathbf{B}_0$ *(here we are neglecting both the Zeeman coupling between* $\mathbf{I}$ *and* $\mathbf{B}_0$ *and the hyperfine coupling between* $\mathbf{I}$ *and* $\mathbf{S}$*, so that* $\mathbf{I}$ *remains motionless).*

Expression (E-19) for the hyperfine coupling remains valid for classical vectors. Because of the very rapid precession of $\mathbf{S}$, the terms S_+ and S_- oscillate very fast and have, on the average, no effect, so that only the term $I_z S_z$ counts. The effect of the hyperfine coupling is therefore to add a small field parallel to Oz and proportional to I_z (*cf.* comment of the preceding section), which accelerates or slows down the precession of $\mathbf{S}$ about Oz, depending on the sign of I_z. The vector model of the atom thus predicts that S_z will be static in a strong field.

We shall show that quantum theory gives an analogous result for the average value $\langle S_z \rangle$ of the observable S_z. In a strong field, the well-defined energy states are, as we have seen, the states $|m_S, m_I\rangle$. Now, in this basis, the operator S_z has only diagonal matrix elements. No non-zero Bohr frequency can therefore appear in $\langle S_z \rangle$, which, consequently, is a static quantity[12], unlike its weak-field counterpart (*cf.* § E-2-c).

E-4. The intermediate-field Zeeman effect

E-4-a. The matrix which represents the total perturbation in the $\{|F, m_F\rangle\}$ basis

The $\{|F, m_F\rangle\}$ states are eigenstates of the operator $\mathcal{A}\mathbf{I} \cdot \mathbf{S}$. The matrix which represents this operator in the $\{|F, m_F\rangle\}$ basis is therefore diagonal. The diagonal elements corresponding to $F = 1$ are equal to $\mathcal{A}\hbar^2/4$, and those corresponding to $F = 0$, to $-3\mathcal{A}\hbar^2/4$. Furthermore, we have already written, in (E-10), the matrix representation of S_z in the same basis. It is now very simple to write the matrix which represents the total perturbation (E-8). Arranging the basis vectors in the order $|1,1\rangle$, $|1,-1\rangle$, $|1,0\rangle$, $|0,0\rangle$, we thus obtain:

$$\left(\begin{array}{c|c|c|c} \frac{\mathcal{A}\hbar^2}{4} + \hbar\omega_0 & 0 & 0 & 0 \\ \hline 0 & \frac{\mathcal{A}\hbar^2}{4} - \hbar\omega_0 & 0 & 0 \\ \hline 0 & 0 & \frac{\mathcal{A}\hbar^2}{4} & \hbar\omega_0 \\ \hline 0 & 0 & \hbar\omega_0 & -\frac{3\mathcal{A}\hbar^2}{4} \end{array}\right) \tag{E-23}$$

Comment:

S_z and F_z commute; $2\omega_0 S_z$ can therefore have non-zero matrix elements only between two states with the same m_F. Thus, we could have predicted all the zeros of matrix (E-23).

[12] The study of $\langle S_x \rangle$ and $\langle S_y \rangle$ presents no difficulty. We find two Bohr angular frequencies: one, $\omega_0 + \mathcal{A}\hbar/2$, slightly larger than ω_0, and the other one, $\omega_0 - \mathcal{A}\hbar/2$, slightly smaller. They correspond to the two possible orientations of the "internal field", produced by I_z, which adds to the external field B_0. Similarly, we find that $\mathbf{I}$ precesses about the "internal field" produced by S_z.

E-4-b. Energy values in an arbitrary field

Matrix (E-23) can be broken into two 1×1 matrices and one 2×2 matrix. The two 1×1 matrices immediately yield two eigenvalues:

$$\begin{cases} E_1 = \dfrac{\mathcal{A}\hbar^2}{4} + \hbar\omega_0 \\ E_2 = \dfrac{\mathcal{A}\hbar^2}{4} - \hbar\omega_0 \end{cases} \tag{E-24}$$

corresponding respectively to the state $|1,1\rangle$ (that is, the state $|+,+\rangle$) and to the state $|1,-1\rangle$ (that is, the state $|-,-\rangle$). In Figure 9, the two straight lines of slopes ,+1 and -1 passing through the point whose ordinate is $+\mathcal{A}\hbar^2/4$ for a zero field (for which the perturbation theory treatment gave only the initial and asymptotic behavior) therefore represent, for any B_0, two of the Zeeman sublevels.

The eigenvalue equation of the remaining 2×2 matrix can be written:

$$\left(\frac{\mathcal{A}\hbar^2}{4} - E\right)\left(-\frac{3\mathcal{A}\hbar^2}{4} - E\right) - \hbar^2\omega_0^2 = 0 \tag{E-25}$$

The two roots of this equation can easily be found:

$$E_3 = -\frac{\mathcal{A}\hbar^2}{4} + \sqrt{\left(\frac{\mathcal{A}\hbar^2}{2}\right)^2 + \hbar^2\omega_0^2} \tag{E-26}$$

$$E_4 = -\frac{\mathcal{A}\hbar^2}{4} - \sqrt{\left(\frac{\mathcal{A}\hbar^2}{2}\right)^2 + \hbar^2\omega_0^2} \tag{E-27}$$

When $\hbar\omega_0$ varies, the two points of abscissas $\hbar\omega_0$ and ordinates E_3 and E_4 follow the two branches of a hyperbola (Fig. 9). The asymptotes of this hyperbola are the two straight lines whose equation is $E = -(\mathcal{A}\hbar^2/4) \pm \hbar\omega_0$, obtained in § 3 above. The two turning points of the hyperbola have abscissas of $\omega_0 = 0$ and ordinates of $-(\mathcal{A}\hbar^2/4) \pm \mathcal{A}\hbar^2/2$, that is, $\mathcal{A}\hbar^2/4$ and $-3\mathcal{A}\hbar^2/4$. The tangents at both these points are horizontal. This is in agreement with the results of § 2 for the states $|F=1; m_F=0\rangle$ and $|F=0; m_F=0\rangle$.

The preceding results are summarized in Figure 9, which is the Zeeman diagram of the $1s$ ground state.

E-4-c. Partial hyperfine decoupling

In a weak field, the well-defined energy states are the states $|F, m_F\rangle$; in a strong field, the states $|m_S, m_I\rangle$; in an intermediate field, the eigenstates of matrix (E-23), which are intermediate between the states $|F, m_F\rangle$ and the states $|m_S, m_I\rangle$.

One thus moves continuously from a strong coupling between **I** and **S** (coupled bases) to a total decoupling (uncoupled bases) via a partial coupling.

Comment:

An analogous phenomenon exists for the Zeeman fine structure effect. If, for simplicity, we neglect W_{hf}, we know (§ C) that, in a zero field, the eigenstates of the Hamiltonian H are the $|J, m_J\rangle$ states corresponding to a strong coupling

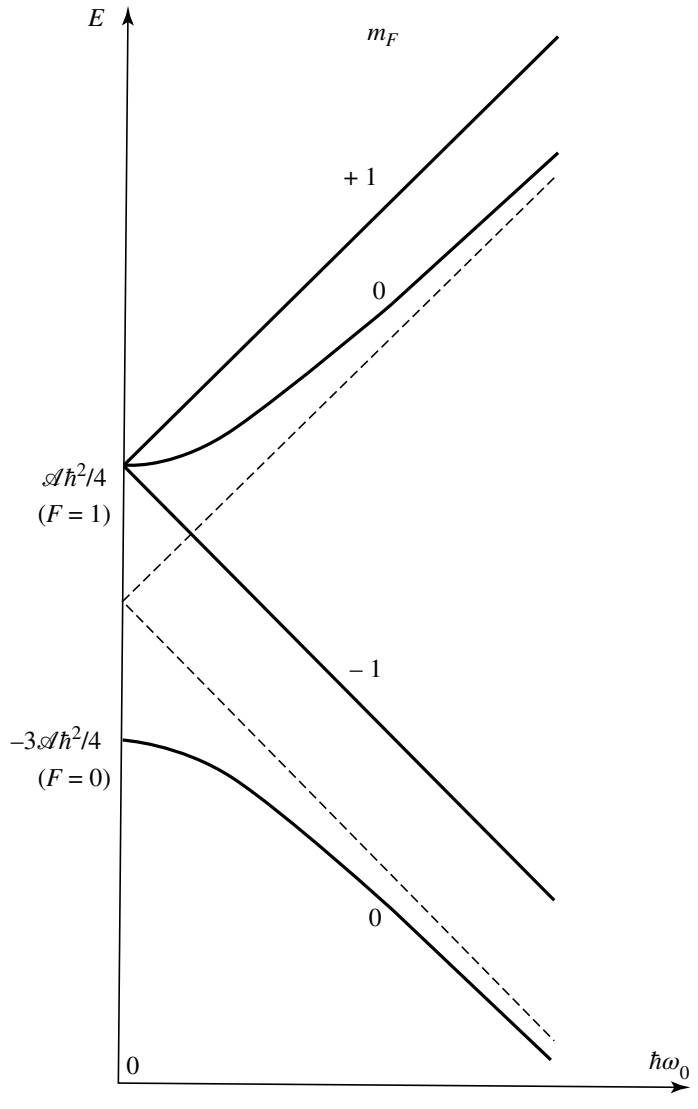


Figure 9: The Zeeman diagram (for an arbitrary field) of the $1s$ *ground state of the hydrogen atom:* m_F *remains a good quantum number for any value of the field. We obtain two straight lines, of opposite slopes, corresponding to the values,* $+1$ *and* -1*, of* m_F*, as well as a hyperbola whose two branches are associated with the two* $m_F = 0$ *levels. Figures 5 and 7 give, respectively, the tangents at the origin and the asymptotes of the levels shown in this diagram.*

between **L** and **S** (the spin-orbit coupling). This property remains valid as long as $W_Z \ll W_f$. If, on the other hand, B_0 is strong enough to make $W_Z \gg W_f$, we find that the eigenstates of H are the $|m_L, m_S\rangle$ states corresponding to a total decoupling of **L** and **S**. The intermediate zone ($W_Z \simeq W_f$) corresponds to a partial

coupling of **L** and **S**. See, for example, Complement D_{XII}, in which we study the Zeeman effect of the $2p$ level (without taking W_{hf} into account).

References and suggestions for further reading:

The hydrogen atom spectrum: Series (11.7), Bethe and Salpeter (11.10).

The Dirac equation: the subsection "Relativistic quantum mechanics" of section 2 of the bibliography and Messiah (1.17), Chap. XX, especially §§ V and IV-27.

The fine structure of the $n = 2$ level and the Lamb shift: Lamb and Retherford (3.11); Frisch (3.13); Series (11.7), Chaps. VI, VII and VIII.

The hyperfine structure of the ground state: Crampton et al. (3.12).

The Zeeman effect and the vector model of the atom: Cagnac and Pebay-Peyroula (11.2), Chap. XVII, §§ 3E and 4C; Born (11.4), Chap. 6, § 2.

Interstellar hydrogen: Roberts (11.17); Encrenaz (12.11), Chap. IV.

A_{XII} : THE MAGNETIC HYPERFINE HAMILTONIAN	Derivation of the expression for the hyperfine Hamiltonian used in Chapter XII. Gives the physical interpretation of the various terms appearing in the Hamiltonian – in particular the contact term. Rather difficult.
B_{XII} : CALCULATION OF THE AVERAGE VALUES OF THE FINE-STRUCTURE HAMILTONIAN IN THE $1s$, $2s$ AND $2p$ STATES	The detailed calculations of certain radial integrals appearing in the expression obtained in Chapter XII for the energy shifts. Not conceptually difficult.
C_{XII} : THE HYPERFINE STRUCTURE AND THE ZEEMAN EFFECT FOR MUONIUM AND POSITRONIUM	Extension of the study of §§ D and E of Chapter XII to two important hydrogen-like systems, muonium and positronium, already presented in Complement A_{VII}. Brief description of experimental methods for studying these two systems. Simple if the calculations of §§ D and E of Chapter XII have been well understood.
D_{XII} : THE INFLUENCE OF THE ELECTRON SPIN ON THE ZEEMAN EFFECT OF THE HYDROGEN RESONANCE LINE	Study of the effect of the electronic spin on the frequencies and polarizations of the Zeeman components of the resonance line of hydrogen. Improves the results obtained in Complement D_{VII}, in which the electron spin was ignored (uses certain results of that complement). Moderately difficult.
E_{XII} : THE STARK EFFECT OF THE HYDROGEN ATOM	Study of the effect of a static electric field on the ground state ($n = 1$) and of the first excited state ($n = 2$) of the hydrogen atom (Stark effect). Shows the importance for the Stark effect of the existence of a degeneracy between two states of different parities. Rather simple.

Complement A_{XII}

The magnetic hyperfine Hamiltonian

1 Interaction of the electron with the scalar and vector potentials created by the proton 1267

2 The detailed form of the hyperfine Hamiltonian 1268

2-a Coupling of the magnetic moment of the proton with the orbital angular momentum of the electron 1268

2-b Coupling with the electron spin 1270

3 Conclusion: the hyperfine-structure Hamiltonian 1274

The aim of this complement is to justify the expression for the hyperfine Hamiltonian given in Chapter XII [relation (B-20)]. As in that chapter, we shall confine our reasoning to the hydrogen atom, which is composed of a single electron and a proton, although most of the ideas remain valid for any atom. We have already said that the origin of the hyperfine Hamiltonian is the coupling between the electron and the electromagnetic field created by the proton. We shall therefore call $\mathbf{A}_I(\mathbf{r})$ and $U_I(\mathbf{r})$ the vector and scalar potentials associated with this electromagnetic field. We shall begin by considering the Hamiltonian of an electron subjected to these potentials.

1. Interaction of the electron with the scalar and vector potentials created by the proton

Let $\mathbf{R}$ and $\mathbf{P}$ be the position and momentum of the electron, $\mathbf{S}$, its spin; m_e, its mass; and q, its charge; $\mu_B = q\hbar/2m_e$ is the Bohr magneton.

The Hamiltonian H of the electron in the field of the proton can be written:

$$H = \frac{1}{2m_e}\left[\mathbf{P} - q\mathbf{A}_I(\mathbf{R})\right]^2 + qU_I(\mathbf{R}) - 2\mu_B\left(\frac{\mathbf{S}}{\hbar}\right)\cdot\Big(\boldsymbol{\nabla}\times\mathbf{A}_I(\mathbf{R})\Big) \tag{1}$$

This operator is obtained from expression (B-46) of Chapter III (the Hamiltonian of a spinless particle) by adding to it the coupling energy between the magnetic moment $2\mu_B\mathbf{S}/\hbar$ associated with the spin and the magnetic field $\boldsymbol{\nabla}\times\mathbf{A}_I(\mathbf{r})$.

We shall begin by studying the terms which, in (1), arise from the scalar potential $U_I(\mathbf{r})$. According to Complement E_X, this potential results from the superposition of several contributions, each of them associated with one of the electric multipole moments of the nucleus. For an arbitrary nucleus, we must consider:

(i) The total charge $-Zq$ of the nucleus (the moment of order $k = 0$), which yields a potential energy:

$$V_0(\mathbf{r}) = qU_0(\mathbf{r}) = -\frac{Zq^2}{4\pi\varepsilon_0 r} \tag{2}$$

(with, for the proton, $Z = 1$). Now, the Hamiltonian which we chose in Chapter VII for the study of the hydrogen atom is precisely:

$$H_0 = \frac{\mathbf{P}^2}{2m_e} + V_0(\mathbf{R}) \tag{3}$$

$V_0(\mathbf{R})$ has therefore already been taken into account in the Hamiltonian H_0.

(*ii*) The electric quadrupole moment ($k = 2$) of the nucleus. The corresponding potential adds to the potential V_0 and yields a term of the hyperfine Hamiltonian, called the electric quadrupole term. The results of Complement E_X enable us to write this term without difficulty. In the case of the hydrogen atom, it is zero, since the proton, which is a spin 1/2 particle, has no electric quadrupole moment (*cf.* § 2-c-α of Complement E_X).

(*iii*) The electric multipole moments of order $k = 4$, 6, etc... which are theoretically involved as long as $k \leq 2I$; for the proton, they are all zero.

Thus, for the hydrogen atom, potential (2) is really the potential seen by the electron[1]. There is no need to add any corrections to it (by hydrogen atom, we mean the electron-proton system, excluding isotopes such as deuterium: since the deuterium nucleus has a spin $I = 1$, we would have to take into account an electric quadrupole hyperfine Hamiltonian – see comment (*i*) at the end of this complement).

Now let us consider the terms arising from the vector potential $\mathbf{A}_I(\mathbf{r})$ in (1). We denote by $\mathbf{M}_I$ the magnetic dipole moment of the proton (which, for the same reason as above, cannot have magnetic multipole moments of order $k > 1$). We have:

$$\mathbf{A}_I(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{\mathbf{M}_I \times \mathbf{r}}{r^3} \tag{4}$$

The hyperfine Hamiltonian W_{hf} can now be obtained by retaining in (1) the terms which are linear in $\mathbf{A}_I$:

$$W_{hf} = -\frac{q}{2m_e} \left[\mathbf{P} \cdot \mathbf{A}_I(\mathbf{R}) + \mathbf{A}_I(\mathbf{R}) \cdot \mathbf{P}\right] - 2\mu_B \left(\frac{\mathbf{S}}{\hbar}\right) \cdot \boldsymbol{\nabla} \times \mathbf{A}_I(\mathbf{R}) \tag{5}$$

and by replacing $\mathbf{A}_I$ by expression (4) (since W_{hf} already makes a very small correction to the energy levels of H_0, it is perfectly legitimate to ignore the second-order term, in $\mathbf{A}_I^2$). This is what we shall do in the following section.

2. The detailed form of the hyperfine Hamiltonian

2-a. Coupling of the magnetic moment of the proton with the orbital angular momentum of the electron

First of all, we shall calculate the first term of (5). Using (4), we have:

$$\mathbf{P} \cdot \mathbf{A}_I(\mathbf{R}) + \mathbf{A}_I(\mathbf{R}) \cdot \mathbf{P} = \frac{\mu_0}{4\pi} \left\{ \mathbf{P} \cdot (\mathbf{M}_I \times \mathbf{R}) \frac{1}{R^3} + \frac{1}{R^3} (\mathbf{M}_I \times \mathbf{R}) \cdot \mathbf{P} \right\} \tag{6}$$

[1] We are concerned here only with the potential outside the nucleus, where the multipole moment expansion is possible. Inside the nucleus, we know that the potential does not have form (2). This causes a shift in the atomic levels called the "volume effect". This effect was studied in Complement D_{XI}, and we shall not take it into account here.

We can apply the rules for a mixed vector product to vector operators as long as we do not change the order of two non-commuting operators. The components of $\mathbf{M}_I$ commute with $\mathbf{R}$ and $\mathbf{P}$, so we have:

$$(\mathbf{M}_I \times \mathbf{R}) \cdot \mathbf{P} = (\mathbf{R} \times \mathbf{P}) \cdot \mathbf{M}_I = \mathbf{L} \cdot \mathbf{M}_I \tag{7}$$

where:

$$\mathbf{L} = \mathbf{R} \times \mathbf{P} \tag{8}$$

is the orbital angular momentum of the electron. It can easily be shown that:

$$\left[\mathbf{L}, \frac{1}{R^3}\right] = \mathbf{0} \tag{9}$$

(any function of $|\mathbf{R}|$ is a scalar operator), so that:

$$\frac{1}{R^3} (\mathbf{M}_I \times \mathbf{R}) \cdot \mathbf{P} = \frac{\mathbf{L} \cdot \mathbf{M}_I}{R^3} \tag{10}$$

Similarly:

$$\mathbf{P} \cdot (\mathbf{M}_I \times \mathbf{R}) \frac{1}{R^3} = -\mathbf{M}_I \cdot (\mathbf{P} \times \mathbf{R}) \frac{1}{R^3} = \frac{\mathbf{M}_I \cdot \mathbf{L}}{R^3} \tag{11}$$

since:

$$-\mathbf{P} \times \mathbf{R} = \mathbf{L} \tag{12}$$

Thus, the first term of (5) makes a contribution W_{hf}^L to W_{hf} which is equal to:

$$W_{hf}^L = -\frac{\mu_0}{4\pi} \frac{q}{2m_e} 2 \frac{\mathbf{M}_I \cdot \mathbf{L}}{R^3} = -\frac{\mu_0}{4\pi} 2\mu_B \frac{\mathbf{M}_I \cdot (\mathbf{L}/\hbar)}{R^3} \tag{13}$$

This term corresponds to the coupling between the nuclear magnetic moment $\mathbf{M}_I$ and the magnetic field:

$$\mathbf{B}_L = \frac{\mu_0}{4\pi} \frac{q\mathbf{L}}{m_e r^3}$$

created by the current loop associated with the rotation of the electron (*cf.* Fig. 1).

Comment:

The presence of the $1/R^3$ term in (13) might lead us to believe that there is a singularity at the origin, and that certain matrix elements of W_{hf}^L are infinite. Actually, this is not the case. Consider the matrix element $\left\langle \varphi_{k,l,m} \mid W_{hf}^L \mid \varphi_{k',l',m'} \right\rangle$, where $|\varphi_{k,l,m}\rangle$ and $|\varphi_{k',l',m'}\rangle$ are the stationary states of the hydrogen atom found in Chapter VII. In the $\{|\mathbf{r}\rangle\}$ representation, we have:

$$\langle \mathbf{r} \mid \varphi_{k,l,m} \rangle = \varphi_{k,l,m}(\mathbf{r}) = R_{k,l}(r)\ Y_l^m(\theta, \varphi) \tag{14}$$

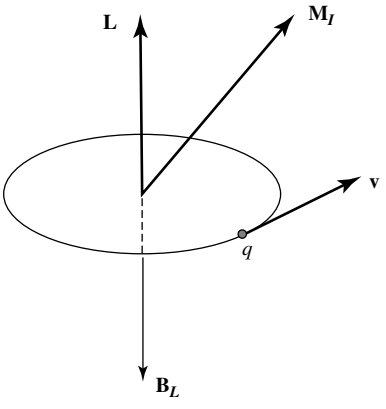


Figure 1: Relative disposition of the magnetic moment $\mathbf{M}_I$ *of the proton and the field* $\mathbf{B}_L$ *created by the current loop associated with the motion of the electron of charge* q *and velocity* $\mathbf{v}$ *(*$\mathbf{B}_L$ *is antiparallel to the orbital angular momentum* $\mathbf{L}$ *of the electron).*

with [*cf.* Chap. VII, relation (A-28)]:

$$R_{k,l}(r) \underset{r\to 0}{\sim} C\, r^l \tag{15}$$

With the presence of the $r^2\,\mathrm{d}r$ term in the integration volume element taken into account, the function to be integrated over r behaves at the origin like $r^{l+l'+2-3} = r^{l+l'-1}$. Furthermore, the presence of the Hermitian operator $\mathbf{L}$ in (13) means that the matrix element $\langle \varphi_{k,l,m} \mid W_{hf}^L \mid \varphi_{k',l',m'} \rangle$ is zero when l or l' is zero. We then have $l + l' \geq 2$, and $r^{l+l'-1}$ remains finite at the origin.

2-b. Coupling with the electron spin

We shall see that, for the last term of (5), the problems related to the singularity at the origin of the vector potential (4) are important. This is why, in studying this term, we shall choose a proton of finite size, letting its radius approach zero at the end of the calculations. Furthermore, from a physical point of view, we now know that the proton does possess a certain spatial extension and that its "magnetism" is spread over a certain volume. However, the dimensions of the proton are much smaller than the Bohr radius a_0. This justifies our treating the proton as a point particle in the final stage of the calculation.

α. *The magnetic field associated with the proton*

Consider the proton to be a particle of radius ρ_0 (Fig. 2), placed at the origin. The distribution of magnetism inside the proton creates, at a distant point, a field $\mathbf{B}$ which can be calculated by attributing to the proton a magnetic moment $\mathbf{M}_I$ which we shall choose parallel to Oz. For $r \gg \rho_0$, we obtain the components of $\mathbf{B}$ from the curl of the

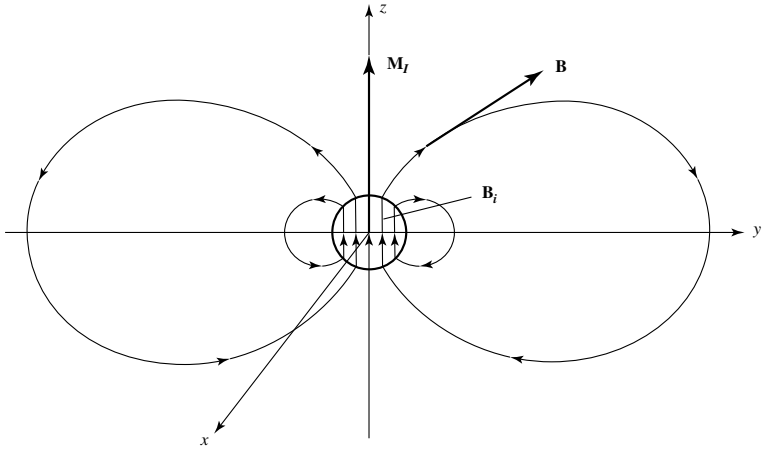


Figure 2: The magnetic field created by the proton. Outside the proton, the field is that of a dipole. Inside, the field depends on the exact distribution of the magnetism of the proton, but we can, in a first approximation, consider it to be uniform. The contact term corresponds to the interaction between the spin magnetic moment of the electron and this uniform field $\mathbf{B}_i$ inside the proton.

vector potential written in (4):

$$\begin{cases} B_x = \dfrac{\mu_0}{4\pi} 3M_I \dfrac{xz}{r^5} \\ B_y = \dfrac{\mu_0}{4\pi} 3M_I \dfrac{yz}{r^5} \\ B_z = \dfrac{\mu_0}{4\pi} M_I \dfrac{3z^2 - r^2}{r^5} \end{cases} \tag{16}$$

Expressions (16), moreover, remain valid even if r is not very large compared to ρ_0. We have already emphasized that the proton, since it is a spin 1/2 particle, has no magnetic multipole moments of order $k > 1$. The field outside the proton is therefore a pure dipole field.

Inside the proton, the magnetic field depends on the exact magnetic distribution. We shall assume this field $\mathbf{B}_i$ to be uniform (by symmetry, it must then be parallel to $\mathbf{M}_I$ and, therefore, to Oz)[2].

To calculate the field $\mathbf{B}_i$ inside the proton, we shall write the equation stating that the flux of the magnetic field through a closed surface, bounded by the xOy plane and the upper hemisphere centered at O and of infinite radius, is zero. Since, as $r \to \infty$, $|\mathbf{B}|$ decreases as $1/r^3$, the flux through this hemisphere is zero. Therefore, if $\Phi_i(\rho_0)$ denotes the flux through a disk centered at O of radius ρ_0 in the xOy plane, and $\Phi_e(\rho_0)$, the flux

[2]The following argument can be generalized to cases where $\mathbf{B}_i$ varies in a more complicated fashion (see comment (*ii*) at the end of this complement).

through the rest of the xOy plane, we have:

$$\Phi_i(\rho_0) + \Phi_e(\rho_0) = 0 \tag{17}$$

Relations (16) enable us to calculate $\Phi_e(\rho_0)$ easily, and we get:

$$\begin{aligned}\Phi_e(\rho_0) &= 2\pi \int_{\rho_0}^{+\infty} r\,\mathrm{d}r \left[-\frac{\mu_0}{4\pi} M_I \frac{1}{r^3}\right] \\ &= -\frac{\mu_0}{4\pi} M_I \frac{2\pi}{\rho_0}\end{aligned} \tag{18}$$

As for the flux $\Phi_i(\rho_0)$ of $\mathbf{B}_i$, it is equal to:

$$\Phi_i(\rho_0) = \pi \rho_0^2 B_i \tag{19}$$

so that (17) and (18) yield:

$$B_i = \frac{\mu_0}{4\pi} M_I \frac{2}{\rho_0^3} \tag{20}$$

Thus, we know the values of the field created by the proton at all points in space. We can now calculate the part of W_{hf} related to the electron spin $\mathbf{S}$.

β. The magnetic dipole term

If we substitute (16) into the term $-2\mu_B \left(\frac{\mathbf{S}}{\hbar}\right) \cdot \boldsymbol{\nabla} \times \mathbf{A}_I$, we obtain the operator:

$$W_{hf}^{\mathrm{dip}} = -\frac{\mu_0}{4\pi} \frac{2\mu_B M_I}{\hbar} \left\{ 3Z \frac{XS_x + YS_y + ZS_z}{R^5} - \frac{S_z}{R^3} \right\} \tag{21}$$

that is, taking into account the fact that $\mathbf{M}_I$ is, by hypothesis, parallel to Oz:

$$W_{hf}^{\mathrm{dip}} = \frac{\mu_0}{4\pi} \frac{2\mu_B}{\hbar} \frac{1}{R^3} \left\{ \mathbf{S} \cdot \mathbf{M}_I - 3 \frac{(\mathbf{S} \cdot \mathbf{R})(\mathbf{M}_I \cdot \mathbf{R})}{R^2} \right\} \tag{22}$$

This is the expression for the Hamiltonian of the dipole-dipole interaction between two magnetic moments $\mathbf{M}_I$ and $\mathbf{M}_S = 2\mu_B\, \mathbf{S}/\hbar$ (*cf.* Complement B_{XI}, § 1).

Actually, expression (16) for the magnetic field created by the proton is valid only for $r \geq \rho_0$, and (22) should be applied only to the part of the wave functions which satisfies this condition. However, when we let ρ_0 approach zero, expression (22) gives no singularity at the origin; it is therefore valid in all space.

Consider the matrix element:

$$\left\langle \varphi_{k,l,m,\varepsilon} \mid W_{hf}^{\mathrm{dip}} \mid \varphi_{k',l',m',\varepsilon'} \right\rangle$$

(we are adding here the indices ε and ε' to the states $\mid \varphi_{k,l,m} \rangle$ considered above in order to label the eigenvalues $\varepsilon\hbar/2$ and $\varepsilon'\hbar/2$ of S_z) and, in particular, the radial integral which corresponds to it. At the origin, the function of r to be integrated behaves like $r^{l+l'+2-3} = r^{l+l'-1}$. Now, according to condition (8-c) of Complement B_{XI}, the non-zero matrix elements are obtained for $l + l' \geq 2$. There is therefore no divergence at the origin. In the limit where $\rho_0 \longrightarrow 0$, the integral over r becomes an integral from 0 to infinity, and expression (22) is valid in all space.

γ. The contact term

We shall now substitute (20) into the last term of (5), so as to obtain the contribution of the internal field of the proton to W_{hf}. We then obtain an operator W_{hf}^c, which we shall call the "contact term", and whose matrix elements in the $\{|\varphi_{k,l,m,\varepsilon}\rangle\}$ representation are:

$$\langle\varphi_{k,l,m,\varepsilon}|W_{hf}^c\,|\varphi_{k',l',m',\varepsilon'}\rangle = -\frac{\mu_0}{4\pi}\frac{2\mu_B M_I}{\hbar}\,\langle\varepsilon\mid S_z\mid\varepsilon'\rangle\,\frac{2}{\rho_0^3}\iiint_{r\leq\rho_0}\mathrm{d}^3r\;\varphi_{k,l,m}^*(\mathbf{r})\;\varphi_{k',l',m'}(\mathbf{r}) \tag{23}$$

Let ρ_0 approach zero. The integration volume, $4\pi\rho_0^3/3$, also approaches zero, and the right-hand side of (23) becomes:

$$-\frac{\mu_0}{4\pi}\frac{2\mu_B M_I}{\hbar}\,\langle\varepsilon\mid S_z\mid\varepsilon'\rangle\,\frac{8\pi}{3}\;\varphi_{k,l,m}^*(\mathbf{r}=\mathbf{0})\;\varphi_{k',l',m'}(\mathbf{r}=\mathbf{0}) \tag{24}$$

The contact term is therefore given by:

$$W_{hf}^c = -\frac{\mu_0}{4\pi}\frac{8\pi}{3}\;\mathbf{M}_I\cdot\left(\frac{2\mu_B\mathbf{S}}{\hbar}\right)\delta(\mathbf{R}) \tag{25}$$

Although the volume containing an internal magnetic field (20) approaches zero when $\rho_0 \longrightarrow 0$, the value of W_{hf}^c remains finite, since this internal field approaches infinity as $1/\rho_0^3$.

Comments:

(*i*) In (25), the function $\delta(\mathbf{R})$ of the operator $\mathbf{R}$ is simply the projector:

$$\delta(\mathbf{R}) = |\mathbf{r}=\mathbf{0}\rangle\langle\mathbf{r}=\mathbf{0}| \tag{26}$$

(*ii*) The matrix element written in (23) is different from zero only if $l = l' = 0$. This is a necessary condition for $\varphi_{k,l,m}(\mathbf{r}=\mathbf{0})$ and $\varphi_{k',l',m'}(\mathbf{r}=\mathbf{0})$ to be non-zero (*cf.* Chap. VII, § C-4-c-β). The contact term therefore exists only for the *s* states.

(*iii*) In order to study, in § 2-a, the coupling between $\mathbf{M}_I$ and the orbital angular momentum of the electron, we assumed expression (4) for $\mathbf{A}_I(\mathbf{r})$ to be valid in all space. This amounts to ignoring the fact that the field $\mathbf{B}$ actually has the form (20) inside the proton. We might wonder if this procedure is correct, or if there is not also an orbital contact term in W_{hf}^L.

Actually, this is not the case. The term in $\mathbf{P}\cdot\mathbf{A}_I + \mathbf{A}_I\cdot\mathbf{P}$ would lead, for the field $\mathbf{B}_i$, to an operator proportional to:

$$\mathbf{B}_i\cdot\mathbf{L} = \frac{\mu_0}{4\pi}\mathbf{M}_I\frac{2}{\rho_0^3}L_z \tag{27}$$

Let us calculate the matrix element of such an operator in the $\{|\varphi_{k,l,m}\rangle\}$ representation. The presence of the operator L_z requires, as above, $l, l' \geq 1$. The radial function to be integrated between 0 and ρ_0 then behaves at the origin like $r^{l+l'+2}$ and therefore goes to zero at least as rapidly as r^4. Despite the presence of the $1/\rho_0^3$ term in (27), the integral between $r = 0$ and $r = \rho_0$ therefore goes to zero in the limit where $\rho_0 \longrightarrow 0$.

3. Conclusion: the hyperfine-structure Hamiltonian

Now, let us take the sum of the operators W_{hf}^{L}, W_{hf}^{dip} and W_{hf}^{c}. We use the fact that the magnetic dipole moment $\mathbf{M}_I$ of the proton is proportional to its angular momentum $\mathbf{I}$:

$$\mathbf{M}_I = g_p \mu_n \left(\frac{\mathbf{I}}{\hbar} \right) \tag{28}$$

(*cf.* § B-2-a of Chapter XII). We obtain:

$$W_{hf} = -\frac{\mu_0}{4\pi} \frac{2\mu_B \mu_n g_p}{\hbar^2} \left\{ \frac{\mathbf{I} \cdot \mathbf{L}}{R^3} + 3\frac{(\mathbf{I} \cdot \mathbf{R})(\mathbf{S} \cdot \mathbf{R})}{R^5} - \frac{\mathbf{I} \cdot \mathbf{S}}{R^3} + \frac{8\pi}{3} \mathbf{I} \cdot \mathbf{S}\, \delta(\mathbf{R}) \right\} \tag{29}$$

This operator acts both in the state space of the electron and in the state space of the proton. It can be seen that this is indeed the operator introduced in Chapter XII [*cf.* (B-20)].

Comments:

(*i*) We will now discuss the generalization of formula (29) to the case of an atom having a nuclear spin $I > 1/2$.

First of all, if $I = 1$, we have already seen that the nucleus can have an electric quadrupole moment which adds a contribution to the potential $V_0(\mathbf{r})$ given by (2). An electric quadrupole hyperfine term is therefore present in the hyperfine Hamiltonian, in addition to the magnetic dipole term (29). Since an electrical interaction does not directly affect the electron spin, this quadrupole term only acts on the orbital variables of the electrons.

If now $I > 1$, other nuclear electric or magnetic multipole moments can exist, increasing in number as I increases. The electric moments give rise to hyperfine terms acting only on the orbital electron variables, while the magnetic terms act on both the orbital and the spin variables. For elevated values of I, the hyperfine Hamiltonian has therefore a complex structure. In practice however, for the great majority of cases, one can limit the hyperfine Hamiltonian to magnetic dipole and electric quadrupole terms. This is due to the fact that the multipole nuclear moments of an order superior to 2 make extremely small contributions to the hyperfine atomic structures. These contributions are therefore difficult to observe experimentally. This arises essentially from the extremely small size of the nuclei compared to the spatial extent a_0 of the electronic wave functions.

(*ii*) The simplifying hypothesis which we have made concerning the field $\mathbf{B}(\mathbf{r})$ created by the proton (a perfectly uniform field within a sphere, a dipole field outside) is not essential. The form (25) of the magnetic dipole Hamiltonian remains valid whenever the nuclear magnetism has an arbitrary repartition, giving rise to more complicated internal fields $\mathbf{B}_i(\mathbf{r})$ (assuming however that the spatial extent of the nucleus is negligible compared to a_0; *cf.* the following comment). The argument is actually a direct generalization of that given in this complement. Consider a sphere S_ε centered at the origin, containing the nucleus and having a radius $\varepsilon \ll a_0$.

If $I = \frac{1}{2}$, the field outside S_ε has the form (16) and, since ε is very small compared to a_0, its contribution leads to the terms (13) and (22). As for the contribution of the field $\mathbf{B}(\mathbf{r})$ inside S_ε, it depends only on the value at the origin of the electronic wave functions and on the integral of $\mathbf{B}(\mathbf{r})$ inside S_ε. Since the flux of $\mathbf{B}(\mathbf{r})$ across all closed surfaces is zero, the integral in S_ε of each component of $\mathbf{B}(\mathbf{r})$ can be transformed into an integral outside of S_ε, where $\mathbf{B}(\mathbf{r})$ has the form (16). A simple calculation will again give exactly expression (25) which is therefore independent of the simplifying hypothesis that we have made.

If $I > \frac{1}{2}$, the nuclear contribution to the electromagnetic field outside of S_ε gives rise to the multipole hyperfine Hamiltonian which we have discussed in comment (*i*) above. On the other hand, one can easily show that the contribution of the field inside S_ε does not give rise to any new term: only the magnetic dipole possesses a contact term.

(*iii*) In all of the above, we have totally neglected the dimensions of the nucleus compared to those of the electronic wave functions (we have taken the limit $\rho_0/a_0 \to 0$). This is obviously not always realistic, in particular for heavy atoms whose nuclei have a relatively large spatial extension. If one studies these "volume effects" (keeping for example several of the lower order terms in ρ_0/a_0), a series of new terms appears in the electron-nucleus interaction Hamiltonian. We have already encountered this type of effect in Complement $\mathrm{D_{XI}}$ where we studied the effects of the radial distribution of the nuclear charge (nuclear multipole moments of order $k = 0$). Analogous phenomena occur concerning the spatial distribution of nuclear magnetism and lead to modifications of different terms of the hyperfine Hamiltonian (29). In particular, a new term must be added to the contact term (25) when the electronic wave functions vary significantly within the nucleus. This new term is neither simply proportional to $\delta(\mathbf{R})$, nor to the total magnetic moment of the nucleus. It depends on the spatial distribution of the nuclear magnetism. From a practical point of view, such a term is interesting since, using precise measurements of the hyperfine structure of heavy atoms, it permits obtaining information concerning the variation of the magnetism within the corresponding nuclei.

References and suggestions for further reading:

The hyperfine Hamiltonian including the electric quadrupole interaction term: Abragam (14.1), Chap. VI; Kuhn (11.1), Chap. VI, § B; Sobel'man (11.12), Chap. 6.

Complement B_{XII}

Calculation of the average values of the fine-structure Hamiltonian in the $1s$, $2s$ and $2p$ states

1 Calculation of $\langle 1/R \rangle$, $\langle 1/R^2 \rangle$ and $\langle 1/R^3 \rangle$ 1276
2 The average values $\langle W_{mv} \rangle$ 1278
3 The average values $\langle W_D \rangle$ 1279
4 Calculation of the coefficient ξ_{2p} associated with W_{SO} in the $2p$ level 1279

For the hydrogen atom, the fine-structure Hamiltonian W_f is the sum of three terms:

$$W_f = W_{mv} + W_{SO} + W_D \tag{1}$$

studied in § B-1 of Chapter XII.

The aim of this complement is to give the calculation of the average values of these three operators for the $1s$, $2s$ and $2p$ states of the hydrogen atom, a calculation which was omitted in Chapter XII for the sake of simplicity. We shall begin by calculating the average values of $\langle 1/R \rangle$, $\langle 1/R^2 \rangle$ and $\langle 1/R^3 \rangle$ in these states.

1. Calculation of $\langle 1/R \rangle$, $\langle 1/R^2 \rangle$ and $\langle 1/R^3 \rangle$

The wave function associated with a stationary state of the hydrogen atom is (*cf.* Chap. VII, § C):

$$\varphi_{n,l,m}(\mathbf{r}) = R_{n,l}(r) Y_l^m(\theta, \varphi) \tag{2}$$

$Y_l^m(\theta, \varphi)$ is a spherical harmonic. The expressions for the radial functions $R_{n,l}(r)$ corresponding to the $1s$, $2s$, $2p$ states are:

$$\begin{cases} R_{1,0}(r) = 2(a_0)^{-3/2}\, \mathrm{e}^{-r/a_0} \\ R_{2,0}(r) = 2(2a_0)^{-3/2} \left(1 - \dfrac{r}{2a_0}\right) \mathrm{e}^{-r/2a_0} \\ R_{2,1}(r) = (2a_0)^{-3/2} (3)^{-1/2} \dfrac{r}{a_0}\, \mathrm{e}^{-r/2a_0} \end{cases} \tag{3}$$

where a_0 is the Bohr radius:

$$a_0 = 4\pi\varepsilon_0 \frac{\hbar^2}{m_e q^2} = \frac{\hbar^2}{m_e e^2} \tag{4}$$

The Y_l^m are normalized with respect to θ and φ, so that the average value $\langle R^q \rangle$ of the qth power (where q is a positive or negative integer) of the operator R associated with

$r = |\mathbf{r}|$ in the state $|\varphi_{n,l,m}\rangle$ can be written[1]:

$$\langle R^q \rangle_{n,l,m} = \int_0^\infty r^{q+2} \, |R_{n,l}(r)|^2 \, \mathrm{d}r \tag{5}$$

It therefore does not depend on m. If (3) is substituted into (5), there appear integrals of the form:

$$I(k,p) = \int_0^\infty r^k \mathrm{e}^{-p\mathbf{r}/a_0} \, \mathrm{d}r \tag{6}$$

where p and k are integers. We shall assume here that $k \geq 0$, that is, $q \geq -2$. An integration by parts then yields directly:

$$\begin{aligned} I(k,p) &= \left[-\frac{a_0}{p} \mathrm{e}^{-p\mathbf{r}/a_0} \, r^k \right]_0^\infty + \frac{k a_0}{p} \int_0^\infty r^{k-1} \mathrm{e}^{-p\mathbf{r}/a_0} \, \mathrm{d}r \\ &= \frac{k a_0}{p} I(k-1,p) \end{aligned} \tag{7}$$

Since, furthermore:

$$I(0,p) = \int_0^\infty \mathrm{e}^{-p\mathbf{r}/a_0} \, \mathrm{d}r = \frac{a_0}{p} \tag{8}$$

we obtain, by recurrence:

$$I(k,p) = k! \left(\frac{a_0}{p} \right)^{k+1} \tag{9}$$

Now, let us apply this result to the average values to be determined. We obtain:

$$\begin{aligned} \langle 1/R \rangle_{1s} &= \frac{4}{a_0^3} \int_0^\infty r \, \mathrm{e}^{-2\mathbf{r}/a_0} \, \mathrm{d}r \\ &= \frac{4}{a_0^3} I(1,2) = \frac{1}{a_0} \end{aligned} \tag{10a}$$

$$\begin{aligned} \langle 1/R \rangle_{2s} &= \frac{4}{8a_0^3} \int_0^\infty r \left[1 - \frac{r}{2a_0} \right]^2 \mathrm{e}^{-\mathbf{r}/a_0} \, \mathrm{d}r \\ &= \frac{1}{2a_0^3} \left[I(1,1) - \frac{1}{a_0} I(2,1) + \frac{1}{4a_0^2} I(3,1) \right] \\ &= \frac{1}{4a_0} \end{aligned} \tag{10b}$$

$$\begin{aligned} \langle 1/R \rangle_{2p} &= \frac{1}{8a_0^3} \frac{1}{3} \int_0^\infty r \left(\frac{r}{a_0} \right)^2 \mathrm{e}^{-\mathbf{r}/a_0} \, \mathrm{d}r \\ &= \frac{1}{24a_0^5} I(3,1) = \frac{1}{4a_0} \end{aligned} \tag{10c}$$

[1]Of course, this average value exists only for values of q which make integral (5) convergent.

Similarly:

$$\langle 1/R^2\rangle_{1s} = \frac{4}{a_0^3} I(0,2) = \frac{2}{a_0^2} \tag{11a}$$

$$\langle 1/R^2\rangle_{2s} = \frac{1}{2a_0^3}\left[I(0,1) - \frac{1}{a_0} I(1,1) + \frac{1}{4a_0^2} I(2,1)\right] = \frac{1}{4a_0^2} \tag{11b}$$

$$\langle 1/R^2\rangle_{2p} = \frac{1}{24a_0^5} I(2,1) = \frac{1}{12a_0^2} \tag{11c}$$

It is clear that the expression for the average value of $1/R^3$ is meaningless for the $1s$ and $2s$ states [since integral (5) is divergent]. For the $2p$ state, it is equal to:

$$\langle 1/R^3\rangle_{2p} = \frac{1}{24a_0^5} I(1,1) = \frac{1}{24a_0^3} \tag{12}$$

2. The average values $\langle W_{mv}\rangle$

Let:

$$H_0 = \frac{\mathbf{P}^2}{2m_e} + V \tag{13}$$

be the Hamiltonian of the electron subjected to the Coulomb potential. We have:

$$\mathbf{P}^4 = 4m_e^2 \left[H_0 - V\right]^2 \tag{14a}$$

with:

$$V = -\frac{e^2}{R} \tag{14b}$$

so that:

$$W_{mv} = -\frac{1}{2m_e c^2} \left[H_0 - V\right]^2 \tag{15}$$

We shall take the average values of both sides of this expression in a state $|\varphi_{n,l,m}\rangle$. Since H_0 and V are Hermitian operators, we obtain:

$$\langle W_{mv}\rangle_{n,l,m} = -\frac{1}{2m_e c^2}\left[(E_n)^2 + 2E_n\, e^2 \langle 1/R\rangle_{n,l} + e^4 \langle 1/R^2\rangle_{n,l}\right] \tag{16}$$

In this expression, we have set:

$$E_n = -\frac{E_I}{n^2} = -\frac{1}{2n^2}\alpha^2 m_e c^2 \tag{17}$$

where:

$$\alpha = \frac{e^2}{\hbar c} \tag{18}$$

is the fine-structure constant.

If we apply relation (16) to the case of the $1s$ state, we obtain, using (10a) and (11a):

$$\langle W_{mv}\rangle_{1s} = -\frac{1}{2m_e c^2}\left[\frac{1}{4}\alpha^4 m_e^2 c^4 - \alpha^2 m_e c^2 \frac{e^2}{a_0} + 2\frac{e^4}{a_0^2}\right] \tag{19}$$

that is, since, according to (4) and (18), $e^2/a_0 = \alpha^2 m_e c^2$:

$$\langle W_{mv}\rangle_{1s} = -\frac{1}{2}\alpha^4 m_e c^2\left[\frac{1}{4} - 1 + 2\right] = -\frac{5}{8}\alpha^4 m_e c^2 \tag{20}$$

The same type of calculation, for the $2s$ state, leads to:

$$\langle W_{mv}\rangle_{2s} = -\frac{1}{2}\alpha^4 m_e c^2\left[\left(\frac{1}{8}\right)^2 - 2\frac{1}{8}\frac{1}{4} + \frac{1}{4}\right] = -\frac{13}{128}\alpha^4 m_e c^2 \tag{21}$$

and, for the $2p$ state, to:

$$\langle W_{mv}\rangle_{2p} = -\frac{1}{2}\alpha^4 m_e c^2\left[\left(\frac{1}{8}\right)^2 - 2\frac{1}{8}\frac{1}{4} + \frac{1}{12}\right] = -\frac{7}{384}\alpha^4 m_e c^2 \tag{22}$$

3. The average values $\langle W_D\rangle$

With (14b) and the fact that $\Delta(1/r) = -4\pi\delta(\mathbf{r})$ taken into account, the average value of W_D in the $|\varphi_{n,l,m}\rangle$ state can be written [see also formula (B-14) of Chapter XII]:

$$\langle W_D\rangle_{n,l,m} = \frac{\hbar^2}{8m_e^2 c^2} 4\pi e^2 \left|\varphi_{n,l,m}(\mathbf{r}=\mathbf{0})\right|^2 \tag{23}$$

This expression goes to zero if $\varphi_{n,l,m}(\mathbf{r}=\mathbf{0}) = 0$, that is, if $l \neq 0$. Therefore:

$$\langle W_D\rangle_{2p} = 0 \tag{24a}$$

For the $1s$ and $2s$ levels, we obtain, using (2), (23) and the fact that $Y_0^0 = 1/\sqrt{4\pi}$:

$$\langle W_D\rangle_{1s} = \frac{\hbar^2}{8m_e^2 c^2}\, e^2 \left|R_{1,0}(0)\right|^2 = \frac{1}{2}\alpha^4 m_e c^2 \tag{24b}$$

as well as:

$$\langle W_D\rangle_{2s} = \frac{\hbar^2}{8m_e^2 c^2}\, e^2 \left|R_{2,0}(0)\right|^2 = \frac{1}{16}\alpha^4 m_e c^2 \tag{24c}$$

4. Calculation of the coefficient ξ_{2p} associated with W_{SO} in the $2p$ level

In § C-2-c-β of Chapter XII, we defined the coefficient:

$$\xi_{2p} = \frac{e^2}{2m_e^2 c^2}\int_0^\infty \frac{\left|R_{2,1}(r)\right|^2}{r}\mathrm{d}r \tag{25}$$

According to (3):

$$\xi_{2p} = \frac{e^2}{2m_e^2c^2}\frac{1}{24a_0^5}I(1,1) \tag{26}$$

Relation (9) then yields:

$$\xi_{2p} = \frac{e^2}{2m_e^2c^2}\frac{1}{24a_0^3} = \frac{1}{48\hbar^2}\alpha^4 m_e c^2 \tag{27}$$

References:

Several radial integrals for hydrogen-like atoms are given in Bethe and Salpeter (11.10).

Complement C_{XII}

The hyperfine structure and the Zeeman effect for muonium and positronium

1 The hyperfine structure of the $1s$ ground state 1281

2 The Zeeman effect in the $1s$ ground state 1282

2-a The Zeeman Hamiltonian 1282

2-b Stationary state energies 1283

2-c The Zeeman diagram for muonium 1284

2-d The Zeeman diagram for positronium 1286

In Complement A_{VII}, we studied some hydrogen-like systems, composed, like the hydrogen atom, of two oppositely charged particles electrostatically attracted to each other. Of all these systems, two are particularly interesting: muonium (composed of an electron, e^-, and a positive muon, μ^+) and positronium (composed of an electron, e^-, and a positron, e^+). Their importance lies in the fact that the various particles which come into play (the electron, the positron and the muon) are not directly affected by strong interactions (while the proton is). The theoretical and experimental study of muonium and positronium therefore permits a very direct test of the validity of quantum electrodynamics.

Actually, a very precise information we now possess about these two systems comes from the study of the hyperfine structure of their $1s$ ground state [the optical lines joining the $1s$ ground state to the various excited states have been observed for positronium; *cf.* Ref. (11.25)]. This hyperfine structure is the result, as in the case of the hydrogen atom, of magnetic interactions between the spins of the two particles. We shall describe some interesting features of the hyperfine structure and the Zeeman effect for muonium and positronium in this complement.

1. The hyperfine structure of the $1s$ ground state

Let $\mathbf{S}_1$ be the electron spin and $\mathbf{S}_2$, the spin of the other particle (the muon or the positron, which are both spin 1/2 particles). The degeneracy of the $1s$ ground state is then, as for hydrogen, four-fold.

We can use stationary perturbation theory to study the effect on the $1s$ ground state of the magnetic interactions between $\mathbf{S}_1$ and $\mathbf{S}_2$. The calculation is analogous to the one in § D of Chapter XII. We are left with a problem of two spin 1/2's coupled by an interaction of the form:

$$\mathcal{A}\ \mathbf{S}_1 \cdot \mathbf{S}_2 \tag{1}$$

where $\mathcal{A}$ is a constant which depends on the system under study. We shall denote by $\mathcal{A}_H$, $\mathcal{A}_M$, $\mathcal{A}_P$ the three values of $\mathcal{A}$ which correspond respectively to hydrogen, muonium and positronium.

It is easy to see that:

$$\mathcal{A}_H < \mathcal{A}_M < \mathcal{A}_P \tag{2}$$

since the smaller the mass of particle (2), the larger its magnetic moment. Now the positron is about 200 times lighter than the muon, which is close to 10 times lighter than the proton.

Comment:

The theory of Chapter XII is insufficient for the extremely precise study of the hyperfine structure of hydrogen, muonium and positronium. In particular, the hyperfine Hamiltonian W_{hf} given in § B-2 of this chapter describes only part of the interactions between particles (1) and (2). For example, the fact that the electron and the positron are antiparticles of each other (they can annihilate to produce photons) is responsible for an additional coupling between them which has no equivalent for hydrogen and muonium. In addition, a series of corrections (relativistic, radiative, recoil effects, etc.) must be taken into account. These are complicated to calculate and must be treated by quantum electrodynamics. Finally, for hydrogen, nuclear corrections are also involved which are related to the structure and polarizability of the proton. However, it can be shown that the form (1) of the coupling between $\mathbf{S}_1$ and $\mathbf{S}_2$ remains valid, the constant $\mathcal{A}$ being given by an expression which is much more complicated than formula (D-8) of Chapter XII. The hydrogen-like systems studied in this complement are important precisely because they enable us to compare the theoretical value of $\mathcal{A}$ with experimental results.

The eigenstates of $\mathcal{A}\ \mathbf{S}_1 \cdot \mathbf{S}_2$ are the $|F, m_F\rangle$ states, where F and m_F are the quantum numbers related to the total angular momentum:

$$\mathbf{F} = \mathbf{S}_1 + \mathbf{S}_2 \tag{3}$$

As in the case of the hydrogen atom, F can take on two values, $F = 1$ and $F = 0$. The two levels, $F = 1$ and $F = 0$, have energies equal to $\mathcal{A}\hbar^2/4$ and $-3\mathcal{A}\hbar^2/4$, respectively. Their separation $\mathcal{A}\hbar^2$ gives the hyperfine structure of the $1s$ ground state. Expressed in MHz, this interval is equal to:

$$\frac{\hbar}{2\pi}\mathcal{A}_M = 4\,463.317 \pm 0.021 \text{ MHz} \tag{4}$$

for muonium, and:

$$\frac{\hbar}{2\pi}\mathcal{A}_P = 203\,403 \pm 12 \text{ MHz} \tag{5}$$

for positronium.

2. The Zeeman effect in the $1s$ ground state

2-a. The Zeeman Hamiltonian

If we apply a static field $\mathbf{B}_0$ parallel to Oz, we must add, to the hyperfine Hamiltonian (1), the Zeeman Hamiltonian which describes the coupling of $\mathbf{B}_0$ to the magnetic

moments:

$$\mathbf{M}_1 = \gamma_1 \mathbf{S}_1 \tag{6}$$

and:

$$\mathbf{M}_2 = \gamma_2 \mathbf{S}_2 \tag{7}$$

of the two spins, with gyromagnetic ratios γ_1 and γ_2. If we set:

$$\omega_1 = -\gamma_1 B_0 \tag{8}$$

$$\omega_2 = -\gamma_2 B_0 \tag{9}$$

this Zeeman Hamiltonian can be written:

$$\omega_1 S_{1z} + \omega_2 S_{2z} \tag{10}$$

In the case of hydrogen, the magnetic moment of the proton is much smaller than that of the electron. We used this property in § E-1 of Chapter XII to neglect the Zeeman coupling of the proton, compared to that of the electron[1]. Such an approximation is less justified for muonium, since the magnetic moment of the muon is larger than that of the proton. We shall therefore take both terms of (10) into account. For positronium, furthermore, they are equally important: the electron and positron have equal masses and opposite charges, so that:

$$\gamma_1 = -\gamma_2 \qquad \text{(positronium)} \tag{11}$$

or:

$$\omega_1 = -\omega_2 \qquad \text{(positronium)} \tag{12}$$

2-b. Stationary state energies

When B_0 is not zero, it is necessary, in order to find the stationary state energies, to diagonalize the matrix representing the total Hamiltonian:

$$\mathcal{A}\,\mathbf{S}_1 \cdot \mathbf{S}_2 + \omega_1 S_{1z} + \omega_2 S_{2z} \tag{13}$$

in an arbitrary orthonormal basis, for example, the $\{|F, m_F\rangle\}$ basis. A calculation which is analogous to the one in § E-4 of Chapter XII then leads to the following matrix (the four basis vectors are arranged in the order $|1,1\rangle$, $|1,-1\rangle$, $|1,0\rangle$, $|0,0\rangle$):

$$\left(\begin{array}{cc|cc}
\boxed{\frac{\mathcal{A}\hbar^2}{4} + \frac{\hbar}{2}(\omega_1+\omega_2)} & 0 & 0 & 0 \\
0 & \boxed{\frac{\mathcal{A}\hbar^2}{4} - \frac{\hbar}{2}(\omega_1+\omega_2)} & 0 & 0 \\
\hline
0 & 0 & \boxed{\frac{\mathcal{A}\hbar^2}{4}} & \boxed{\frac{\hbar}{2}(\omega_1-\omega_2)} \\
0 & 0 & \boxed{\frac{\hbar}{2}(\omega_1-\omega_2)} & \boxed{-\frac{3\mathcal{A}\hbar^2}{4}}
\end{array}\right) \tag{14}$$

[1] Recall that the gyromagnetic ratio of the electron spin is $\gamma_1 = 2\mu_B/\hbar$ (μ_B: the Bohr magneton). Thus, if we set $\omega_0 = -\mu_B B_0/\hbar$ (the Larmor angular frequency), the constant ω_1 defined by (8) is equal to $2\omega_0$ (this is, furthermore, the notation used in § E of Chapter XII; to obtain the results of that section, it therefore suffices, in this complement, to replace ω_1 by $2\omega_0$ and ω_2 by 0).

Matrix (14) can be broken down into two 1×1 submatrices and a 2×2 submatrix. Two eigenvalues are therefore obvious:

$$\begin{cases} E_1 = \dfrac{\mathcal{A}\hbar^2}{4} + \dfrac{\hbar}{2}(\omega_1 + \omega_2) & (15) \\ E_2 = \dfrac{\mathcal{A}\hbar^2}{4} - \dfrac{\hbar}{2}(\omega_1 + \omega_2) & (16) \end{cases}$$

They correspond, respectively, to the states $|1,1\rangle$ and $|1,-1\rangle$, which, moreover, coincide with the states $|+,+\rangle$ and $|-,-\rangle$ of the $\{|\varepsilon_1,\varepsilon_2\rangle\}$ basis of common eigenstates of S_{1z} and S_{2z}. The other two eigenvalues can be obtained by diagonalizing the remaining 2×2 submatrix. They are equal to:

$$E_3 = -\frac{\mathcal{A}\hbar^2}{4} + \sqrt{\left(\frac{\mathcal{A}\hbar^2}{2}\right)^2 + \frac{\hbar^2}{4}(\omega_1 - \omega_2)^2} \tag{17}$$

$$E_4 = -\frac{\mathcal{A}\hbar^2}{4} - \sqrt{\left(\frac{\mathcal{A}\hbar^2}{2}\right)^2 + \frac{\hbar^2}{4}(\omega_1 - \omega_2)^2} \tag{18}$$

In a weak field, they correspond to the states $|1,0\rangle$ and $|0,0\rangle$, respectively, and, in a strong field, to the states $|+,-\rangle$ and $|-,+\rangle$.

2-c. The Zeeman diagram for muonium

The only differences with the results of § E-4 of Chapter XII arise from the fact that here, we are taking the Zeeman coupling of particle (2) into account. These differences appear only in a sufficiently strong field.

Let us therefore consider the form taken on by the energies E_3 and E_4 when $\hbar(\omega_1 - \omega_2) \gg \mathcal{A}\hbar^2$. In this case: $\omega_1 \geq \omega_2$

$$E_3 \simeq -\frac{\mathcal{A}\hbar^2}{4} + \frac{\hbar}{2}(\omega_1 - \omega_2) \tag{19}$$

$$E_4 \simeq -\frac{\mathcal{A}\hbar^2}{4} - \frac{\hbar}{2}(\omega_1 - \omega_2) \tag{20}$$

Now, compare (19) with (15) and (20) with (16). We see that, in a strong field, the energy levels are no longer represented by pairs of parallel lines, as was the case in § E-3 of Chapter XII. The slopes of the asymptotes of the E_1 and E_3 levels are, respectively, $-\frac{\hbar}{2}(\gamma_1 + \gamma_2)$ and $-\frac{\hbar}{2}(\gamma_1 - \gamma_2)$; those of the E_2 and E_4 levels, $\frac{\hbar}{2}(\gamma_1 + \gamma_2)$ and $\frac{\hbar}{2}(\gamma_1 - \gamma_2)$. Since the two particles (1) and (2) have opposite charges, γ_1 and γ_2 have opposite signs. Consequently, in a sufficiently strong field, the E_3 level (which then corresponds to the $|+,-\rangle$ state) moves above the E_1 level (the $|+,+\rangle$ state), since its slope, $-\frac{\hbar}{2}(\gamma_1 - \gamma_2)$ is greater than $-\frac{\hbar}{2}(\gamma_1 + \gamma_2)$.

The distance between the E_1 and E_3 levels therefore varies in the following way with respect to B_0 (*cf.* Fig. 1): starting from 0, it increases to a maximum for the value of B_0 which makes the derivative of:

$$E_1 - E_3 = \frac{\mathcal{A}\hbar^2}{2} + f(B_0) \tag{21}$$

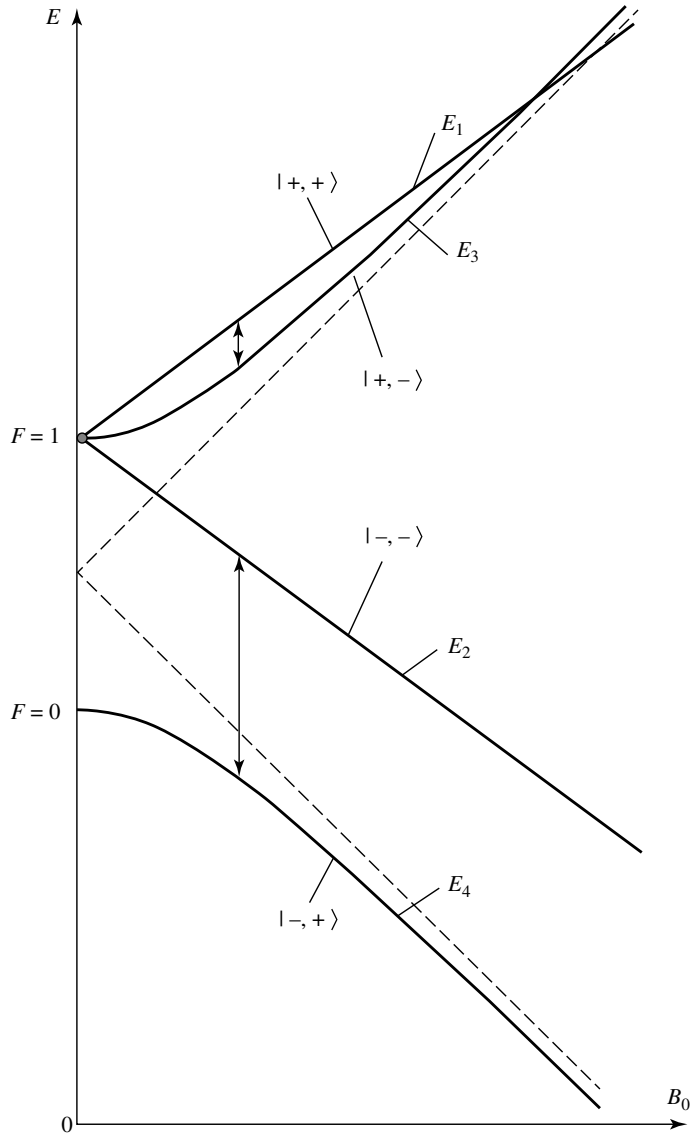


Figure 1: The Zeeman diagram for the $1s$ *ground state of muonium. Since we are not neglecting here the Zeeman coupling between the magnetic moment of the muon and the static field* $\mathbf{B}_0$*, the two straight lines (which correspond, in a strong field, to the same electron spin orientation but different muon spin orientations) are no longer parallel, as was the case for hydrogen (in the Zeeman diagram of Figure 9 of Chapter XII, the Larmor angular frequency* ω_n *of the proton was neglected). For the same value of the static field* B_0*, the splitting between the* E_1 *and* E_3 *levels is maximal and that between the* E_2 *and* E_4 *levels is minimal. The arrows represent the transitions studied experimentally for this value of the field* $\mathbf{B}_0$*.*

equal to zero, with:

$$f(B_0) = -\frac{\hbar}{2}(\gamma_1 + \gamma_2)B_0 - \sqrt{\left(\frac{\mathcal{A}\hbar^2}{2}\right)^2 + \frac{\hbar^2 B_0^2}{4}(\gamma_1 - \gamma_2)^2} \tag{22}$$

The distance then goes to zero again, and finally increases without bound. As for the distance between the E_2 and E_4 levels, it starts with the value $\mathcal{A}\hbar^2$, decreases to a minimum for the value of B_0 which makes the derivative of:

$$E_2 - E_4 = \frac{\mathcal{A}\hbar^2}{2} - f(B_0) \tag{23}$$

equal to zero and then increases without bound.

Since it is the same function $f(B_0)$ that appears in (21) and (23), we can show that, for the same value of B_0 [the one which makes the derivative of $f(B_0)$ go to zero], the distances between the E_1 and E_3 levels and between the E_2 and E_4 levels are either maximal or minimal. This property was recently used to improve the accuracy of experimental determinations of the hyperfine structure of muonium.

By stopping polarized muons (for example, in the $|+\rangle$ state) in a rare gas target, one can prepare, in a strong field, muonium atoms which will be found preferentially in the $|+,+\rangle$ and $|-,+\rangle$ states. If we then apply simultaneously two radio frequency fields whose frequencies are close to $(E_1 - E_3)/\hbar$ and $(E_2 - E_4)/\hbar$, we induce resonant transitions from $|+,+\rangle$ to $|+,-\rangle$ and from $|-,+\rangle$ to $|-,-\rangle$ (arrows in Figure 1). It is these transitions which are detected experimentally, since they correspond to a flip of the muon spin which is revealed by a change in the anisotropy of the positrons emitted during the β-decay of the muons. If we are operating in a field B_0 such that the derivative of $f(B_0)$ is zero, the inhomogeneities of the static field, which may exist from one point to another of the cell containing the rare gas, are not troublesome, since the resonant frequencies of muonium, $(E_1 - E_3)/h$ and $(E_2 - E_4)/h$, are not affected, to first order, by a variation of B_0 [ref. (11.24)].

Comment:

For the ground state of the hydrogen atom, we obtain a Zeeman diagram analogous to the one in Figure 1 when we take into account the Zeeman coupling between the proton spin and the field $\mathbf{B}_0$.

2-d. The Zeeman diagram for positronium

If we set $\omega_1 = -\omega_2$ (this property is a direct consequence of the fact that the positron is the antiparticle of the electron) in (15) and (16), we see that the E_1 and E_2 levels are independent of B_0:

$$E_1 = E_2 = \frac{\mathcal{A}\hbar^2}{4} \tag{24}$$

On the other hand, we obtain from (17) and (18):

$$E_3 = -\frac{\mathcal{A}\hbar^2}{4} + \sqrt{\left(\frac{\mathcal{A}\hbar^2}{2}\right)^2 + \hbar^2\gamma_1^2 B_0^2} \tag{25}$$

$$E_4 = -\frac{\mathcal{A}\hbar^2}{4} - \sqrt{\left(\frac{\mathcal{A}\hbar^2}{2}\right)^2 + \hbar^2\gamma_1^2 B_0^2} \tag{26}$$

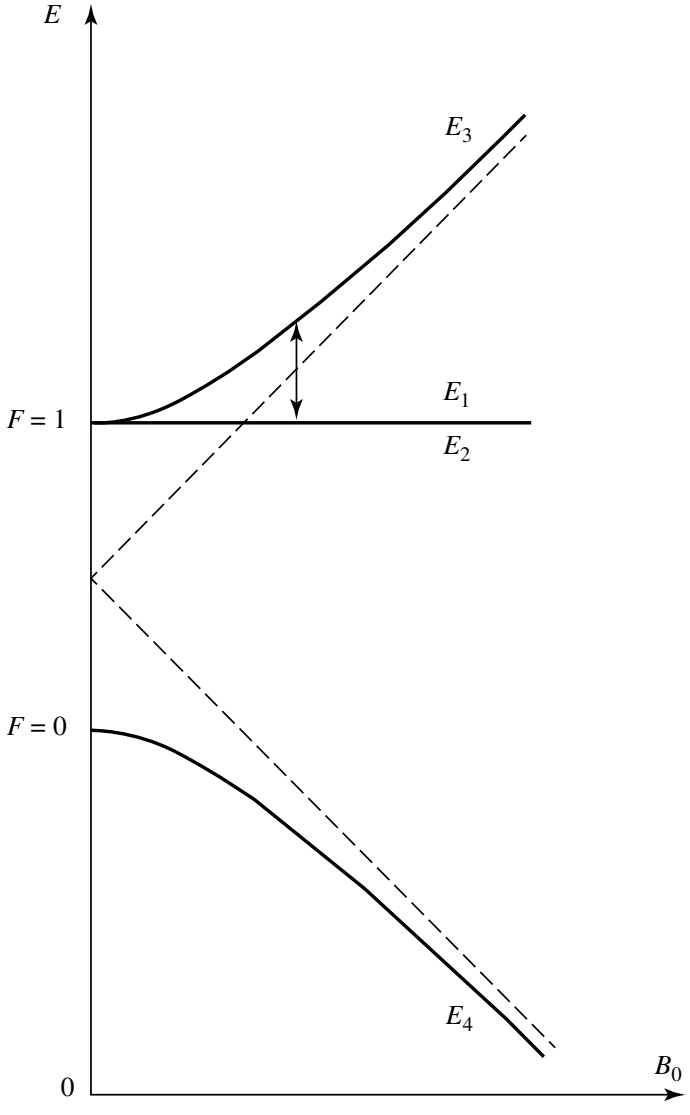


Figure 2: The Zeeman diagram for the $1s$ *ground state of positronium. As in the cases of hydrogen and muonium, this diagram is composed of one hyperbola and two straight lines. However, since the gyromagnetic ratios of the electron and positron are equal and opposite, the two straight lines have a zero slope and, consequently, are superposed (in the two corresponding states, with energy* E_1 *and* E_2*, the total magnetic moment is zero, since the electron and positron spins are parallel). The arrow represents the experimentally studied transition.*

The Zeeman diagram for positronium therefore has the form shown in Figure 2. It is composed of two superposed straight lines parallel to the B_0 axis and one hyperbola.

Actually, positronium is not stable. It decays by emitting photons. In a zero field, it can be shown by symmetry considerations that the $F = 0$ state (the singlet spin state, or "parapositronium") decays by emitting two photons. Its half-life is of the order of $\tau_0 \simeq 1{,}25\,10^{-10}$ s. On the other hand, the $F = 1$ state (the triplet spin state, or "orthopositronium") can decay only by emitting three photons (since the two-photon transition is forbidden). This process is much less probable, and the half-life of the triplet is much longer, on the order of $\tau_1 \simeq 1{,}4\,10^{-7}$ s.

When a static field is applied, the E_1 and E_2 levels retain the same lifetimes since the corresponding eigenstates do not depend on B_0. On the other hand, the $|1,0\rangle$ state is "mixed" with the $|0,0\rangle$ state, and vice versa. Calculations analogous to those of Complement H_{IV} show that the lifetime of the E_3 level is reduced relative to its zero-field value τ_1 (that of the E_4 level is increased relative to the value τ_0). The positronium atoms in the E_3 state then have a certain probability of decaying by emission of two photons.

This inequality of the lifetimes of the three states of energies E_1, E_2, E_3 when B_0 is non-zero is the basis of the methods for determining the hyperfine structure of positronium. Formation of positronium atoms by positron capture by an electron generally populates the four states of energies E_1, E_2, E_3, E_4 equally. In a non-zero field B_0, the two states E_1 and E_2 decay less rapidly than the E_3 state, so that in the stationary state, they are more populated. If we then apply a radiofrequency field oscillating at the frequency $(E_3 - E_1)/h = (E_3 - E_2)/h$, we induce resonant transitions from the E_1 and E_2 states to the E_3 state (the arrow of Figure 2). This increases the decay rate via two-photon emission, which permits the detection of resonance when (with fixed B_0) we vary the frequency of the oscillating field. Determination of $E_3 - E_1$ for a given value of B_0 then allows us to find the constant $\mathcal{A}$ by using (24) and (25).

In a zero field, resonant transitions could also be induced between the unequally populated $F = 1$ and $F = 0$ levels. However, the corresponding resonant frequency, given by (5), is high and not easily produced experimentally. This is why one generally prefers to use the "low frequency" transition represented by the arrow of Figure 2.

References and suggestions for further reading:

See the subsection "Exotic atoms" of section 11 of the bibliography.

The annihilation of positronium is discussed in Feynman III (1.2), § 18-3.

Complement D_{XII}

The influence of the electronic spin on the Zeeman effect of the hydrogen resonance line

1 Introduction 1289
2 The Zeeman diagrams of the $1s$ and $2s$ levels 1290
3 The Zeeman diagram of the $2p$ level 1290
4 The Zeeman effect of the resonance line 1293
4-a Statement of the problem 1293
4-b The weak-field Zeeman components 1295
4-c The strong-field Zeeman components 1295

1. Introduction

The conclusions of Complement D_{VII} relative to the Zeeman effect for the resonance line of the hydrogen atom spectrum (the $1s \leftrightarrow 2p$ transition) must be modified to take into account the electron spin and the associated magnetic interactions. This is what we shall do in this complement, using the results obtained in Chapter XII.

To simplify the discussion, we shall neglect effects related to nuclear spin (which are much smaller than those related to the electron spin). Therefore, we shall not take the hyperfine coupling W_{hf} (chap. XII, § B-2) into account, choosing the Hamiltonian H in the form:

$$H = H_0 + W_f + W_Z \tag{1}$$

H_0 is the electrostatic Hamiltonian studied in Chapter VII (§ C), W_f, the sum of the fine structure terms (chap. XII, § B-1):

$$W_f = W_{mv} + W_D + W_{SO} \tag{2}$$

and W_Z, the Zeeman Hamiltonian (chap. XII, § E-1) describing the interaction of the atom with a magnetic field $\mathbf{B}_0$ parallel to Oz:

$$W_Z = \omega_0(L_z + 2S_z) \tag{3}$$

where the Larmor angular frequency ω_0 is given by:

$$\omega_0 = -\frac{q}{2m_e}B_0 \tag{4}$$

[we shall neglect ω_n relative to ω_0; see formula (E-4) of Chapter XII].

We shall determine the eigenvalues and eigenvectors of H by using a method analogous to that of § E of Chapter XII: we shall treat W_f and W_Z like perturbations of H_0. Although they have the same unperturbed energy, the $2s$ and $2p$ levels can be studied

separately since they are connected neither by W_f (chap. XII, § C-2-a-β) nor by W_Z. In this complement, the magnetic field $\mathbf{B}_0$ will be called weak or strong, depending on whether W_Z is small or large compared to W_f. Note that the magnetic fields considered here to be "weak" are those for which W_Z is small compared to W_f but large compared to W_{hf}, which we have neglected. These "weak fields" are therefore much stronger than those treated in § E of Chapter XII.

Once the eigenstates and eigenvalues of H have been obtained, it is possible to study the evolution of the average values of the three components of the electric dipole moment of the atom. Since an analogous calculation was performed in detail in Complement D_{VII}, we shall not repeat it. We shall merely indicate, for weak fields and for strong fields, the frequencies and polarization states of the various Zeeman components of the resonance line of hydrogen (the Lyman α line).

2. The Zeeman diagrams of the $1s$ and $2s$ levels

We saw in § D-1-b of Chapter XII that W_f shifts the $1s$ level as a whole and gives rise to only one fine-structure level, $1s_{1/2}$. The same is true for the $2s$ level, which becomes $2s_{1/2}$. In each of these two levels, we can choose a basis:

$$\left\{\left|n; l=0; m_L=0; m_S=\pm\frac{1}{2}; m_I=\pm\frac{1}{2}\right\rangle\right\} \tag{5}$$

of eigenvectors common to H_0, $\mathbf{L}^2$, L_z, S_z, I_z (the notation is identical to that of Chapter XII; since H does not act on the proton spin, we shall ignore m_I in all that follows).

The vectors (5) are obviously eigenvectors of W_Z with eigenvalues $2m_S\hbar\omega_0$. Thus, each $1s_{1/2}$ or $2s_{1/2}$ level splits, in a field B_0, into two Zeeman sublevels of energies:

$$E(n; l=0; m_L=0; m_S) = E(ns_{1/2}) + 2m_S\hbar\omega_0 \tag{6}$$

where $E(ns_{1/2})$ is the zero-field energy of the $ns_{1/2}$ level, calculated in §§ C-2-b and D-1-b of Chapter XII. The Zeeman diagram of the $1s_{1/2}$ level (as well as the one for the $2s_{1/2}$ level) is therefore composed of two straight lines of slopes $+1$ and -1 (Fig. 1), corresponding, respectively, to the two possible orientations of the spin relative to $\mathbf{B}_0$ ($m_S = +1/2$ and $m_S = -1/2$).

Comparison of Figure 1 and Figure 9 of Chapter XII shows that to neglect, as we are doing here, the effects related to nuclear spin amounts to considering fields $\mathbf{B}_0$ which are so large that $W_Z \gg W_{hf}$. We are then in the asymptotic region of the diagram of Figure 9 of Chapter XII, where we can ignore the splitting of the energy levels due to the proton spin and hyperfine coupling.

3. The Zeeman diagram of the $2p$ level

In the six-dimensional $2p$ subspace, we can choose one of the two bases:

$$\{|n=2; l=1; m_L; m_S\rangle\} \tag{7}$$

or:

$$\{|n=2; l=1; J; m_J\rangle\} \tag{8}$$

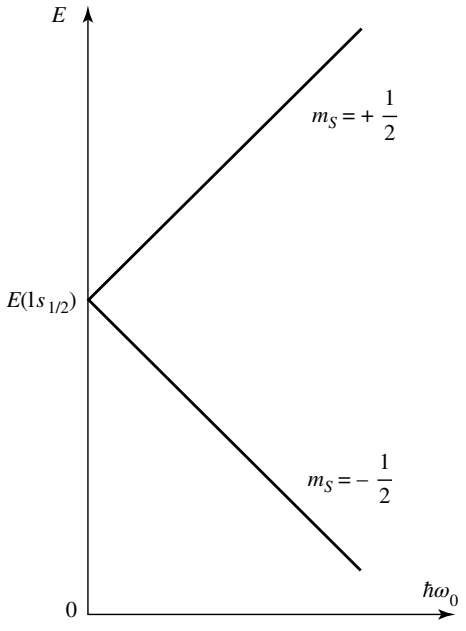


Figure 1: The Zeeman diagram of the $1s_{1/2}$ level when the hyperfine coupling W_{hf} is neglected. The ordinate of the point at which the two levels $m_S = \pm 1/2$ cross is the energy of the $1s_{1/2}$ level (i.e., the eigenvalue $-E_I$ of H_0, corrected for the global shift produced by the fine-structure Hamiltonian W_f). Figure 9 of Chapter XII gives an idea of the modifications of this diagram produced by W_{hf}.

adapted, respectively, to the individual angular momenta $\mathbf{L}$ and $\mathbf{S}$ and to the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ [*cf.* (36a) and (36b) of Complement A_X].

The terms W_{mv} and W_D which appear in expression (2) for W_f shift the $2p$ level as a whole. Therefore, to study the Zeeman diagram of the $2p$ level, we simply diagonalize the 6×6 matrix which represents $W_{SO} + W_Z$ in either one of the two bases, (7) or (8). Actually, since W_Z and $W_{SO} = \xi_{2p}\ \mathbf{L} \cdot \mathbf{S}$ both commute with $J_z = L_z + S_z$, this 6×6 matrix can be broken down into as many submatrices as there are distinct values of m_J. Thus, there appear two one-dimensional submatrices (corresponding respectively to $m_J = +3/2$ and $m_J = -3/2$) and two two-dimensional submatrices (corresponding respectively to $m_J = +1/2$ and $m_J = -1/2$). The calculation of the eigenvalues and associated eigenvectors (which is very much like that of § E-4 of Chapter XII) presents no difficulties and leads to the Zeeman diagram shown in Figure 2. This diagram is composed of two straight lines and four hyperbolic branches.

In a zero field, the energies depend only on J. We obtain the two fine-structure levels, $2p_{3/2}$ and $2p_{1/2}$, already studied in § C of Chapter XII, whose energies are equal to:

$$E(2p_{3/2}) = \tilde{E}(2p) + \frac{1}{2}\xi_{2p}\hbar^2 \tag{9}$$

$$E(2p_{1/2}) = \tilde{E}(2p) - \xi_{2p}\hbar^2 \tag{10}$$

$\tilde{E}(2p)$ is the $2p$ level energy $E(2p)$ corrected for the global shift due to W_{mv} and W_D [*cf.* expressions (C-8) and (C-9) of Chapter XII]. ξ_{2p} is the constant which appears in the restriction $\xi_{2p}\ \mathbf{L} \cdot \mathbf{S}$ of W_{SO} to the $2p$ level [*cf.* expression (C-13) of Chapter XII].

In weak magnetic fields ($W_Z \ll W_{SO}$), the slope of the energy levels can also be obtained by treating W_Z like a perturbation of W_f. It is then necessary to diagonalize the

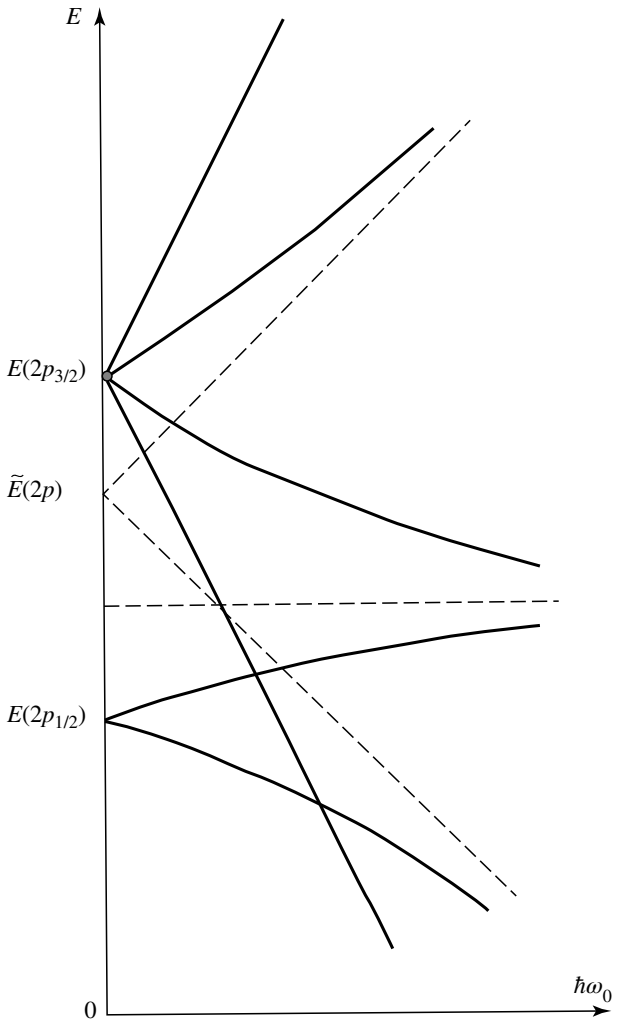


Figure 2: The Zeeman diagram of the 2p level when the hyperfine coupling W_{hf} is neglected. In a zero field, we find the fine-structure levels, $2p_{1/2}$ and $2p_{3/2}$. The Zeeman diagram is composed of two straight lines and two hyperbolas (for which the asymptotes are shown in dashed lines). The hyperfine coupling W_{hf} would significantly modify this diagram only in the neighborhood of $\omega_0 = 0$. $\tilde{E}(2p)$ is the 2p level energy (the eigenvalue $-E_I/4$ of H_0) corrected for the global shift produced by $W_{mv} + W_D$.

4×4 and 2×2 matrices representing W_Z in the $2p_{3/2}$ and $2p_{1/2}$ levels. Calculations analogous to those of § E-2 of chapter XII show that these two submatrices are respectively proportional to those which represent $\omega_0 J_z$ in the same subspaces. The proportionality coefficients, called "Landé factors" (*cf.* Complement D_X, § 3), are equal, respectively,

to[1]:

$$g(2p_{3/2}) = \frac{4}{3} \tag{11}$$

$$g(2p_{1/2}) = \frac{2}{3} \tag{12}$$

In weak fields, each fine-structure level therefore splits into $2J + 1$ equidistant Zeeman sublevels. The eigenstates are the states of the "coupled" basis, (8), corresponding to the eigenvalues:

$$E(J, m_J) = E(2p_J) + m_J\, g(2p_J)\, \hbar\omega_0 \tag{13}$$

where the $E(2p_J)$ are given by expressions (9) and (10).

In strong fields ($W_Z \gg W_{SO}$), we can, on the other hand, treat $W_{SO} = \xi_{2p}\ \mathbf{L}\cdot\mathbf{S}$ like a perturbation of W_Z, which is diagonal in basis (7). As in § E-3-b of chapter XII, it can easily be shown that only the diagonal elements of $\xi_{2p}\ \mathbf{L}\cdot\mathbf{S}$ are involved when the corrections are calculated to first order in W_{SO}. Thus, we find that in strong fields, the eigenstates are the states of the "decoupled" basis, (7), and the corresponding eigenvalues are:

$$E(m_L, m_S) = \tilde{E}(2p) + (m_L + 2m_S)\hbar\omega_0 + m_L m_S\ \hbar^2\xi_{2p} \tag{14}$$

Formula (14) gives the asymptotes of the diagram of Figure 2.

As the magnetic field B_0 increases, we pass continuously from basis (8) to basis (7). The magnetic field gradually decouples the orbital angular momentum and the spin. This situation is the analogue of the one studied in § E of Chapter XII, in which the angular momenta **S** and **I** were coupled or decoupled, depending on the relative importance of the hyperfine and Zeeman terms.

4. The Zeeman effect of the resonance line

4-a. Statement of the problem

Arguments of the same type as those of § 2-c of Complement $\mathrm{D_{VII}}$ (see, in particular, the comment at the end of that complement) show that the optical transition between a $2p$ Zeeman sublevel and a $1s$ Zeeman sublevel is possible only if the matrix element of the electric dipole operator $q\mathbf{R}$ between these two states is different from zero[2]. In addition, depending on whether it is the $q(X + iY)$, $q(X - iY)$ or qZ operator which has a non-zero matrix element between the two Zeeman sublevels under consideration, the polarization state of the emitted light is σ^+, σ^- or π. Therefore, we use the previously determined eigenvectors and eigenvalues of H in order to obtain the frequencies of the various Zeeman components of the hydrogen resonance line and their polarization states.

Comment:

[1] These Landé factors can be calculated directly from formula (43) of Complement $\mathrm{D_X}$.

[2] The electric dipole, since it is an odd operator, has no matrix elements between the $1s$ and $2s$ states, which are both even. This is why we are ignoring the $2s$ states here.

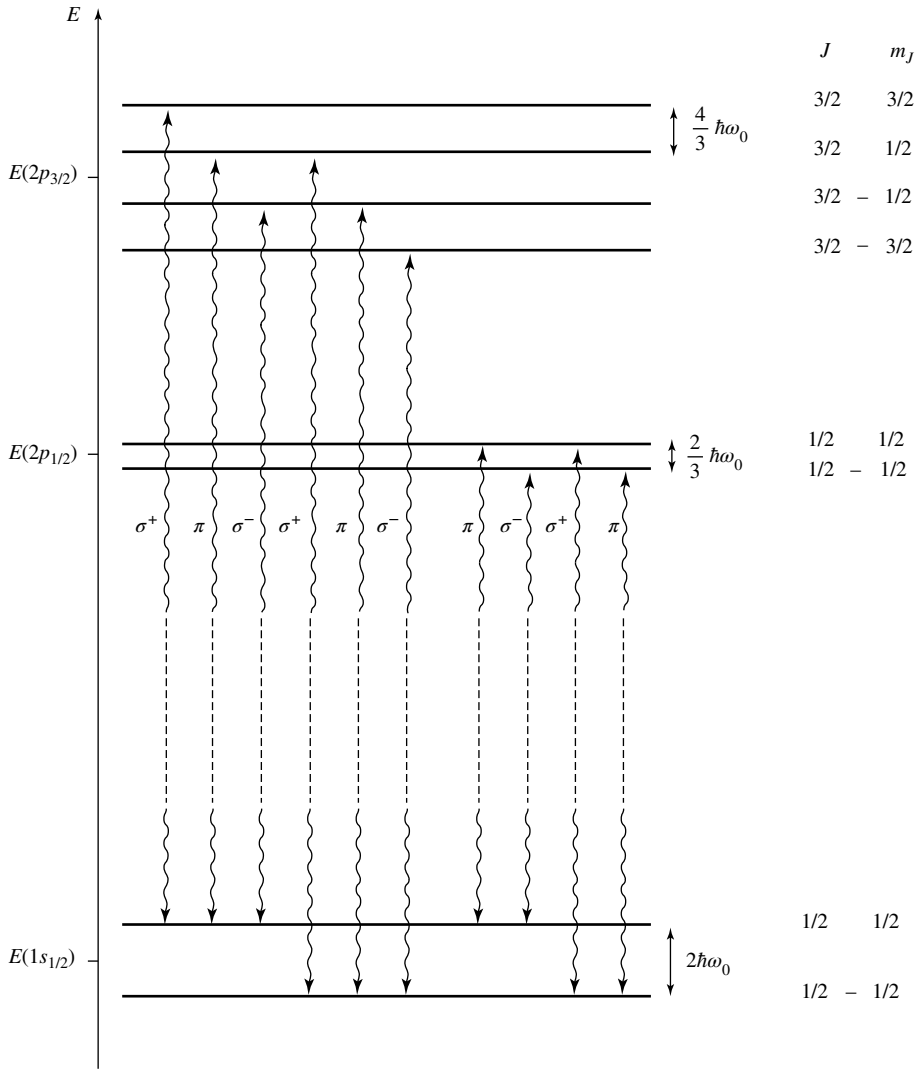


Figure 3: The disposition, in a weak field, of the Zeeman sublevels arising from the fine-structure levels, $1s_{1/2}$, $2p_{1/2}$, $2p_{3/2}$ (whose zero-field energies are marked on the vertical energy scale). On the right-hand side of the figure are indicated the splittings between adjacent Zeeman sublevels (for greater clarity, these splittings have been exaggerated with respect to the fine-structure splitting which separates the $2p_{1/2}$ and $2p_{3/2}$ levels), as well as the values of the quantum numbers J and m_J associated with each sublevel. The arrows indicate the Zeeman components of the resonance line, each of which has a well-defined polarization, σ^+, σ^- or π.

The $q(X + iY)$, $q(X - iY)$ and qZ operators act only on the orbital part of the wave function and cause m_L to vary, respectively, by $+1$, -1 and 0 (*cf.* Complement D_{VII}, § 2-c); m_S is not affected. Since $m_J = m_L + m_S$ is a good quantum number (whatever the strength of the field B_0), $\Delta m_J = +1$ transitions have a σ^+ polarization; $\Delta m_J = -1$ transitions, a σ^- polarization; and $\Delta m_J = 0$ transitions, a π polarization.

4-b. The weak-field Zeeman components

Figure 3 shows the weak-field positions of the various Zeeman sublevels resulting from the $1s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$ levels, obtained from expressions (6), (13), (11) and (12). The vertical arrows indicate the various Zeeman components of the resonance line. The polarization is σ^+, σ^- or π, depending on whether $\Delta m_J = +1, -1$ or 0.

Figure 4 shows the positions of these various components on a frequency scale, relative to the zero-field positions of the lines. The result differs notably from that of Complement D_{VII} (see Figure 2 of that complement), where, observing in a direction perpendicular to $\mathbf{B}_0$, we had three equidistant components of polarization σ^+, π, σ^-, separated by a frequency difference $\omega_0/2\pi$.

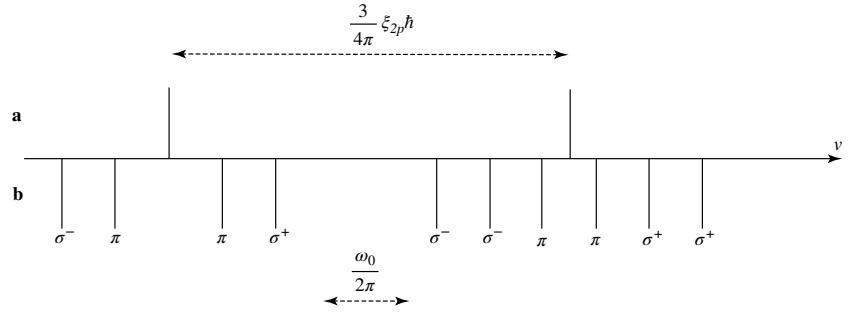


Figure 4: Frequencies of the various Zeeman components of the hydrogen resonance line. a) In a zero field: two lines are observed, separated by the fine-structure splitting $3\xi_{2p}\hbar/4\pi$ (ξ_{2p} is the spin-orbit coupling constant of the 2p level) and corresponding respectively to the transitions $2p_{3/2} \leftrightarrow 1s_{1/2}$ (the line on the right-hand side of the figure) and $2p_{1/2} \leftrightarrow 1s_{1/2}$ (the line on the left-hand side).
b) In a weak field B_0: each line splits into a series of Zeeman components whose polarizations are indicated; $\omega_0/2\pi$ is the Larmor frequency in the field B_0.

4-c. The strong-field Zeeman components

Figure 5 shows the strong-field positions of the Zeeman sublevels arising from the $1s$ and $2p$ levels [see expressions (6) and (14)]. To first order in W_{SO}, the degeneracy between the states $|m_L = -1, m_S = 1/2\rangle$ and $|m_L = 1, m_S = -1/2\rangle$ is not removed. The vertical arrows indicate the Zeeman components of the resonance line. The polarization is σ^+, σ^- or π, depending on whether $\Delta m_L = +1, -1$ or 0 (recall that in an electric dipole transition, the quantum number m_S is not affected).

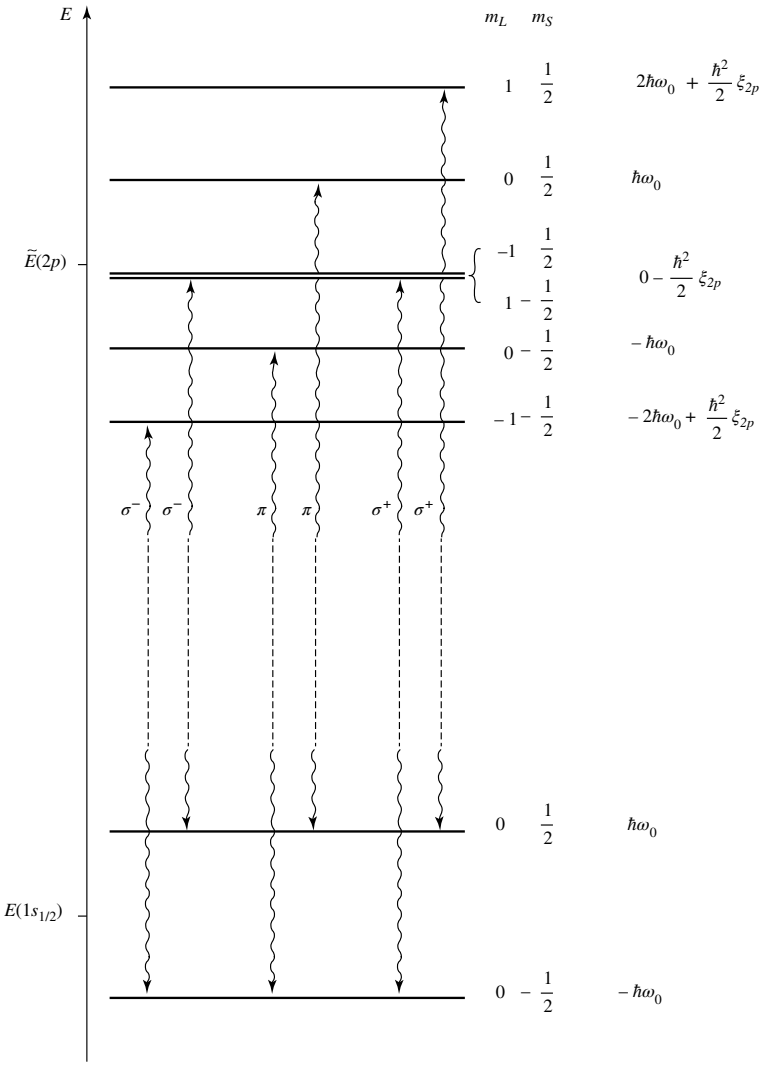


Figure 5: The disposition, in a strong field (decoupled fine structure), of the Zeeman sublevels arising from the $1s$ and $2p$ levels. On the right-hand side of the figure are indicated the values of the quantum numbers m_L and m_S associated with each Zeeman sublevel, as well as the corresponding energy, given relative to $E(1s_{1/2})$ or $\tilde{E}(2p)$. The vertical arrows indicate the Zeeman components of the resonance line.

The corresponding optical spectrum is shown in Figure 6. The two π transitions have the same frequency (*cf.* Fig. 5). On the other hand, there is a small splitting, $\hbar\xi_{2p}/2\pi$, between the frequencies of the two σ^+ transitions and between those of the two σ^- transitions. The mean distance between the σ^+ doublet and the π line (or between the π line and the σ^- doublet) is equal to $\omega_0/2\pi$. The spectrum of Figure 6 is therefore

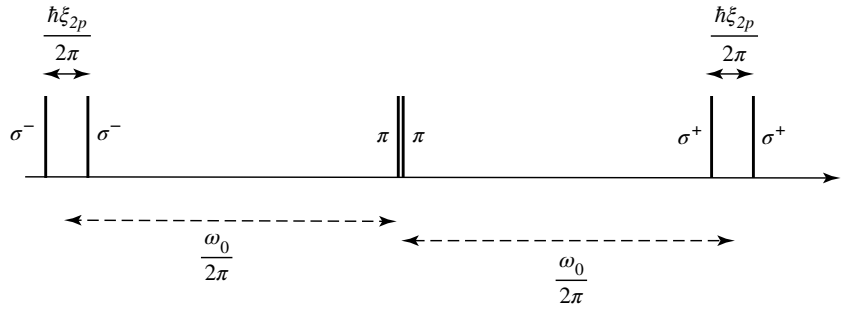


Figure 6: The strong-field positions of the Zeeman components of the hydrogen resonance line. Aside from the splitting of the σ^+ and σ^- lines, this spectrum is identical to the one obtained in Complement D_{VII}, where the effects related to electron spin were ignored.

similar to that of Figure 2 of Complement D_{VII}. Furthermore, the splitting of the σ^+ and σ^- lines, due to the existence of the electron spin, is easy to understand.

In strong fields, **L** and **S** are decoupled. Since the $1s \leftrightarrow 2p$ transition is an electric dipole transition, only the orbital angular momentum **L** of the electron is affected by the optical transition. An argument analogous to the one in § E-3-b of Chapter XII shows that the magnetic interactions related to the spin can be described by an "internal field" which adds to the external field $\mathbf{B}_0$ and whose sign changes, depending on whether the spin points up or down. It is this internal field that causes the splitting of the σ^+ and σ^- lines (the π line is not affected, since its quantum number m_L is zero).

References and suggestions for further reading:

Cagnac and Pebay-Peyroula (11.2), Chaps. XI and XVII (especially § 5-A of that chapter); White (11.5), Chap. X; Kuhn (11.1), Chap. III, § F; Sobel'man (11.12), Chap. 8, § 29.

Complement E_{XII}

The Stark effect for the hydrogen atom

1 The Stark effect on the $n = 1$ level 1298

1-a The shift of the $1s$ state is quadratic in $\mathcal{E}$ 1298

1-b Polarizability of the $1s$ state 1299

2 The Stark effect on the $n = 2$ level 1300

Consider a hydrogen atom placed in a uniform static electric field $\boldsymbol{\mathcal{E}}$ parallel to Oz. To the Hamiltonian studied in Chapter XII must be added the Stark Hamiltonian W_S, which describes the interaction energy of the electric dipole moment $q\mathbf{R}$ of the atom with the field $\mathcal{E}$. W_S can be written:

$$W_S = -q \ \boldsymbol{\mathcal{E}} \cdot \mathbf{R} = -q \, \mathcal{E} Z \tag{1}$$

Even for the strongest electric fields that can be produced in the laboratory, we always have $W_S \ll H_0$. On the other hand, if $\mathcal{E}$ is strong enough, W_S can have the same order of magnitude as W_f and W_{hf} or be even larger. To simplify the discussion, we shall assume throughout this complement that $\mathcal{E}$ is strong enough for the effect of W_S to be much larger than that of W_f or W_{hf}. We shall therefore calculate directly, using perturbation theory, the effect of W_S on the eigenstates of H_0 found in Chapter VII (the next step, which we shall not consider here, would consist of evaluating the effect of W_f, and then of W_{hf}, on the eigenstates of $H_0 + W_S$).

Since both H_0 and W_S do not act on the spin variables, we shall ignore the quantum numbers m_S and m_I.

1. The Stark effect on the $n = 1$ level

1-a. The shift of the $1s$ state is quadratic in $\mathcal{E}$

According to perturbation theory, the effect of the electric field can be obtained to first order by calculating the matrix element:

$$-q\mathcal{E}\langle n = 1, l = 0, m_L = 0|Z|n = 1, l = 0, m_L = 0\rangle$$

Since the operator Z is odd, and since the ground state has a well-defined parity (it is even), the preceding matrix element is zero.

There is therefore no effect which is linear in $\mathcal{E}$, and we must go on to the next term of the perturbation series:

$$\varepsilon_2 = q^2\mathcal{E}^2 \sum_{n\neq 1,l,m} \frac{|\langle 1, 0, 0|Z|n, l, m\rangle|^2}{E_1 - E_n} \tag{2}$$

where $E_n = -E_I/n^2$ is the eigenvalue of H_0 associated with the eigenstate $|n, l, m\rangle$ (*cf.* Chap. VII, § C). The preceding sum is certainly not zero, since there exist states $|n, l, m\rangle$

whose parity is opposite to that of $|1,0,0\rangle$. We conclude that, to lowest order in $\mathcal{E}$, the Stark shift of the $1s$ ground state is quadratic. Since $E_1 - E_n$ is always negative, the ground state is lowered.

1-b. Polarizability of the $1s$ state

We have already mentioned that, for reasons of parity, the average values of the components of the operator $q\mathbf{R}$ are zero in the state $|1,0,0\rangle$ (the unperturbed ground state).

In the presence of an electric field $\boldsymbol{\mathcal{E}}$ parallel to Oz, the ground state is no longer $|1,0,0\rangle$, but rather (according to the results of § B-1-b of Chapter XI):

$$|\psi_0\rangle = |1,0,0\rangle - q\mathcal{E} \sum_{n\neq 1,l,m} |n,l,m\rangle \frac{\langle n,l,m|Z|1,0,0\rangle}{E_1 - E_n} + ... \tag{3}$$

This shows that the average value of the electric dipole moment $q\mathbf{R}$ in the perturbed ground state is, to first order in $\mathcal{E}$, $\langle\psi_0|q\mathbf{R}|\psi_0\rangle$. Using expression (3) for $|\psi_0\rangle$, we then obtain:

$$\langle\psi_0|q\mathbf{R}|\psi_0\rangle = -q^2\mathcal{E} \sum_{n\neq 1,l,m} \frac{\langle 1,0,0|\mathbf{R}|n,l,m\rangle\langle n,l,m|Z|1,0,0\rangle + \langle 1,0,0|Z|n,l,m\rangle\langle n,l,m|\mathbf{R}|1,0,0\rangle}{E_1 - E_n} \tag{4}$$

Thus, we see that the electric field $\mathcal{E}$ causes an "induced" dipole moment to appear, proportional to $\mathcal{E}$. It can easily be shown, by using the spherical harmonic orthogonality relation[1], that $\langle\psi_0|qX|\psi_0\rangle$ and $\langle\psi_0|qY|\psi_0\rangle$ are zero, and that the only non-zero average value is:

$$\langle\psi_0|qZ|\psi_0\rangle = -2q^2\mathcal{E} \sum_{\substack{n\neq 1\\ l,m}} \frac{|\langle n,l,m|Z|1,0,0\rangle|^2}{E_1 - E_n} \tag{5}$$

In other words, the induced dipole moment is parallel to the applied field $\mathcal{E}$.This is not surprising, given the spherical symmetry of the $1s$ state. The coefficient of proportionality χ between the induced dipole moment and the field is called the linear electric susceptibility. We see that quantum mechanics permits the calculation of this susceptibility for the $1s$ state:

$$\chi_{1s} = -2q^2 \sum_{\substack{n\neq 1\\ l,m}} \frac{|\langle n,l,m|Z|1,0,0\rangle|^2}{E_1 - E_n} \tag{6}$$

[1]This relation implies that $\langle 1,0,0|Z|n,l,m\rangle$ is different from zero only if $l = 1$, $m = 0$ (the argument is the same as the one given for $\langle 2,1,m|Z|2,0,0\rangle$ in the beginning of § 2 below). Consequently, in (2), (3), (4), (5), (6), the summation is actually carried out only over n (it includes, furthermore, the states of the positive energy continuum).

2. The Stark effect on the $n = 2$ level

The effect of W_S on the $n = 2$ level can be obtained to first order by diagonalizing the restriction of W_S to the subspace spanned by the four states of the $\{|2,0,0\rangle; |2,1,m\rangle, m = -1,0,+1\}$ basis.

The $|2,0,0\rangle$ state is even; the three $|2,1,m\rangle$ states are odd. Since W_S is odd, the matrix element $\langle 2,0,0|W_S|2,0,0\rangle$ and the nine matrix elements $\langle 2,1,m'|W_S|2,1,m\rangle$ are zero (*cf.* Complement F_{II}). On the other hand, since the $|2,0,0\rangle$ and $|2,1,m\rangle$ states have opposite parities, $\langle 2,1,m|W_S|2,0,0\rangle$ can be different from zero.

Let us show that actually only $\langle 2,1,0|W_S|2,0,0\rangle$ is non-zero. W_S is proportional to $Z = R\cos\theta$ and, therefore, to $Y_1^0(\theta)$. The angular integral which enters into the matrix elements $\langle 2,1,m|W_S|2,0,0\rangle$ is therefore of the form:

$$\int Y_1^{m*}(\Omega)\ Y_1^0(\Omega)\ Y_0^0(\Omega)\ \mathrm{d}\Omega$$

Since Y_0^0 is a constant, this integral is proportional to the scalar product of Y_1^0 and Y_1^m and is therefore different from zero only if $m = 0$. Moreover, since Y_1^0, $R_{21}(r)$ and $R_{20}(r)$ are real, the corresponding matrix element of W_S is real. We shall set:

$$\langle 2,1,0|W_S|2,0,0\rangle = \gamma\mathcal{E} \tag{7}$$

without concerning ourselves with the exact value of γ [which could be calculated without difficulty since we know the wave functions $\varphi_{2,1,0}(\mathbf{r})$ and $\varphi_{2,0,0}(\mathbf{r})$].

The matrix which represents W_S in the $n = 2$ level, therefore, has the following form (the basis vectors are arranged in the order $|2,1,1\rangle$, $|2,1,-1\rangle$, $|2,1,0\rangle$, $|2,0,0\rangle$):

0	0	0	0
0	0	0	0
0	0	0	$\gamma\mathcal{E}$
0	0	$\gamma\mathcal{E}$	0

(8)

We can immediately deduce the corrections to first order in $\mathcal{E}$ and the eigenstates to zeroth order:

Eigenstates		*Corrections*
$\|2,1,1\rangle$	$\leftrightarrow$	0
$\|2,1,-1\rangle$	$\leftrightarrow$	0
$\frac{1}{\sqrt{2}}(\|2,1,0\rangle + \|2,0,0\rangle)$	$\leftrightarrow$	$\gamma\mathcal{E}$
$\frac{1}{\sqrt{2}}(\|2,1,0\rangle - \|2,0,0\rangle)$	$\leftrightarrow$	$-\gamma\mathcal{E}$

(9)

Thus, we see that the degeneracy of the $n = 2$ level is partially removed and that the energy shifts are *linear,* and not quadratic, in $\mathcal{E}$. The appearance of a linear Stark effect is a typical result of the existence of two levels of opposite parities and the same energy, here the $2s$ and $2p$ levels. This situation exists only in the case of hydrogen (because of the l-fold degeneracy of the $n \neq 1$ shells).

Comment:

The states of the $n = 2$ level are not stable. Nevertheless, the lifetime of the $2s$ state is considerably longer than that of the $2p$ states, since the atom passes easily from $2p$ to $1s$ by spontaneous emission of a Lyman α photon (lifetime of the order of 10^{-9} s), while decay from the $2s$ state requires the emission of two photons (lifetime of the order of a second). For this reason, the $2p$ states are said to be unstable and the $2s$ state, metastable.

Since the Stark Hamiltonian W_S has a non-zero matrix element between $2s$ and $2p$, any electric field (static or oscillating) "mixes" the metastable $2s$ state with the unstable $2p$ state, greatly reducing the $2s$ state's lifetime. This phenomenon is called "metastability quenching" (see also Complement H_{IV}, in which we study the effect of a coupling between two states of different lifetimes).

References and suggestions for further reading:

The Stark effect in atoms: Kuhn (11.1), Chap. III, §§ A-6 and G. Ruark and Urey (11.9), Chap. V, §§ 12 and 13; Sobel'man (11.12), Chap. 8, § 28.

The summation over the intermediate states which appears in (2) and (6) can be calculated exactly by the method of Dalgarno and Lewis; see Borowitz (1.7), § 14-5; Schiff (1.18), § 33. Original references: (2.34), (2.35), (2.36).

Quenching of metastability: see Lamb and Retherford (3.11), App. II; Sobel'man (11.12), Chap. 8, § 28.5.

Chapter XIII

Approximation methods for time-dependent problems

A Statement of the problem 1303
B Approximate solution of the Schrödinger equation 1305
B-1 The Schrödinger equation in the $\{|\varphi_n\rangle\}$ representation 1305
B-2 Perturbation equations 1306
B-3 Solution to first order in λ 1307
C An important special case: a sinusoidal or constant perturbation 1309
C-1 Application of the general equations 1309
C-2 Sinusoidal perturbation coupling two discrete states: the resonance phenomenon 1311
C-3 Coupling with the states of the continuous spectrum 1316
D Random perturbation 1320
D-1 Statistical properties of the perturbation 1320
D-2 Perturbative computation of the transition probability 1321
D-3 Validity of the perturbation treatment 1323
E Long-time behavior for a two-level atom 1324
E-1 Sinusoidal perturbation 1324
E-2 Random perturbation 1325
E-3 Broadband optical excitation of an atom 1332

A. Statement of the problem

Consider a physical system with Hamiltonian H_0. The eigenvalues and eigenvectors of H_0 will be denoted by E_n and $|\varphi_n\rangle$:

$$H_0|\varphi_n\rangle = E_n|\varphi_n\rangle \tag{A-1}$$

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

For the sake of simplicity, we shall consider the spectrum of H_0 to be discrete and non-degenerate; the formulas obtained can easily be generalized (see, for example, § C-3). We assume that H_0 is not explicitly time-dependent, so that its eigenstates are stationary states.

At $t = 0$, a perturbation is applied to the system. Its Hamiltonian then becomes:

$$H(t) = H_0 + W(t) \tag{A-2}$$

with:

$$W(t) = \lambda \hat{W}(t) \tag{A-3}$$

where λ is a real dimensionless parameter much smaller than 1 and $\hat{W}(t)$ is an observable (which can be explicitly time-dependent) of the same order of magnitude as H_0, and zero for $t < 0$.

The system is assumed to be initially in the stationary state $|\varphi_i\rangle$, an eigenstate of H_0 of eigenvalue E_i. Starting at $t = 0$ when the perturbation is applied, the system evolves: the state $|\varphi_i\rangle$ is no longer, in general, an eigenstate of the perturbed Hamiltonian. We propose, in this chapter, to calculate the probability $\mathscr{P}_{if}(t)$ of finding the system in another eigenstate $|\varphi_f\rangle$ of H_0 at time t. In other words, we want to study the transitions that can be induced by the perturbation $W(t)$ between the stationary states of the unperturbed system.

The treatment is very simple. Between the times 0 and t, the system evolves in accordance with the Schrödinger equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = \left[H_0 + \lambda \hat{W}(t) \right] |\psi(t)\rangle \tag{A-4}$$

The solution $|\psi(t)\rangle$ of this first-order differential equation corresponding to the initial condition:

$$|\psi(t = 0)\rangle = |\varphi_i\rangle \tag{A-5}$$

is unique. The desired probability $\mathscr{P}_{if}(t)$ can be written:

$$\mathscr{P}_{if}(t) = |\langle \varphi_f | \psi(t) \rangle|^2 \tag{A-6}$$

The whole problem therefore consists of finding the solution $|\psi(t)\rangle$ of (A-4) that corresponds to the initial condition (A-5). However, such a problem is not generally rigorously soluble. This is why we resort to approximation methods. We shall show in this chapter how, if λ is sufficiently small, the solution $|\psi(t)\rangle$ can be found in the form of a power series expansion in λ. Thus, we shall calculate $|\psi(t)\rangle$ explicitly to first order in λ, as well as the corresponding probability (§ B). The general formulas obtained will then be applied (§ C) to the study of an important special case, the one in which the perturbation is a sinusoidal function of time or a constant (the interaction of an atom with an electro-magnetic wave, which falls into this category, is treated in detail in Complement $\mathrm{A_{XIII}}$). This is an example of the *resonance* phenomenon. Two situations will be considered: the one in which the spectrum of H_0 is discrete, and the one in which the initial state $|\varphi_i\rangle$ is coupled to a continuum of final states. In the latter case, we shall prove an important formula known as "*Fermi's golden rule*". In § D we will

consider another important case in which the perturbation fluctuates randomly; it is then characterized by its time-dependent correlation function, and will be treated with a perturbative calculation that is valid for short times. We will then show in § E how to extend the valitidy of this calculation to long times, within a general approximation called "motional narrowing approximation".

Comment:

The situation treated in § C-3 of Chapter IV can be considered to be a special case of the general problem discussed in this chapter. Recall that, in Chapter IV, we discussed a two-level system (the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$), initially in the state $|\varphi_1\rangle$, subjected, beginning at time $t = 0$, to a constant perturbation W. The probability $\mathscr{P}_{12}(t)$ can then be calculated exactly, leading to *Rabi's formula*.

The problem we are taking up here is much more general. We shall consider a system with an arbitrary number of levels (sometimes, as in § C-3, with a continuum of states) and a perturbation $W(t)$ which is an arbitrary function of the time. This explains why, in general, we can obtain only an approximate solution.

B. Approximate solution of the Schrödinger equation

B-1. The Schrödinger equation in the $\{|\varphi_n\rangle\}$ representation

The probability $\mathscr{P}_{if}(t)$ explicitly involves the eigenstates $|\varphi_i\rangle$ and $|\varphi_f\rangle$ of H_0. It is therefore reasonable to choose the $\{|\varphi_n\rangle\}$ representation.

B-1-a. The system of differential equations for the components of the state vector

Let $c_n(t)$ be the components of the ket $|\psi(t)\rangle$ in the $\{|\varphi_n\rangle\}$ basis:

$$|\psi(t)\rangle = \sum_n c_n(t)|\varphi_n\rangle \tag{B-1}$$

with:

$$c_n(t) = \langle\varphi_n|\psi(t)\rangle \tag{B-2}$$

$\hat{W}_{nk}(t)$ denotes the matrix elements of the observable $\hat{W}(t)$ in the same basis:

$$\langle\varphi_n|\hat{W}(t)|\varphi_k\rangle = \hat{W}_{nk}(t) \tag{B-3}$$

Recall that H_0 is represented in the $\{|\varphi_n\rangle\}$ basis by a diagonal matrix:

$$\langle\varphi_n|H_0|\varphi_k\rangle = E_n\delta_{nk} \tag{B-4}$$

We shall project both sides of Schrödinger equation (A-4) onto $|\varphi_n\rangle$. To do so, we insert the closure relation:

$$\sum_k |\varphi_k\rangle\langle\varphi_k| = 1 \tag{B-5}$$

and use relations (B-2), (B-3) and (B-4). We obtain:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}c_n(t) = E_n c_n(t) + \sum_k \lambda\hat{W}_{nk}(t)c_k(t) \tag{B-6}$$

The set of equations (B-6), written for the various values of n, constitutes a system of coupled linear differential equations of first order in t, which enables us, in theory, to determine the components $c_n(t)$ of $|\psi(t)\rangle$. The coupling between these equations arises solely from the existence of the perturbation $\lambda\hat{W}(t)$, which, by its non-diagonal matrix elements $\lambda\hat{W}_{nk}(t)$, relates the evolution of $c_n(t)$ to that of all the other coefficients $c_k(t)$.

B-1-b. Changing functions

When $\lambda\hat{W}(t)$ is zero, equations (B-6) are no longer coupled, and their solution is very simple. It can be written:

$$c_n(t) = b_n \mathrm{e}^{-iE_n t/\hbar} \tag{B-7}$$

where b_n is a constant which depends on the initial conditions.

Now, if $\lambda\hat{W}(t)$ is not zero, while remaining much smaller than H_0 because of the condition $\lambda \ll 1$, we expect the solution $c_n(t)$ of equations (B-6) to be very close to solution (B-7). In other words, if we perform the change of functions:

$$c_n(t) = b_n(t)\mathrm{e}^{-iE_n t/\hbar} \tag{B-8}$$

we can predict that the $b_n(t)$ will be slowly varying functions of time.

We substitute (B-8) into equation (B-6); we obtain:

$$\begin{aligned} i\hbar\,\mathrm{e}^{-iE_n t/\hbar}\frac{\mathrm{d}}{\mathrm{d}t}\,b_n(t) + E_n\,b_n(t)\;\mathrm{e}^{-iE_n t/\hbar} \\ = E_n\,b_n(t)\;\mathrm{e}^{-iE_n t/\hbar} + \sum_k \lambda\hat{W}_{nk}(t)\;b_k(t)\;\mathrm{e}^{-iE_k t/\hbar} \end{aligned} \tag{B-9}$$

We now multiply both sides of this relation by $\mathrm{e}^{+iE_n t/\hbar}$, and introduce the Bohr angular frequency:

$$\omega_{nk} = \frac{E_n - E_k}{\hbar} \tag{B-10}$$

related to the pair of states E_n and E_k. We obtain:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,b_n(t) = \lambda\sum_k \mathrm{e}^{i\omega_{nk}t}\;\hat{W}_{nk}(t)\,b_k(t) \tag{B-11}$$

B-2. Perturbation equations

The system of equations (B-11) is rigorously equivalent to Schrödinger equation (A-4). In general, we do not know how to find its exact solution. This is why we shall use the fact that λ is much smaller than 1 to try to determine this solution in the form

of a power series expansion in λ (which we can hope to be rapidly convergent if λ is sufficiently small):

$$b_n(t) = b_n^{(0)}(t) + \lambda b_n^{(1)}(t) + \lambda^2 b_n^{(2)}(t) + ... \tag{B-12}$$

If we substitute this expansion into (B-11), and if we set equal the coefficients of λ^r on both sides of the equation, we find:

(i) for $r = 0$:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} b_n^{(0)}(t) = 0 \tag{B-13}$$

since the right-hand side of (B-11) has a common factor λ. Relation (B-13) expresses the fact that $b_n^{(0)}$ does not depend on t. Thus, if λ is zero, $b_n(t)$ reduces to a constant [*cf.* (B-7)].

(ii) for $r \neq 0$:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} b_n^{(r)}(t) = \sum_k \mathrm{e}^{i\omega_{nk}t}\, \hat{W}_{nk}(t)\, b_k^{(r-1)}(t) \tag{B-14}$$

We see that, with the zeroth-order solution determined by (B-13) and the initial conditions, recurrence relation (B-14) enables us to obtain the first-order solution ($r = 1$). It then furnishes the second-order solution ($r = 2$) in terms of the first-order one and, by recurrence, the solution to any order r in terms of the one to order $r - 1$.

B-3. Solution to first order in λ

B-3-a. The state of the system at time t

For $t < 0$, the system is assumed to be in the state $|\varphi_i\rangle$. Of all the coefficients $b_n(t)$, only $b_i(t)$ is different from zero (and, furthermore, independent of t since $\lambda\hat{W}$ is then zero). At time $t = 0$, $\lambda\hat{W}(t)$ may become discontinuous in passing from a zero value to the value $\lambda\hat{W}(0)$. However, since $\lambda\hat{W}(t)$ remains finite, the solution of the Schrödinger equation is continuous at $t = 0$. It follows that:

$$b_n(t = 0) = \delta_{ni} \tag{B-15}$$

and this relation is valid for all λ. Consequently, the coefficients of expansion (B-12) must satisfy:

$$b_n^{(0)}(t = 0) = \delta_{ni} \tag{B-16}$$

$$b_n^{(r)}(t = 0) = 0 \qquad \text{if} \quad r \geq 1 \tag{B-17}$$

Equation (B-13) then immediately yields, for all positive t:

$$b_n^{(0)}(t) = \delta_{ni} \tag{B-18}$$

which completely determines the zeroth-order solution.

This result now permits us to write (B-14), for $r = 1$, in the form:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}b_n^{(1)}(t) = \sum_k \mathrm{e}^{i\omega_{nk}t}\hat{W}_{nk}(t)\delta_{ki}$$
$$= \mathrm{e}^{i\omega_{ni}t}\hat{W}_{ni}(t) \tag{B-19}$$

an equation which can be integrated without difficulty. Taking into account initial condition (B-17), we find:

$$b_n^{(1)}(t) = \frac{1}{i\hbar}\int_0^t \mathrm{e}^{i\omega_{ni}t'}\hat{W}_{ni}(t')\,\mathrm{d}t' \tag{B-20}$$

If we now substitute (B-18) and (B-20) into (B-8) and then into (B-1), we obtain the state $|\psi(t)\rangle$ of the system at time t, calculated to first order in λ.

B-3-b. The transition probability $\mathscr{P}_{if}(t)$

According to (A-6) and definition (B-2) of $c_f(t)$, the transition probability $\mathscr{P}_{if}(t)$ is equal to $|c_f(t)|^2$, that is, since $b_f(t)$ and $c_f(t)$ have the same modulus [*cf.* (B-8)]:

$$\mathscr{P}_{if}(t) = |b_f(t)|^2 \tag{B-21}$$

where:

$$b_f(t) = b_f^{(0)}(t) + \lambda b_f^{(1)}(t) + ... \tag{B-22}$$

can be calculated from the formulas established in the preceding section.

From now on, we shall assume that the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are different. We shall therefore be concerned only with the transitions induced by $\lambda\hat{W}(t)$ between two distinct stationary states of H_0. We then have $b_f^{(0)}(t) = 0$, and, consequently:

$$\mathscr{P}_{if}(t) = \lambda^2|b_f^{(1)}(t)|^2 \tag{B-23}$$

Using (B-20) and replacing $\lambda\hat{W}(t)$ by $W(t)$ [*cf.* (A-3)], we finally obtain:

$$\boxed{\mathscr{P}_{if}(t) = \frac{1}{\hbar^2}\left|\int_0^t \mathrm{e}^{i\omega_{fi}t'}W_{fi}(t')\,\mathrm{d}t'\right|^2} \tag{B-24}$$

Consider the function $\tilde{W}_{fi}(t')$, which is zero for $t' < 0$ and $t' > t$, and equal to $W_{fi}(t')$ for $0 \leq t' \leq t$ (*cf.* Fig. 1). $\tilde{W}_{fi}(t')$ is the matrix element of the perturbation "seen" by the system between the time $t = 0$ and the measurement time t, when we try to determine if the system is in the state $|\varphi_f\rangle$. Result (B-24) shows that $\mathscr{P}_{if}(t)$ is proportional to the square of the modulus of the Fourier transform of the perturbation actually "seen", $\tilde{W}_{fi}(t')$. This Fourier transform is evaluated at an angular frequency equal to the Bohr angular frequency associated with the transition under consideration.

Note also that the transition probability $\mathscr{P}_{if}(t)$ is zero to first order if the matrix element $W_{fi}(t)$ is zero for all t.

Comment:

We have not discussed the validity conditions of the approximation to first order in λ. Comparison of (B-11) with (B-19) shows that this approximation simply amounts to replacing, on the right-hand side of (B-11), the coefficients $b_k(t)$ by their values $b_k(0)$ at time $t = 0$. It is therefore clear that, so long as t remains small enough for $b_k(0)$ not to differ very much from $b_k(t)$, the approximation remains valid. On the other hand, when t becomes large, there is no reason why the corrections of order 2, 3, etc. in λ should be negligible.

C. An important special case: a sinusoidal or constant perturbation

C-1. Application of the general equations

Now assume that $W(t)$ has one of the two simple forms:

$$\hat{W}(t) = \hat{W} \sin \omega t \tag{C-1a}$$

$$\hat{W}(t) = \hat{W} \cos \omega t \tag{C-1b}$$

where $\hat{W}$ is a time-independent observable and ω, a constant angular frequency. Such a situation is often encountered in physics. For example, in Complements A_{XIII} and B_{XIII}, we consider the perturbation of a physical system by an electromagnetic wave of angular frequency ω; $\mathscr{P}_{if}(t)$ then represents the probability, induced by the incident monochromatic radiation, of a transition between the initial state $|\varphi_i\rangle$ and the final state $|\varphi_f\rangle$.

With the particular form (C-1a) of $\hat{W}(t)$, the matrix elements $\hat{W}_{fi}(t)$ take on the form:

$$\hat{W}_{fi}(t) = \hat{W}_{fi} \sin \omega t = \frac{\hat{W}_{fi}}{2i}(\mathrm{e}^{i\omega t} - \mathrm{e}^{-i\omega t}) \tag{C-2}$$

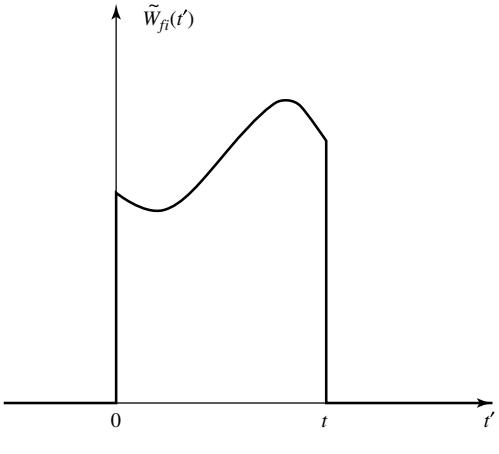


Figure 1: The variation of the function $\tilde{W}_{fi}(t')$ with respect to t'. This function coincides with $W_{fi}(t')$ in the interval $0 \leq t' \leq t$, and goes to zero outside this interval. It is the Fourier transform of $\tilde{W}_{fi}(t')$ that enters into the transition probability $\mathscr{P}_{if}(t)$ to lowest order.

where $\hat{W}_{fi}$ is a time-independent complex number. Let us now calculate the state vector of the system to first order in λ. If we substitute (C-2) into general formula (B-20), we obtain:

$$b_n^{(1)}(t) = -\frac{\hat{W}_{ni}}{2\hbar}\int_0^t \left[\mathrm{e}^{i(\omega_{ni}+\omega)t'} - \mathrm{e}^{i(\omega_{ni}-\omega)t'}\right] \mathrm{d}t' \tag{C-3}$$

The integral which appears on the right-hand side of this relation can easily be calculated and yields:

$$b_n^{(1)}(t) = \frac{\hat{W}_{ni}}{2i\hbar}\left[\frac{1-\mathrm{e}^{i(\omega_{ni}+\omega)t}}{\omega_{ni}+\omega} - \frac{1-\mathrm{e}^{i(\omega_{ni}-\omega)t}}{\omega_{ni}-\omega}\right] \tag{C-4}$$

Therefore, in the special case we are treating, general equation (B-24) becomes:

$$\mathscr{P}_{if}(t;\omega) = \lambda^2\left|b_f^{(1)}(t)\right|^2 = \frac{|W_{fi}|^2}{4\hbar^2}\left|\frac{1-\mathrm{e}^{i(\omega_{fi}+\omega)t}}{\omega_{fi}+\omega} - \frac{1-\mathrm{e}^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega}\right|^2 \tag{C-5a}$$

(we have added the variable ω in the probability $\mathscr{P}_{if}$, since the latter depends on the frequency of the perturbation).

If we choose the special form (C-1b) for $\hat{W}(t)$ instead of (C-1a), a calculation analogous to the preceding one yields:

$$\mathscr{P}_{if}(t;\omega) = \frac{|W_{fi}|^2}{4\hbar^2}\left|\frac{1-\mathrm{e}^{i(\omega_{fi}+\omega)t}}{\omega_{fi}+\omega} + \frac{1-\mathrm{e}^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega}\right|^2 \tag{C-5b}$$

The operator $\hat{W}\cos\omega t$ becomes time-independent if we choose $\omega = 0$. The transition probability $\mathscr{P}_{if}(t)$ induced by a constant perturbation W can therefore be obtained by replacing ω by 0 in (C-5b):

$$\begin{aligned}\mathscr{P}_{if}(t) &= \frac{|W_{fi}|^2}{\hbar^2\omega_{fi}^2}\left|1-\mathrm{e}^{i\omega_{fi}t}\right|^2 \\ &= \frac{|W_{fi}|^2}{\hbar^2}F(t,\omega_{fi})\end{aligned} \tag{C-6}$$

with:

$$F(t,\omega_{fi}) = \left[\frac{\sin(\omega_{fi}t/2)}{(\omega_{fi}/2)}\right]^2 \tag{C-7}$$

In order to study the physical content of equations (C-5b) and (C-6), we shall first consider the case in which $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are two discrete levels (§ C-2), and then that in which $|\varphi_f\rangle$ belongs to a continuum of final states (§ C-3). In the first case, $\mathscr{P}_{if}(t;\omega)$ [or $\mathscr{P}_{if}(t)$] really represents a transition probability which can be measured, while, in the second case, we are actually dealing with a probability density (the truly measurable quantities then involve a summation over a set of final states). From a physical point of view, there is a distinct difference between these two cases. We shall see in Complements $\mathrm{C_{XIII}}$ and $\mathrm{D_{XIII}}$ that, over a sufficiently long time interval, the system

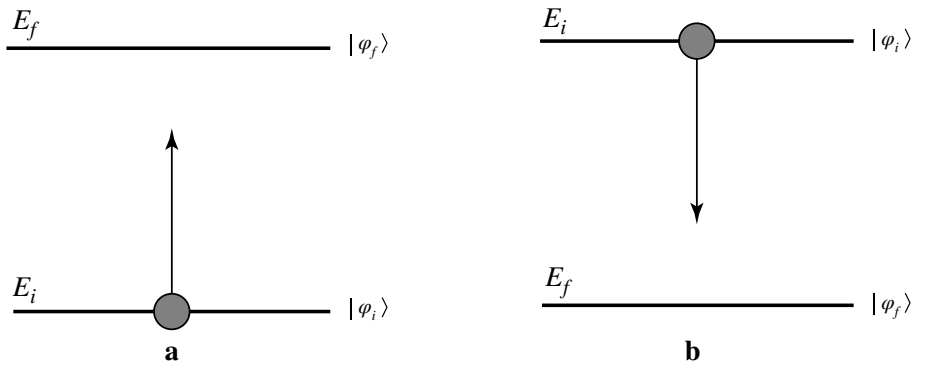


Figure 2: The relative disposition of the energies E_i and E_f associated with the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$. If $E_i < E_f$ (fig. a), the $|\varphi_i\rangle \longrightarrow |\varphi_f\rangle$ transition occurs through absorption of an energy quantum $\hbar\omega$. If, on the other hand, $E_i > E_f$ (fig. b), the $|\varphi_i\rangle \longrightarrow |\varphi_f\rangle$ transition occurs through induced emission of an energy quantum $\hbar\omega$.

oscillates between the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ in the first case, while it leaves the state $|\varphi_i\rangle$ irreversibly in the second case.

In § C-2, in order to concentrate on the resonance phenomenon, we shall choose a sinusoidal perturbation, but the results obtained can easily be transposed to the case of a constant perturbation. On the other hand, we shall use this latter case for the discussion of § C-3.

C-2. Sinusoidal perturbation coupling two discrete states: the resonance phenomenon

C-2-a. Resonant nature of the transition probability

When the time t is fixed, the transition probability $\mathscr{P}_{if}(t;\omega)$ is a function only of the variable ω. We shall see that this function has a maximum for:

$$\omega \simeq \omega_{fi} \tag{C-8a}$$

or:

$$\omega \simeq -\omega_{fi} \tag{C-8b}$$

A resonance phenomenon therefore occurs when the angular frequency of the perturbation coincides with the Bohr angular frequency associated with the pair of states $|\varphi_i\rangle$ and $|\varphi_f\rangle$. If we agree to choose $\omega \geq 0$, relations (C-8) give the resonance conditions corresponding respectively to the cases $\omega_{fi} > 0$ and $\omega_{fi} < 0$. In the first case (*cf.* Fig. 2-a), the system goes from the lower energy level E_i to the higher level E_f by the resonant absorption of an energy quantum $\hbar\omega$. In the second case (*cf.* Fig. 2-b), the resonant perturbation stimulates the passage of the system from the higher level E_i to the lower level E_f (accompanied by the induced emission of an energy quantum $\hbar\omega$). Throughout this section, we shall assume that ω_{fi} is positive (the situation of Figure 2-a). The case in which ω_{fi} is negative could be treated analogously.

To reveal the resonant nature of the transition probability, we note that both expressions (C-5a) and (C-5b) for $\mathscr{P}_{if}(t;\omega)$ involve the square of the modulus of a sum of two complex terms. The first of these terms is proportional to:

$$A_+ = \frac{1-\mathrm{e}^{i(\omega_{fi}+\omega)t}}{\omega_{fi}+\omega} = -i\mathrm{e}^{i(\omega_{fi}+\omega)t/2}\frac{\sin\left[(\omega_{fi}+\omega)t/2\right]}{(\omega_{fi}+\omega)/2} \tag{C-9a}$$

and the second one, to:

$$A_- = \frac{1-\mathrm{e}^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega} = -i\mathrm{e}^{i(\omega_{fi}-\omega)t/2}\frac{\sin\left[(\omega_{fi}-\omega)t/2\right]}{(\omega_{fi}-\omega)/2} \tag{C-9b}$$

The denominator of the A_- term goes to zero for $\omega = \omega_{fi}$, and that of the A_+ term, for $\omega = -\omega_{fi}$. Consequently, for ω close to ω_{fi}, we expect only the A_- term to be important; this is why it is called the "resonant term", while the A_+ term is called the "anti-resonant term" (A_+ would become resonant if, for negative ω_{fi}, ω were close to $-\omega_{fi}$).

Let us then consider the case in which:

$$|\omega - \omega_{fi}| \ll |\omega_{fi}| \tag{C-10}$$

neglecting the anti-resonant term A_+ (the validity of this approximation will be discussed in § C-2-c below). Taking (C-9b) into account, we then obtain:

$$\mathscr{P}_{if}(t;\omega) = \frac{|W_{fi}|^2}{4\hbar^2}F(t,\omega-\omega_{fi}) \tag{C-11}$$

with:

$$F(t,\omega-\omega_{fi}) = \left\{\frac{\sin\left[(\omega_{fi}-\omega)t/2\right]}{(\omega_{fi}-\omega)/2}\right\}^2 \tag{C-12}$$

Figure 3 represents the variation of $\mathscr{P}_{if}(t;\omega)$ with respect to ω, for a given time t. It clearly shows the resonant nature of the transition probability. This probability presents a maximum for $\omega = \omega_{fi}$, when it is equal to $|W_{fi}|^2t^2/4\hbar^2$. As we move away from ω_{fi}, it decreases, going to zero for $|\omega-\omega_{fi}| = 2\pi/t$. When $|\omega-\omega_{fi}|$ continues to increase, it oscillates between the value $|W_{fi}|^2/\hbar^2(\omega-\omega_{fi})^2$ and zero ("diffraction pattern").

C-2-b. The resonance width and the time-energy uncertainty relation

The resonance width $\Delta\omega$ can be approximately defined as the distance between the first two zeros of $\mathscr{P}_{if}(t;\omega)$ on each side of $\omega = \omega_{fi}$. It is inside this interval that the transition probability takes on its largest values [the first secondary maximum $\mathscr{P}_{if}$, attained when $(\omega-\omega_{fi})t/2 = 3\pi/2$, is equal to $|W_{fi}|^2/t^2/9\pi^2\hbar^2$, that is, less than 5% of the transition probability at resonance]. We then have:

$$\Delta\omega \simeq \frac{4\pi}{t} \tag{C-13}$$

The larger the time t, the smaller this width.

Result (C-13) presents a certain analogy with the time-energy uncertainty relation (*cf.* Chap. III, § D-2-e). Assume that we want to measure the energy difference $E_f - E_i =$

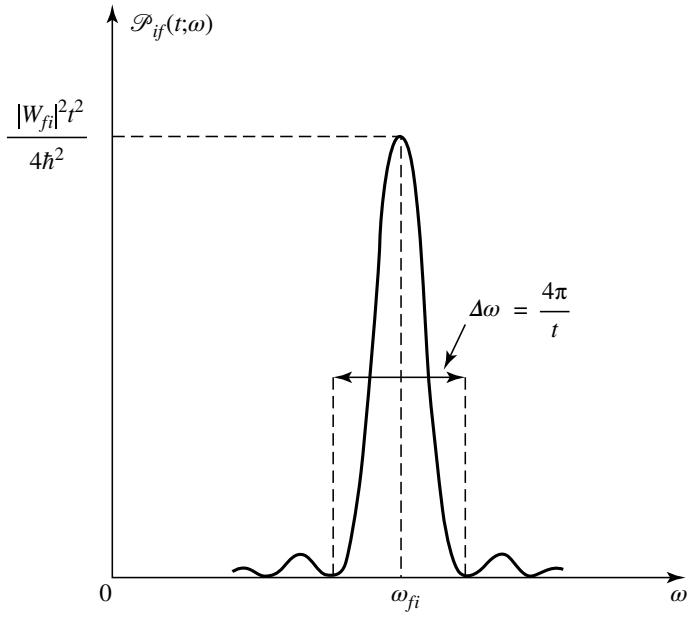


Figure 3: Variation, with respect to ω, of the first-order transition probability $\mathscr{P}_{if}(t;\omega)$ associated with a sinusoidal perturbation of angular frequency ω; t is fixed. When $\omega \simeq \omega_{fi}$, a resonance appears whose intensity is proportional to t^2 and whose width is inversely proportional to t.

$\hbar\omega_{fi}$ by applying a sinusoidal perturbation of angular frequency ω to the system and varying ω so as to detect the resonance. If the perturbation acts during a time t, the uncertainty ΔE on the value $E_f - E_i$ will be, according to (C-13), of the order of:

$$\Delta E = \hbar\Delta\omega \simeq \frac{\hbar}{t} \tag{C-14}$$

Therefore, the product $t\Delta E$ cannot be smaller than $\hbar$. This recalls the time-energy uncertainty relation, although t here is not a time interval characteristic of the free evolution of the system, but is externally imposed.

C-2-c. Validity of the perturbation treatment

Now let us examine the limits of validity of the calculations leading to result (C-11). We shall first discuss the resonant approximation, which consists of neglecting the anti-resonant term A_+, and then the first-order approximation in the perturbation expansion of the state vector.

α. Discussion of the resonant approximation

Using the hypothesis $\omega \simeq \omega_{fi}$, we have neglected A_+ relative to A_-. We shall therefore compare the moduli of A_+ and A_-.

The shape of the function $|A_-(\omega)|^2$ is shown in Figure 3. Since $|A_+(\omega)|^2 = |A_-(-\omega)|^2$, $|A_+(\omega)|^2$ can be obtained by plotting the curve which is symmetric with respect to the preceding one relative to the vertical axis $\omega = 0$. If these two curves, of width $\Delta\omega$, are centered at points whose separation is much larger than $\Delta\omega$, it is clear that, in the neighborhood of $\omega = \omega_{fi}$, the modulus of A_+ is negligible compared to that of A_-. The resonant approximation is therefore justified on the condition[1] that:

$$2|\omega_{fi}| \gg \Delta\omega \tag{C-15}$$

that is, using (C-13):

$$t \gg \frac{1}{|\omega_{fi}|} \simeq \frac{1}{\omega} \tag{C-16}$$

Result (C-11) is therefore valid only if the sinusoidal perturbation acts during a time t which is large compared to $1/\omega$. The physical meaning of such a condition is clear: during the interval $[0, t]$, the perturbation must perform numerous oscillations to appear to the system as a sinusoidal perturbation. If, on the other hand, t were small compared to $1/\omega$, the perturbation would not have enough time to oscillate and would be equivalent to a perturbation varying linearly in time [in the case (C-1a)] or constant [in the case (C-1b)].

Comment:

For a constant perturbation, condition (C-16) can never be satisfied, since ω is zero. However, it is not difficult to adapt the calculations of § C-2-b above to this case. We have already obtained [in (C-6)] the transition probability $\mathscr{P}_{if}(t)$ for a constant perturbation by directly setting $\omega = 0$ in (C-5b). Note that the two terms A_+ and A_- are then equal, which shows that if (C-16) is not satisfied, the anti-resonant term is not negligible.

The variation of the probability $\mathscr{P}_{if}(t)$ with respect to the energy difference $\hbar\omega_{fi}$ (with the time t fixed) is shown in Figure 4. This probability is maximal when $\omega_{fi} = 0$, which corresponds to what we found in § C-2-b above: if its angular frequency is zero, the perturbation is resonant when $\omega_{fi} = 0$ (degenerate levels). More generally, the considerations of § C-2-b concerning the features of the resonance can be transposed to this case.

β. Limits of the first-order calculation

We have already noted (*cf.* comment at the end of § B-3-b) that the first-order approximation can cease to be valid when the time t becomes too large. This can indeed be seen from expression (C-11), which, at resonance, can be written:

$$\mathscr{P}_{if}(t; \omega = \omega_{fi}) = \frac{|W_{fi}|^2}{4\hbar^2} t^2 \tag{C-17}$$

[1] Note that if condition (C-15) is not satisfied, the resonant and anti-resonant terms interfere: it is not correct to simply add $|A_+|^2$ and $|A_-|^2$.

This function becomes infinite when $t \longrightarrow \infty$, which is absurd, since a probability can never be greater than 1.

In practice, for the first-order approximation to be valid at resonance, the probability in (C-17) must be much smaller than 1, that is[2]:

$$t \ll \frac{\hbar}{|W_{fi}|} \tag{C-18}$$

To show precisely why this inequality is related to the validity of the first-order approximation, it would be necessary to calculate the higher-order corrections from (B-14) and to examine under what conditions they are negligible. We would then see that, although inequality (C-18) is necessary, it is not rigorously sufficient. For example, in the terms of second or higher order,

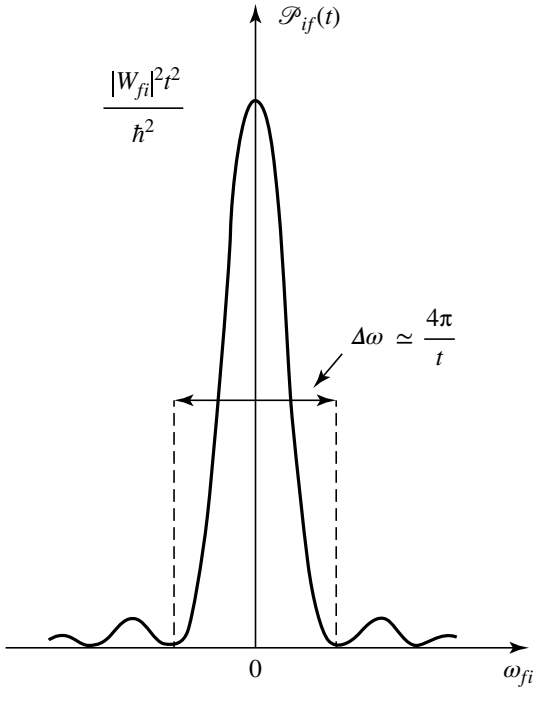


Figure 4: Variation of the transition probability $\mathscr{P}_{if}(t)$ associated with a constant perturbation with respect to $\omega_{fi} = (E_f - E_i)/\hbar$, for fixed t. A resonance appears, centered about $\omega_{fi} = 0$ (conservation of energy), with the same width as the resonance of Figure 3, but an intensity four times greater (because of the constructive interference of the resonant and anti-resonant terms, which, for a constant perturbation, are equal).

[2]For this theory to be meaningful, it is obviously necessary for conditions (C-16) and (C-18) to be compatible. That is, we must have:

$$\frac{1}{|\omega_{fi}|} \ll \frac{\hbar}{|W_{fi}|}$$

This inequality means that the energy difference $|E_f - E_i| = \hbar|\omega_{fi}|$ is much larger than the matrix element of $W(t)$ between $|\varphi_i\rangle$ and $|\varphi_f\rangle$.

there appear matrix elements $\hat{W}_{kn}$ of $\hat{W}$ other than $\hat{W}_{fi}$, on which certain conditions must be imposed for the corresponding corrections to be small.

Note that the problem of calculating the transition probability when t does not satisfy (C-18) is taken up in Complement C_{XIII}, in which an approximation of a different type is used (the secular approximation).

C-3. Coupling with the states of the continuous spectrum

If the energy E_f belongs to a continuous part of the spectrum of H_0, that is, if the final states are labeled by continuous indices, we cannot measure the probability of finding the system in a *well-defined* state $|\varphi_f\rangle$ at time t. The postulates of Chapter III indicate that in this case the quantity $|\langle\varphi_f|\psi(t)\rangle|^2$ which we determined above (approximately) is a probability density. The physical predictions for a given measurement then involve an integration of this probability density over a certain group of final states (which depends on the measurement to be made). We shall consider what happens to the results of the preceding sections in this case.

C-3-a. Integration over a continuum of final states: density of states

α. Example

To understand how this integration is performed over the final states, we shall first consider a concrete example.

We shall discuss the problem of the scattering of a spinless particle of mass m by a potential $W(\mathbf{r})$ (*cf.* Chap. VIII). The state $|\psi(t)\rangle$ of the particle at time t can be expanded on the states $|\mathbf{p}\rangle$ of well-defined momenta $\mathbf{p}$ and energies:

$$E = \frac{\mathbf{p}^2}{2m} \tag{C-19}$$

The corresponding wave functions are the plane waves:

$$\langle\mathbf{r}|\mathbf{p}\rangle = \left(\frac{1}{2\pi\hbar}\right)^{3/2} \mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \tag{C-20}$$

The probability density associated with a measurement of the momentum is $|\langle\mathbf{p}|\psi(t)\rangle|^2$ [$|\psi(t)\rangle$ is assumed to be normalized].

The detector used in the experiment (see, for example, Figure 2 of Chapter VIII) gives a signal when the particle is scattered with the momentum $\mathbf{p}_f$. Of course, this detector always has a finite angular aperture, and its energy selectivity is not perfect: it emits a signal whenever the momentum $\mathbf{p}$ of the particle points within a solid angle $\delta\Omega_f$ about $\mathbf{p}_f$ and its energy is included in the interval δE_f centered at $E_f = \mathbf{p}_f^2/2m$. If D_f denotes the domain of $\mathbf{p}$-space defined by these conditions, the probability of obtaining a signal from the detector is therefore:

$$\delta\mathscr{P}(\mathbf{p}_f, t) = \int_{\mathbf{p}\in D_f} \mathrm{d}^3p\ |\langle\mathbf{p}|\psi(t)\rangle|^2 \tag{C-21}$$

To use the results of the preceding sections, we shall have to perform a change of variables which results in an integral over the energies. This does not present any difficulties, since

we can write:

$$\mathrm{d}^3p = p^2 \mathrm{d}p \, \mathrm{d}\Omega \tag{C-22}$$

and replace the variable p by the energy E, to which it is related by (C-19). We thus obtain:

$$\mathrm{d}^3p = \rho(E) \mathrm{d}E \, \mathrm{d}\Omega \tag{C-23}$$

where the function $\rho(E)$, called the *density of final states*, can be written, according to (C-19), (C-22) and (C-23):

$$\rho(E) = p^2 \frac{\mathrm{d}p}{\mathrm{d}E} = p^2 \frac{m}{p} = m\sqrt{2mE} \tag{C-24}$$

(C-21) then becomes:

$$\delta\mathscr{P}(\mathbf{p}_f, t) = \int_{\Omega \in \delta\Omega_f; \quad E \in \delta E_f} \mathrm{d}\Omega \, \mathrm{d}E \, \rho(E) \, |\langle \mathbf{p}|\psi(t)\rangle|^2 \tag{C-25}$$

β. The general case

Assume that, in a particular problem, certain eigenstates of H_0 are labeled by a continuous set of indices, symbolized by α, such that the orthonormalization relation can be written:

$$\langle \alpha|\alpha'\rangle = \delta(\alpha - \alpha') \tag{C-26}$$

The system is described at time t by the normalized ket $|\psi(t)\rangle$. We want to calculate the probability $\delta\mathscr{P}(\alpha_f, t)$ of finding the system, in a measurement, in a given group of final states. We characterize this group of states by a domain D_f of values of the parameters α, centered at α_f, and we assume that their energies form a continuum. The postulates of quantum mechanics then yield:

$$\delta\mathscr{P}(\alpha_f, t) = \int_{\alpha \in D_f} \mathrm{d}\alpha \, |\langle \alpha|\psi(t)\rangle|^2 \tag{C-27}$$

As in the example of § α above, we shall change variables, and introduce the density of final states. Instead of characterizing these states by the parameters α, we shall use the energy E and a set of other parameters β (which are necessary when H_0 alone does not constitute a C.S.C.O.). We can then express $\mathrm{d}\alpha$ in terms of $\mathrm{d}E$ and $\mathrm{d}\beta$:

$$\mathrm{d}\alpha = \rho(\beta, E) \mathrm{d}\beta \, \mathrm{d}E \tag{C-28}$$

in which the density of final states[3] $\rho(\beta, E)$ appears. If we denote by $\delta\beta_f$ and δE_f the range of values of the parameters β and E defined by D_f, we obtain:

$$\delta\mathscr{P}(\alpha_f, t) = \int_{\beta \in \delta\beta_f; \; E \in \delta E_f} \mathrm{d}\beta \, \mathrm{d}E \, \rho(\beta, E) \, |\langle \beta, E|\psi(t)\rangle|^2 \tag{C-29}$$

where the notation $|\alpha\rangle$ has been replaced by $|\beta, E\rangle$ in order to point up the E- and β-dependence of the probability density $|\langle \alpha|\psi(t)\rangle|^2$.

[3] In the general case, the density of states ρ depends on both E and β. However, it often happens (*cf.* example of § α above) that ρ depends only on E.

C-3-b. Fermi's golden rule

In expression (C-29), $|\psi(t)\rangle$ is the normalized state vector of the system at time t. As in § A of this chapter, we shall consider a system which is initially in an eigenstate $|\varphi_i\rangle$ of H_0 [$|\varphi_i\rangle$ therefore belongs to the discrete spectrum of H_0, since the initial state of the system must, like$|\psi(t)\rangle$, be normalizable]. In (C-29), we shall replace the notation $\delta\mathscr{P}(\alpha_f, t)$ by $\delta\mathscr{P}(\varphi_i, \alpha_f, t)$ in order to remember that the system starts from the state $|\varphi_i\rangle$.

The calculations of § B and their application to the case of a sinusoidal or constant perturbation (§§ C-1 and C-2) remain valid when the final state of the system belongs to the continuous spectrum of H_0. If we assume W to be constant, we can therefore use (C-6) to find the probability density $|\langle\beta, E|\psi(t)\rangle|^2$ to first order in W. We then get:

$$|\langle\beta, E|\psi(t)\rangle|^2 = \frac{1}{\hbar^2}|\langle\beta, E|W|\varphi_i\rangle|^2 F\left(t, \frac{E-E_i}{\hbar}\right) \tag{C-30}$$

where E and E_i are the energies of the states $|\beta, E\rangle$ and $|\varphi_i\rangle$ respectively, and F is the function defined by (C-7). We get for $\delta\mathscr{P}(\varphi_i, \alpha_f, t)$, finally:

$$\delta\mathscr{P}(\varphi_i, \alpha_f, t) = \frac{1}{\hbar^2}\int_{\beta\in\delta\beta_f;\ E\in\delta E_f} \mathrm{d}\beta\,\mathrm{d}E\,\rho(\beta, E)\ |\langle\beta, E|W|\varphi_i\rangle|^2 F\left(t, \frac{E-E_i}{\hbar}\right) \tag{C-31}$$

The function $F\left(t, \frac{E-E_i}{\hbar}\right)$ varies rapidly about $E = E_i$ (*cf.* Fig. 4). If t is sufficiently large, this function can be approximated, to within a constant factor, by the function $\delta(E - E_i)$, since, according to (11) and (20) of Appendix II, we have:

$$\lim_{t\to\infty} F\left(t, \frac{E-E_i}{\hbar}\right) = \pi t\,\delta\left(\frac{E-E_i}{2\hbar}\right) = 2\pi\hbar t\,\delta(E-E_i) \tag{C-32}$$

On the other hand, the function $\rho(\beta, E)|\langle\beta, E|W|\varphi_i\rangle|^2$ generally varies much more slowly with E. We shall assume here that t is sufficiently large for the variation of this function over an energy interval of width $4\pi\hbar/t$ centered at $E = E_i$ to be negligible[4]. We can then in (C-31) replace $F\left(t, \frac{E-E_i}{\hbar}\right)$ by its limit (C-32). This enables us to perform the integral over E immediately. If, in addition, $\delta\beta_f$ is very small, integration over β is unnecessary, and we finally get:

$$\begin{cases} \text{– when the energy } E_i \text{ belongs to the domain } \delta E_f : \\ \delta\mathscr{P}(\varphi_i, \alpha_f, t) = \delta\beta_f \dfrac{2\pi}{\hbar} t\ |\langle\beta_f, E_f = E_i|W|\varphi_i\rangle|^2 \rho(\beta_f, E_f = E_i) & \text{(C-33a)} \\ \text{– when the energy } E_i \text{ does not belong to this domain:} \\ \delta\mathscr{P}(\varphi_i, \alpha_f, t) = 0 & \text{(C-33b)} \end{cases}$$

As we saw in the comment of § C-2-c-α, a constant perturbation can induce transitions only between states of equal energies. The system must have the same energy (to within $2\pi\hbar/t$) in the initial and final states. This is why, if the domain δE_f excludes the energy E_i, the transition probability is zero.

[4] $\rho(\beta, E)|\langle\beta, E|W|\varphi_i\rangle|^2$ must vary slowly enough to enable the finding of values of t that satisfy the stated condition but remain small enough for the perturbation treatment of W to be valid. Here, we also assume that $\delta E_f \gg 4\pi\hbar/t$.

The probability (C-33a) increases linearly with time. Consequently, *the transition probability per unit time*, $\delta\mathcal{W}(\varphi_i, \alpha_f)$, defined by:

$$\delta\mathcal{W}(\varphi_i, \alpha_f) = \frac{\mathrm{d}}{\mathrm{d}t}\delta\mathscr{P}(\varphi_i, \alpha_f, t) \tag{C-34}$$

is time-independent. We introduce the transition probability density per unit time and per unit interval of the variable β_f:

$$w(\varphi_i, \alpha_f) = \frac{\delta\mathcal{W}(\varphi_i, \alpha_f)}{\delta\beta_f} \tag{C-35}$$

It is equal to:

$$\boxed{w(\varphi_i, \alpha_f) = \frac{2\pi}{\hbar}\ |\langle\beta_f, E_f = E_i|W|\varphi_i\rangle|^2\ \rho(\beta_f, E_f = E_i)} \tag{C-36}$$

This important result is known as *Fermi's golden rule*.

Comments:

(*i*) Assume that W is a sinusoidal perturbation of the form (C-1a) or (C-1b), which couples a state ($|\varphi_i\rangle$ to a continuum of states $|\beta_f, E_f\rangle$ with energies E_f close to $E_i + \hbar\omega$. Starting with (C-11), we can carry out the same procedure as above, which yields:

$$w(\varphi_i, \alpha_f) = \frac{\pi}{2\hbar}|\langle\beta_f, E_f = E_i + \hbar\omega|W|\varphi_i\rangle|^2\rho(\beta_f, E_f = E_i + \hbar\omega) \tag{C-37}$$

(*ii*) Let us return to the problem of the scattering of a particle by a potential W whose matrix elements in the $\{|\mathbf{r}\rangle\}$ representation are given by:

$$\langle\mathbf{r}|W|\mathbf{r}'\rangle = W(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') \tag{C-38}$$

Now assume that the initial state of the system is a well-defined momentum state:

$$|\psi(t = 0)\rangle = |\mathbf{p}_i\rangle \tag{C-39}$$

and let us calculate the scattering probability of an incident particle of momentum $\mathbf{p}_i$ into the states of momentum $\mathbf{p}$ grouped about a given value $\mathbf{p}_f$ (with $|\mathbf{p}_f| = |\mathbf{p}_i|$). (C-36) gives the scattering probability $w(\mathbf{p}_i, \mathbf{p}_f)$ per unit time and per unit solid angle about $\mathbf{p} = \mathbf{p}_f$:

$$w(\mathbf{p}_i, \mathbf{p}_f) = \frac{2\pi}{\hbar}|\langle\mathbf{p}_f|W|\mathbf{p}_i\rangle|^2\rho(E_f = E_i) \tag{C-40}$$

Taking into account (C-20), (C-38) and expression (C-24) for $\rho(E)$, we then get:

$$w(\mathbf{p}_i, \mathbf{p}_f) = \frac{2\pi}{\hbar}m\sqrt{2mE_i}\left(\frac{1}{2\pi\hbar}\right)^6\left|\int \mathrm{d}^3r\, \mathrm{e}^{i(\mathbf{p}_i - \mathbf{p}_f)\cdot\mathbf{r}/\hbar}W(\mathbf{r})\right|^2 \tag{C-41}$$

On the right-hand side of this relation, we recognize the Fourier transform of the potential $W(\mathbf{r})$, evaluated for the value of $\mathbf{p}$ equal to $\mathbf{p}_i - \mathbf{p}_f$.

Note that the initial state $|\mathbf{p}_i\rangle$ chosen here is not normalizable, and it cannot represent the physical state of a particle. However, although the norm of $|\mathbf{p}_i\rangle$ is infinite, the right-hand side of (C-41) maintains a finite value. Intuitively, we can therefore expect to obtain a correct physical result from this relation. If we divide the probability obtained by the probability current:

$$J_i = \left(\frac{1}{2\pi\hbar}\right)^3 \frac{\hbar k_i}{m} = \left(\frac{1}{2\pi\hbar}\right)^3 \sqrt{\frac{2E_i}{m}} \tag{C-42}$$

associated, according to (C-20), with the state $|\mathbf{p}_i\rangle$, we obtain:

$$\frac{w(\mathbf{p}_i, \mathbf{p}_f)}{J_i} = \frac{m^2}{4\pi^2\hbar^4} \left| \int \mathrm{d}^3 r \, \mathrm{e}^{i(\mathbf{p}_i - \mathbf{p}_f)\cdot\mathbf{r}/\hbar} W(\mathbf{r}) \right|^2 \tag{C-43}$$

which is the expression for the scattering cross section in the Born approximation (§ B-4 of Chap. VIII).

Although it is not rigorous, the preceding treatment enables us to show that the scattering cross sections of the Born approximation can also be obtained by a time-dependent approach, using Fermi's golden rule.

D. Random perturbation

Another interesting case occurs when the perturbation applied to the system fluctuates in a random fashion. Consider for example an atom (a) having a spin magnetic moment, and moving in a gas of particles (b) which also have magnetic moments. As atom (a) undergoes a series of random collisions with particles (b), it is subjected to a magnetic field that varies randomly from one collision to another. The resulting interactions can change the orientation of the atom's magnetic moment. This type of situation is treated here (§ D). We shall go back to the calculation of § B assuming that the matrix element $\hat{W}_{fi}(t)$ of the perturbation is a random function of time. Our aim is to study the transition probability $\mathscr{P}_{if}(t)$ for going from the state $|\varphi_i\rangle$ to the state $|\varphi_f\rangle$ after a time t, and determine how it differs from the result found in the previous section.

D-1. Statistical properties of the perturbation

Here we consider the evolution of a single quantum system, atom (a) in the example described above, and study its evolution averaged over time. We thus need to consider the properties of statistical averages over time[5] of the perturbation $W(t)$. We note $\overline{W_{if}(t)}$ the average value of the matrix element $W_{if}(t)$, and assume it is equal to zero:

$$\overline{W_{if}(t)} = 0 \tag{D-1}$$

This means that $W_{if}(t)$ fluctuates between values that can be opposite. Since the matrix elements W_{if} and W_{fi} are two complex conjugate numbers, their product is necessarily positive or zero, hence having in general a non-zero average value:

$$\overline{W_{fi}(t)W_{if}(t)} \geq 0 \tag{D-2}$$

[5]The point of view of Complement $\mathrm{E_{XIII}}$ is more directly related to most experimental situations: we study an ensemble of N individual quantum systems described by their density operator. The two approaches are nevertheless equivalent since, in statistical mechanics, averaging over "a Gibbs ensemble" is equivalent to averaging a single system over a long time.

It will be useful in what follows to also consider the average value of such a product taken at different instants t and $t+\tau$, called the correlation function $C_{if}(\tau)$:

$$\overline{W_{fi}(t+\tau)W_{if}(t)} = C_{if}(\tau) \neq 0 \tag{D-3}$$

The τ dependence of $C_{if}(\tau)$ characterizes the time during which the perturbation keeps a "memory" of its value: $C_{if}(\tau)$ is non-zero as long as $W_{fi}(t+\tau)$ remains correlated with $W_{if}(t)$. The correlation function $C_{if}(\tau)$ goes to 0 when the time difference $|\tau|$ is longer than a characteristic time τ_c called the "correlation time" :

$$\overline{W_{fi}(t)W_{if}(t-\tau)} = C_{if}(\tau) \longrightarrow \quad 0 \quad \text{if} \quad |\tau| \gg \tau_c \tag{D-4}$$

We shall consider the case where τ_c is very short compared to all the other evolution times of the system. For instance, in the example mentioned above of an atom (a) diffusing in a gas of particles (b), the correlation time is of the order of the duration of a single collision, generally (much) shorter than 10^{-10} s.

We assume the random perturbation to be stationary, meaning that the correlation functions depend only on the difference τ between the two instants $t+\tau$ and t. Consequently, we can also write:

$$\overline{W_{fi}(t)W_{if}(t-\tau)} = C_{if}(\tau) \tag{D-5}$$

Using complex conjugation, relation (D-3) can be written:

$$\overline{W_{if}(t+\tau)W_{fi}(t)} = \overline{W_{fi}(t)W_{if}(t+\tau)} = C^*_{if}(\tau) \tag{D-6}$$

Comparing with (D-5) yields:

$$C^*_{if}(\tau) = C_{if}(-\tau) \tag{D-7}$$

Changing the sign of the variable τ transforms the function $C_{if}(\tau)$ into its complex conjugate; in particular, $C_{if}(0)$ is real.

In the following computations, il will be useful to introduce the Fourier transform $\tilde{C}_{if}(\omega)$ of the function $C_{if}(\tau)$:

$$\tilde{C}_{if}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}\tau \ \mathrm{e}^{-i\omega\tau} C_{if}(\tau) \tag{D-8}$$

leading to its inverse relation:

$$C_{if}(\tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}\omega \ \mathrm{e}^{+i\omega\tau} \tilde{C}_{if}(\omega) \tag{D-9}$$

Relation (D-7) implies that $\tilde{C}_{if}(\omega)$ is a real function.

D-2. Perturbative computation of the transition probability

As in (B-8), we perform the change of functions that transform $c_n(t)$ into $b_n(t)$, which eliminates the variations of the coefficients due to H_0 alone (this amounts to using the interaction picture, *cf.* exercise 15 of Complement $\mathrm{L_{III}}$). We assume that $b_i(0) = 1$

and $b_f(0) = 0$. We then look for the probability amplitude $b_f(t)$ for the system, starting at time $t = 0$ from the state $|\varphi_i\rangle$, to be found at time t in the state $|\varphi_f\rangle$. Equation (B-20) is now written:

$$b_f^{(1)}(t) = \frac{1}{i\hbar}\int_0^t \mathrm{e}^{i\omega_{fi}t'} W_{fi}(t')\mathrm{d}t' \tag{D-10}$$

The probability of finding the system in the state $|\varphi_f\rangle$ at time t is obtained by multiplying (D-10) by its complex conjugate. Since $\omega_{fi} = -\omega_{if}$ and $W_{fi}^* = W_{if}$, we get:

$$[b_f^{(1)}(t)]^* \times [b_f^{(1)}(t)] = \frac{1}{\hbar^2}\int_0^t \mathrm{e}^{i\omega_{if}t''} W_{if}(t'')\mathrm{d}t'' \int_0^t \mathrm{e}^{i\omega_{fi}t'} W_{fi}(t')\mathrm{d}t' \tag{D-11}$$

The transition probability $\mathscr{P}_{if}(t)$ is the average of (D-11) over the various values of the random perturbation. This leads to:

$$\mathscr{P}_{if}(t) = \overline{[b_f^{(1)}(t)]^* \times [b_f^{(1)}(t)]} = \frac{1}{\hbar^2}\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \mathrm{e}^{-i\omega_{if}(t'-t'')}\overline{W_{fi}(t')W_{if}(t'')} \tag{D-12}$$

Setting:

$$\tau = t' - t'' \tag{D-13}$$

and using (D-5) enable us to write (D-12) as:

$$\mathscr{P}_{if}(t) = \frac{1}{\hbar^2}\int_0^t \mathrm{d}t' \int_{t'-t}^{t'} \mathrm{d}\tau\, \mathrm{e}^{-i\omega_{if}\tau} C_{if}(\tau) \tag{D-14}$$

(the change of sign coming from $\mathrm{d}\tau = -\mathrm{d}t''$ is accounted for by interchanging the integration limits).

We assume in what follows that:

$$t \gg \tau_c \tag{D-15}$$

The integral over $\mathrm{d}t'$ is taken over a time interval from 0 to t, very large compared to τ_c. Its value will not be significantly modified if we shorten that interval at both end by a few τ_c. If n is of the order of a few units ($n = 2$ or $n = 3$ for instance), we can write:

$$\mathscr{P}_{if}(t) \simeq \frac{1}{\hbar^2}\int_{n\tau_c}^{t-n\tau_c} \mathrm{d}t' \int_{t'-t}^{t'} \mathrm{d}\tau\, \mathrm{e}^{-i\omega_{if}\tau} C_{if}(\tau) \tag{D-16}$$

In the integral over $\mathrm{d}\tau$, the upper limit is $t' > n\tau_c$; this upper limit may be extended to infinity since $C_{if}(\tau)$ goes to zero when $\tau \gg \tau_c$, and hence the additional contribution to the integral is zero. In the same way, the negative lower limit $t' - t$ can be replaced by $-\infty$, since the condition $t' - t \leq -n\tau_c$ ensures that the function to be integrated is zero in the additional integration domain. The integral over $\mathrm{d}\tau$ becomes independent of t', so that the integral over $\mathrm{d}t'$ is easily computed and leads to:

$$\int_{n\tau_c}^{t-n\tau_c} \mathrm{d}t' = (t - 2n\tau_c) \simeq t \tag{D-17}$$

We then get:

$$\mathscr{P}_{if}(t) \simeq \Gamma\, t \tag{D-18}$$

where the constant Γ is defined by:

$$\Gamma = \frac{1}{\hbar^2}\int_{-\infty}^{+\infty} C_{if}(\tau)\ \mathrm{e}^{-i\omega_{if}\tau}\mathrm{d}\tau = \frac{\sqrt{2\pi}}{\hbar^2}\tilde{C}_{if}(\omega_{if}) \tag{D-19}$$

This result involves the Fourier transform $\tilde{C}_{if}(\omega_{if})$ of the correlation function $C_{if}(\tau)$ defined by relation (D-8), taken at the (angular) frequency $\omega_{if} = -\omega_{fi}$ of the transition between the initial state $|\varphi_i\rangle$ and the final state $|\varphi_f\rangle$. As already mentioned, relation (D-7) shows that the constant Γ is real.

The transition probability from $|\varphi_i\rangle$ to $|\varphi_f\rangle$ after a time t is thus proportional to that time. This means that when the perturbation is random, one can define (at least for short times[6] where the perturbative treatment to lowest order is valid) a transition probability per unit time from $|\varphi_i\rangle$ to $|\varphi_f\rangle$. It is proportional to the Fourier transform of the correlation function of the perturbation, computed at the angular frequency ω_{fi}. This is a very different result from the one obtained in (C-11) and (C-12) for a sinusoidal perturbation. In that case, the transition probability increased as t^2 for short times, and then oscillated as a function of time.

D-3. Validity of the perturbation treatment

Result (D-18), obtained by a perturbative treatment, is valid as long as the transition probability remains small, that is if:

$$t \ll \frac{1}{\Gamma} \tag{D-20}$$

On the other hand, to establish (D-18) we assumed in (D-15) that t was much larger than τ_c. The two conditions (D-20) and (D-15) are compatible only if:

$$\tau_c \ll \frac{1}{\Gamma} \tag{D-21}$$

The calculation we just presented implies the existence of two very different time scales: the evolution time of the system, of the order of $1/\Gamma$, often called the "relaxation time"; the correlation time, τ_c, which is much shorter and characterizes the memory of the fluctuations of the random perturbation.

Equation (D-21) simply expresses the fact that, during this correlation time, the system barely evolves. Using for Γ relation (D-19), this inequality can be written solely with parameters concerning the perturbation. This inequality is often called the "motional narrowing condition", for reasons that will be explained in § 2-c-δ of Complement $\mathrm{E}_{\mathrm{XIII}}$.

[6] We shall see in the next section under which conditions this result remains valid for long times.

E. Long-time behavior for a two-level atom

Until now, we have limited ourselves in this chapter to perturbative calculations of the transition probability after a time t. We found that it increases as t^2 for a sinusoidal perturbation, but linearly as t for a random perturbation with a short memory. As a probability cannot become larger than one, such approximations are only valid for small values of t. In the last part of this chapter, we shall present treatments that permit us to study and compare the long-time behaviors of a system subjected to these two types of perturbations. For the sake of simplicity, we shall limit our study to the case of a two-level system.

E-1. Sinusoidal perturbation

We already studied in Complement $\mathrm{F_{IV}}$ the long-time behavior of a two-level system subjected to a sinusoidal perturbation. We computed the exact evolution of a spin 1/2 in the special case where the Hamiltonian H obeys relation (14) of $\mathrm{F_{IV}}$:

$$H(t) = \frac{\hbar}{2}\begin{pmatrix} \omega_0 & \omega_1 \mathrm{e}^{-i\omega t} \\ \omega_1 \mathrm{e}^{i\omega t} & -\omega_0 \end{pmatrix} \tag{E-1}$$

(this matrix is written using the basis $\{|+\rangle, |-\rangle\}$ of the eigenvectors of the S_z spin component). The diagonal of this matrix yields the matrix elements of the Hamiltonian H_0; this Hamiltonian comes from the coupling of the spin with a static magnetic field $\mathbf{B}_0$, parallel to the Oz axis. The perturbation Hamiltonian $W(t)$ corresponds to the non-diagonal parts of the matrix; it comes from the coupling of the spin with a radiofrequency field rotating around the Oz axis at the angular frequency ω. We showed in Complement $\mathrm{F_{IV}}$ that the quantum evolution of a spin 1/2 was identical to the classical evolution of a magnetic dipole with a proportional angular momentum. This led to a useful image for the evolution of a spin in a magnetic field, composed of a constant and a rotating field.

Now we saw in Complement $\mathrm{C_{IV}}$ that any two-level system is perfectly isomorphic to a spin 1/2. The states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are associated with the spin states $|+\rangle$ and $|-\rangle$, and the Hamiltonian H_0 leads to two non-perturbed energies $E_i = \hbar\omega_0/2$ and $E_f = -\hbar\omega_0/2$; this means that $\omega_0 = \omega_{if}$. We assume that the perturbation W that couples the two states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ is the analog of the action of a magnetic field $\mathbf{B}_1$ rotating in the plane xOy at the frequency ω; it is thus responsible for the non-diagonal matrix elements of (E-1), with:

$$\begin{aligned} W_{if}(t) &= \frac{\hbar\omega_1}{2}\, \mathrm{e}^{-i\omega t} \\ W_{fi}(t) &= \frac{\hbar\omega_1}{2}\, \mathrm{e}^{i\omega t} \end{aligned} \tag{E-2}$$

(the number ω_1 is supposed to be real; if this is not the case, a change of the relative phase of $|\varphi_i\rangle$ and $|\varphi_f\rangle$ allows this condition to be fulfilled). We can then directly transpose the results of Complement $\mathrm{F_{IV}}$, with no additional computations. Relation (27) of Complement $\mathrm{F_{IV}}$ shows that the transition probability is given by "Rabi's formula":

$$\mathscr{P}_{if}(t) = \frac{(\omega_1)^2}{(\omega_1)^2 + (\omega - \omega_{if})^2} \sin^2\left[\sqrt{(\omega_1)^2 + (\omega - \omega_{if})^2}\ \frac{t}{2}\right] \tag{E-3}$$

Since the quantum and classical evolution coincide in the present situation, they can be simply interpreted in terms of the classical precession of a magnetic moment around an "effective field". At resonance, the effective field is located in the xOy plane, for instance along the Ox axis. The spin, initially parallel to Oz, precesses around Ox, hence tracing large circles in the plane yOz. Relation (E-3) shows that the probability for the spin to reverse its initial orientation is written:

$$\mathscr{P}_{if}(t) = \sin^2\left(\frac{\omega_1 t}{2}\right) \tag{E-4}$$

This probability oscillates between 0 and 1 with a precession angular frequency $\omega_1 = 2W_{fi}/\hbar$, called "Rabi's frequency". This type of long lasting oscillations could not have been obtained[7] by a perturbation treatment.

For a non-resonant perturbation, the effective field has a component along the Oz axis. As it precesses, the magnetic moment now follows a cone; the larger the discrepancy between ω and the resonant frequency, the smaller the cone's aperture becomes (Figure 2 of Complement $\mathrm{F_{IV}}$). We must now use the complete relation (E-3) which, also, predicts a sinusoidal oscillation. It should be noted that, if $|\omega - \omega_{if}| \gg W_{if}/2\hbar$, we find again the result of equations (C-11) and (C-12) of Chapter XIII, which thus provide a good approximation in this case.

Comment:

The previous results assume that the perturbation can be reduced to a single rotating field. Replacing in (E-1) the exponentials $e^{\pm i\omega t}$ by the sinusoidal functions $\sin\omega t$ or $\cos\omega t$, it introduces two rotating fields with opposite frequencies $\pm\omega$; they both act simultaneously on the system, leading to a more complex situation. The results remain, however, valid as long as the perturbation is weak enough (meaning $\omega_1 \ll \omega_0$) and ω not too far from one of the two resonances ($\omega \simeq \omega_0$ or $\omega \simeq -\omega_0$).

E-2. Random perturbation

As for the spin 1/2 case considered above, we assume here that the perturbation W does not have diagonal elements:

$$W_{ii} = W_{ff} = 0 \tag{E-5}$$

(Complement $\mathrm{E_{XIII}}$ presents a more general calculation, where this hypothesis is no longer necessary).

In § E-1, we assumed that the system was initially in the state $|\varphi_i\rangle$, hence only $b_i(t=0)$ was different from zero. This rules out the possibility of any superposition of states at the initial moment. To remove this restriction, we now assume that the system is, at instant t, in any superposition of the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$:

$$|\Psi(t)\rangle = b_i(t)\mathrm{e}^{-iE_i t/\hbar}|\varphi_i\rangle + b_f(t)\mathrm{e}^{-iE_f t/\hbar}|\varphi_f\rangle \tag{E-6}$$

We then consider a later instant, $t+\Delta t$, and compute the evolution of the system between the times t and $t+\Delta t$, to second order in W.

[7] One must sum an infinite number of perturbative terms to reconstruct a sine squared.

E-2-a. State vector of the system at time $t + \Delta t$ to second order in W

To zeroth-order of the perturbation, relation (B-13) shows that neither $b_i(t + \Delta t)$ nor $b_f(t + \Delta t)$ depend on Δt:

$$b_i^{(0)}(t + \Delta t) = b_i(t) \qquad ; \qquad b_f^{(0)}(t + \Delta t) = b_f(t) \tag{E-7}$$

To first order, we use relation (B-14) with $r = 1$; hypothesis (E-5) means that b_i is only coupled to b_f, and vice versa. Integrating over time, we get:

$$\begin{aligned} b_i^{(1)}(t + \Delta t) &= \frac{1}{i\hbar}\, b_f(t) \int_t^{t+\Delta t} \mathrm{e}^{i\omega_{if}t'} W_{if}(t')\, \mathrm{d}t' \\ &= \frac{1}{i\hbar}\, b_f(t) \int_0^{\Delta t} \mathrm{e}^{i\omega_{if}(t+t'')} W_{if}(t + t'')\, \mathrm{d}t'' \end{aligned} \tag{E-8}$$

where we have set $t'' = t' - t$; we also get a similar relation where the indices i and f are interchanged:

$$b_f^{(1)}(t + \Delta t) = \frac{1}{i\hbar}\, b_i(t) \int_0^{\Delta t} \mathrm{e}^{i\omega_{fi}(t+t'')} W_{fi}(t + t'')\, \mathrm{d}t'' \tag{E-9}$$

The term (E-8) describes an atom that was at time t in the state $|\varphi_f\rangle$ and is found at time $t + t''$ in the state $|\varphi_i\rangle$; the term (E-9) describes the inverse process.

To second order, we again use relation (B-14), this time with $r = 2$; after integration, it leads to:

$$\begin{aligned} b_i^{(2)}(t + \Delta t) &= \frac{1}{i\hbar} \int_t^{t+\Delta t} \mathrm{e}^{i\omega_{if}t'} W_{if}(t')\, b_f^{(1)}(t')\, \mathrm{d}t' \\ &= \frac{1}{i\hbar} \int_0^{\Delta t} \mathrm{e}^{i\omega_{if}(t+t'')} W_{if}(t + t'')\, b_f^{(1)}(t + t'')\, \mathrm{d}t'' \end{aligned} \tag{E-10}$$

We now change the integration variable t'' to t', and insert relation (E-9) after replacing Δt by t'; this leads to:

$$\begin{aligned} b_i^{(2)}(t + \Delta t) &= -\frac{1}{\hbar^2}\, b_i(t) \int_0^{\Delta t} \mathrm{e}^{i\omega_{if}(t+t')} W_{if}(t + t')\, \mathrm{d}t' \int_0^{t'} \mathrm{e}^{i\omega_{fi}(t+t'')} W_{fi}(t + t'')\, \mathrm{d}t'' \\ &= -\frac{1}{\hbar^2}\, b_i(t) \int_0^{\Delta t} \mathrm{e}^{i\omega_{if}t'} W_{if}(t + t')\, \mathrm{d}t' \int_0^{t'} \mathrm{e}^{i\omega_{fi}t''} W_{fi}(t + t'')\, \mathrm{d}t'' \end{aligned} \tag{E-11}$$

This perturbative term describes an atom that was at time t in the state $|\varphi_i\rangle$, then made a transition to the state $|\varphi_f\rangle$ at time $t + t''$ (included between t and $t + t'$), then came back to the state $|\varphi_i\rangle$ at time $t + t'$ (included between t and $t + \Delta t$). Here also we can interchange the indices i and f to get the probability amplitude of the inverse process.

E-2-b. Average occupation probabilities to second order

For given values of the random variables W_{if} and W_{fi}, the probability of finding the system in the state $|\varphi_i\rangle$ is $|b_i(t + \Delta t)|^2$. To second order in W, this squared modulus contains the following terms:

– the squared modulus of $b_i^{(0)}(t+\Delta t)$, which is of zeroth order.

– a first-order term containing the product of $b_i^{(0)}(t+\Delta t)$ and the complex conjugate of $b_i^{(1)}(t+\Delta t)$, or the opposite. Due to condition (E-5), this term, linear in W, averages out to zero over all the possible values of W_{if} and W_{fi}. It will not be taken into account.

– the squared modulus of $b_i^{(1)}(t+\Delta t)$, which is of order 2.

– finally, twice the real part of the product of $b_i^{(0)}(t+\Delta t)$ and the complex conjugate of $b_i^{(2)}(t+\Delta t)$, which is also of order 2.

We thus get:

$$\begin{aligned} &\mid b_i(t+\Delta t) \mid^2 \simeq \\ &\quad \mid b_i^{(0)}(t+\Delta t) \mid^2 + \mid b_i^{(1)}(t+\Delta t) \mid^2 + 2\mathrm{Re}\left\{[b_i^{(0)}(t+\Delta t)] \times [b_i^{(2)}(t+\Delta t)]^*\right\} \end{aligned} \quad \text{(E-12)}$$

This expression is rewritten below in a slightly different form. The first and third term are regrouped in a first line, while the second term is rewritten in the second line. Note that the first term is rewritten using (E-7), while for the third term we use the complex conjugate of the second line of (E-11), obtained by replacing ω_{if} by ω_{fi} (and vice versa) as well as W_{if} by W_{fi} (and vice versa). This leads to:

$$\begin{aligned} &\mid b_i(t+\Delta t) \mid^2 \simeq \\ &\quad \mid b_i(t) \mid^2 \left(1 - \frac{2}{\hbar^2}\mathrm{Re}\int_0^{\Delta t} \mathrm{e}^{i\omega_{fi}t'} W_{fi}(t+t')\,\mathrm{d}t' \int_0^{t'} \mathrm{e}^{i\omega_{if}t''} W_{if}(t+t'')\,\mathrm{d}t''\right) \\ &\quad + \mid b_f(t) \mid^2 \frac{1}{\hbar^2}\int_0^{\Delta t} \mathrm{e}^{i\omega_{fi}t'} W_{fi}(t+t')\,\mathrm{d}t' \int_0^{\Delta t} \mathrm{e}^{i\omega_{if}t''} W_{if}(t+t'')\,\mathrm{d}t'' \end{aligned} \quad \text{(E-13)}$$

We now average this probability over the various values of the random variables W_{fi} and W_{if}. We get in (E-13) the product of two matrix elements of W and of the amplitude squared of $b_i(t)$ and $b_f(t)$. Rigorously speaking, these quantities are not mutually independent, since the system's state at time t is determined by the values of the perturbation at an earlier time. This correlation actually lasts over a time much shorter than Δt, of the order of the correlation time τ_c of the functions W_{fi} and W_{if} – *cf.* relation (D-4); therefore, a very short time τ_c after the instant t, W_{fi} and W_{if} are no longer correlated with the values of $|b_{i,f}(t)|^2$. It is then justified to compute separately the two averages:

$$\overline{\mid b_{i,f}(t) \mid^2 \times \text{Double integral}} \simeq \overline{\mid b_{i,f}(t) \mid^2} \times \overline{\text{Double integral}} \quad \text{(E-14)}$$

The averages $\overline{\mid b_i(t) \mid^2}$ and $\overline{\mid b_f(t) \mid^2}$ of the populations of the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are simply the diagonal elements of the density operator (Complement $\mathrm{E_{XIII}}$). Noting $\tilde{\rho}(t)$ this operator, we get:

$$\begin{aligned} \tilde{\rho}_{ii}(t) &= \overline{|b_i(t)|^2} \\ \tilde{\rho}_{ff}(t) &= \overline{|b_f(t)|^2} \end{aligned} \quad \text{(E-15)}$$

The variation rate of $\tilde{\rho}_{ii}(t)$ between the times t and $t+\Delta t$ is:

$$\frac{\Delta\tilde{\rho}_{ii}(t)}{\Delta t} = \frac{\overline{|b_i(t+\Delta t)|^2} - \overline{|b_i(t)|^2}}{\Delta t} \tag{E-16}$$

Relation (E-13) then yields:

$$\frac{\Delta\tilde{\rho}_{ii}(t)}{\Delta t} = -\tilde{\rho}_{ii}(t)\,[J_1 + J_1^*] + \tilde{\rho}_{ff}(t)\;J_2 \tag{E-17}$$

where J_1 and J_2 are the averages of double integrals:

$$\begin{aligned} J_1 &= \frac{1}{\hbar^2}\frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t' \int_0^{t'} \mathrm{d}t''\; \mathrm{e}^{-i\omega_{if}(t'-t'')}\overline{W_{fi}(t+t')W_{if}(t+t'')} \\ J_2 &= \frac{1}{\hbar^2}\frac{1}{\Delta t}\int_0^{\Delta t} \mathrm{d}t' \int_0^{\Delta t} \mathrm{d}t''\; \mathrm{e}^{-i\omega_{if}(t'-t'')}\overline{W_{fi}(t+t')W_{if}(t+t'')} \end{aligned} \tag{E-18}$$

The computation of the average values of these two double integrals is similar to the one performed in § D-2. It is carried out in detail below and leads to:

$$J_2 = \Gamma \tag{E-19}$$

$$J_1 = \frac{\Gamma}{2} + i\,\delta_{if} \tag{E-20}$$

In these relations, Γ and δ_{if} are expressed in terms of the Fourier transform $\tilde{C}_{if}(\omega)$ of $C_{if}(\tau)$, which was introduced in (D-8). The constant Γ was given in (D-19):

$$\Gamma = \frac{\sqrt{2\pi}}{\hbar^2}\tilde{C}_{if}(\omega_{if}) = \frac{1}{\hbar^2}\int_{-\infty}^{+\infty} C_{if}(\tau)\;\mathrm{e}^{-i\omega_{if}\tau}\mathrm{d}\tau \tag{E-21}$$

and δ_{if} is defined by:

$$\delta_{if} = \frac{1}{\sqrt{2\pi}\hbar^2}\;\mathcal{P}\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{1}{\omega-\omega_{if}}\;\tilde{C}_{if}(\omega) \tag{E-22}$$

where $\mathcal{P}$ means the principal part (Appendix II, §1-b).

Computation of J_1 ***and*** J_2:

The double integral appearing in J_2 has already been encountered in (D-12), while computing the value of $\mathscr{P}_{if}(t)$; we must simply replace t by Δt. Its value is given in (D-18), which becomes here $\Gamma\Delta t$. According to the definition (E-18) of J_2, we must then divide this result by Δt, which leads to relation (E-19).

The computation of J_1 is similar to that of J_2, except that the upper limit of the integral over $\mathrm{d}t''$ is Δt instead of t'. In the first line of (E-18), we can make the change of variables $t'-t''=\tau$ to transform the integrations according to:

$$\int_0^{\Delta t}\mathrm{d}t'\int_0^{t'}\mathrm{d}t'' \Rightarrow -\int_0^{\Delta t}\mathrm{d}t'\int_{t'}^{0}\mathrm{d}\tau = \int_0^{\Delta t}\mathrm{d}t'\int_0^{t'}\mathrm{d}\tau \tag{E-23}$$

As we did in (D-16), when $\Delta t \gg \tau_c$ we can replace the lower limit of the integral over $\mathrm{d}t'$ by a few τ_c, without any significant change of the result. The upper limit of the integral over $\mathrm{d}\tau$ is then longer than a few τ_c and can be extended to $+\infty$. The integral over $\mathrm{d}t'$ then reduces to a simple factor Δt, which cancels that same factor in the denominator of (E-18). This leads to:

$$J_1 \simeq \frac{1}{\hbar^2}\int_0^{+\infty} \mathrm{d}\tau\ \mathrm{e}^{-i\omega_{if}\tau}\overline{W_{fi}(t')W_{if}(t'-\tau)} \tag{E-24}$$

or else, using relation (D-9) to introduce the Fourier transform of $C_{if}(\tau)$:

$$J_1 \simeq \frac{1}{\sqrt{2\pi}\hbar^2}\int_{-\infty}^{+\infty} \mathrm{d}\omega\, \tilde{C}_{if}(\omega)\int_0^{+\infty} \mathrm{d}\tau\, \mathrm{e}^{i(\omega-\omega_{if})\tau} \tag{E-25}$$

The integral over τ leads to:

$$\int_0^{+\infty} \mathrm{d}\tau\, \mathrm{e}^{i(\omega-\omega_{if})\tau} \tag{E-26}$$

To make this integral convergent, we introduce an infinitesimal (positive) factor ε and write:

$$\int_0^{+\infty} \mathrm{d}\tau\, \mathrm{e}^{i(\omega-\omega_{if}+i\varepsilon)\tau} = -\frac{1}{i(\omega-\omega_{if}+i\varepsilon)} = \frac{i}{\omega-\omega_{if}+i\varepsilon} \tag{E-27}$$

In the limit $\varepsilon \to 0$ we get, taking into account relation (12) of Appendix II:

$$\int_0^{+\infty} \mathrm{d}\tau\, \mathrm{e}^{i(\omega-\omega_{if}+i\varepsilon)\tau} = \pi\delta(\omega-\omega_{if}) + i\mathcal{P}\frac{1}{\omega-\omega_{if}} \tag{E-28}$$

Inserting this result in (E-25) we find (E-20), where Γ and δ_{if} are given by (E-21) and (E-22).

E-2-c. Time evolution of the populations

According to (E-20), we can write $J_1 + J_1^* = \Gamma$. Inserting this result into (E-17) and using (E-19), we get:

$$\begin{aligned}\frac{\Delta\tilde{\rho}_{ii}(t)}{\Delta t} &= -\Gamma\tilde{\rho}_{ii}(t) + \Gamma\tilde{\rho}_{ff}(t)\\ \frac{\Delta\tilde{\rho}_{ff}(t)}{\Delta t} &= +\Gamma\tilde{\rho}_{ii}(t) - \Gamma\tilde{\rho}_{ff}(t)\end{aligned} \tag{E-29}$$

The interpretation of these two equations is straightforward: at any time t the system goes from $|\varphi_i\rangle$ to $|\varphi_f\rangle$, and from $|\varphi_f\rangle$ to $|\varphi_i\rangle$, with a probability per unit time that is constant and equal to Γ. If, at time t, the two populations of $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are different, they will both tend exponentially towards the same value $1/2$, without ever coming back to their initial values. This long-time irreversibility is clearly very different from what we obtained for a two-level system subjected to a sinusoidal perturbation. We no longer observe the oscillating and reversible behavior, similar to the Rabi precession of the spin $1/2$ associated with the two-level system (§ E-1).

One may wonder how a prediction valid for long times can be obtained while using perturbation calculations limited to second order in W: expressions such as (E-11) and

(E-13) are certainly no longer valid for large values of Δt. This is due to the random character of the perturbation, which has a correlation time τ_c much shorter than the evolution time Γ^{-1}. At any time t (even very distant from the initial time $t = 0$) the system has little memory of its past evolution. Between the instant t and $t + \Delta t$, its evolution only depends on what occurred before t during the time interval $[t - \tau_c, t]$. When τ_c is very short compared to Δt, the system barely evolves during the time t to $t+\Delta t$, and a perturbation treatment can be applied. This amounts to dividing the time into time intervals of width Δt, very long compared to τ_c, but nevertheless very short compared to the characteristic evolution time Γ^{-1}.

E-2-d. Time evolution of the coherences

In addition to the populations (E-15) of the two levels, we must also consider the "coherences" existing between them. It is the non-diagonal matrix elements of the density operator:

$$\tilde{\rho}_{if}(t + \Delta t) = \overline{[b_i(t + \Delta t)]\ [b_f(t + \Delta t)]^*} \tag{E-30}$$

that characterize the existence of coherent linear superpositions of the two levels. Up to second order in W, such a non-diagonal element includes zeroth order, first order and second order terms. The zeroth order term is a constant, since we defined in (E-6) the coefficients $[b_i(t)]$ and $[b_f(t)]^*$ in the interaction representation (in the usual representation, this term would correspond to the free evolution of the coherence at the Bohr frequency). The first order terms cancel out since we assumed that the average values of the perturbation matrix elements are zero. To second order in W, these coherences are obtained by first replacing, in (E-30), $[b_i(t + \Delta t)]$ and $[b_f(t + \Delta t)]^*$ by their series expansion in power of W. One must then take the average over the various values of the random perturbation. This leads to:

$$\begin{aligned}\tilde{\rho}_{if}(t + \Delta t) - \tilde{\rho}_{if}(t) = &\overline{[b_i^{(0)}(t + \Delta t)]\ [b_f^{(2)}(t + \Delta t)]^*} \\ &+ \overline{[b_i^{(2)}(t + \Delta t)]\ [b_f^{(0)}(t + \Delta t)]^*} \\ &+ \overline{[b_i^{(1)}(t + \Delta t)]\ [b_f^{(1)}(t + \Delta t)]^*} + ..\end{aligned} \tag{E-31}$$

(i) Let us consider the first two lines of this relation; we will show in (ii) below why the third line can be left out. The zeroth order coefficients, $[b_i^{(0)}(t + \Delta t)]$ and $[b_f^{(0)}(t + \Delta t)]^*$, remain equal to their initial values, written $[b_i(t)]$ and $[b_f(t)]^*$. Using the second line of (E-11), we can write:

$$\begin{aligned}&[b_i^{(2)}(t + \Delta t)]\ [b_f^{(0)}(t + \Delta t)]^* \\ &= -\frac{1}{\hbar^2} b_i(t)\ b_f^*(t) \int_0^{\Delta t} \mathrm{e}^{i\omega_{if}t'} W_{if}(t + t')\ \mathrm{d}t' \int_0^{t'} \mathrm{e}^{i\omega_{fi}t''} W_{fi}(t + t'')\ \mathrm{d}t''\end{aligned} \tag{E-32}$$

whose average value yields the second line of (E-31).

The first line of (E-31) is obtained by interchanging i and f in the second line of

(E-11), taking its complex conjugate, and multiplying the result by $b_i(t)$. This leads to:

$$[b_i^{(0)}(t+\Delta t)]\ [b_f^{(2)}(t+\Delta t)]^* = -\frac{1}{\hbar^2} b_i(t)\ b_f^*(t) \int_0^{\Delta t} \mathrm{e}^{-i\omega_{fi}t'} W_{fi}^*(t+t')\mathrm{d}t' \int_0^{t'} \mathrm{e}^{-i\omega_{if}t''} W_{if}^*(t+t'')\ \mathrm{d}t'' \tag{E-33}$$

This expression is identical to (E-32) since $\omega_{fi} = -\omega_{if}$ and $W_{fi}^* = W_{if}$.

We now average over the various values of W_{if} and W_{fi}. As we did before, we can average independently $b_i(t)b_f^*(t)$ and the double integral of (E-32). The computation of the average value of this double integral has already been performed in § E-2-b and yields $\Delta t \times J_1^*$, where J_1 is given in (E-20). This result must be doubled since the two terms (E-32) and (E-33) are equal and add up. This finally leads to:

$$\tilde{\rho}_{if}(t+\Delta t) = \tilde{\rho}_{if}(t)[1 - 2J_1^*\,\Delta t] = \tilde{\rho}_{if}(t)[1 - \Delta t(\Gamma - 2i\,\delta_{if})] \tag{E-34}$$

or else:

$$\frac{\Delta\tilde{\rho}_{if}(t)}{\Delta t} = \frac{\tilde{\rho}_{if}(t+\Delta t) - \tilde{\rho}_{if}(t)}{\Delta t} = -(\Gamma - 2i\,\delta_{if})\ \tilde{\rho}_{if}(t) \tag{E-35}$$

Let us go back to the initial components $c_i(t)$ and $c_f(t)$ of the state vector. Relation (B-8) shows that the elements of the density matrix constructed with these components are:

$$\rho_{if}(t) = \overline{[c_i(t)]\ [c_f(t)]^*} = \mathrm{e}^{-i(E_i - E_f)t/\hbar}\ \overline{[b_i(t)]\ [b_f(t)]^*} = \mathrm{e}^{-i\omega_{if}t}\tilde{\rho}_{if}(t) \tag{E-36}$$

which leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{if}(t) = -i\omega_{if}\ \rho_{if}(t)\ + \mathrm{e}^{-i\omega_{if}t}\ \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\rho}_{if}(t) \tag{E-37}$$

Now for short enough values of Δt, the derivative of $\tilde{\rho}_{if}(t)$ is given by (E-35). Using this relation in (E-37), we get:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{if}(t) = -\left[\Gamma + i\,(\omega_{if} - 2\delta_{if})\right]\rho_{if}(t) \tag{E-38}$$

This means that the coherence between $|\varphi_i\rangle$ and $|\varphi_f\rangle$ is damped at a rate Γ, and that its evolution frequency is shifted by $-2\delta_{if}$.

(ii) The third line of (E-31) is proportional to $\overline{b_f(t+\Delta t)b_i^*(t+\Delta t)}$, and is therefore responsible for a coupling between $\tilde{\rho}_{if}(t+\Delta t)$ and $\tilde{\rho}_{fi}(t+\Delta t)$: the rate of variation of $\tilde{\rho}_{if}(t+\Delta t)$ is a priori dependent on $\tilde{\rho}_{if}(t+\Delta t)$. However, if the energy difference $\hbar|\omega_{if}|$ is sufficiently large, the unperturbed evolution frequencies of these nondiagonal elements are very different, so that the effect of this coupling by the perturbation remains negligible (it actually disappears within the secular approximation). Moreover, if the statistical distribution of the random perturbation has a rotational symmetry around the Oz axis[8], this third line is equal to zero; this is demonstrated in a more general case[9] in Complement $\mathrm{E_{XIII}}$.

[8] The Oz axis is defined for the spin 1/2 associated with the two-level system.

[9] The diagonal elements of W are generally not equal to zero. This means that the coherences can also be coupled to the populations. In Complement $\mathrm{E_{XIII}}$, W is supposed to be invariant with respect to a rotation around any axis.

E-2-e. Energy shifts

The previous computation shows that the two states are shifted by the perturbation, but it does not give the value of the shift of each state; relation E-38 only simply predicts that the difference of the shifts $\hbar(\delta_i - \delta_f)$ of the two states must be equal to $-2\hbar\delta_{if}$. We will now prove that the shifts are opposite, $\delta_i = -\delta_{if}$ and $\delta_f = +\delta_{if}$. The most convenient demonstration uses the theory of the "dressed atom", which will be presented in Complement C_{XX}. A more elementary demonstration is given below.

Imagine that there exists a third, so called "spectator", state $|\varphi_k\rangle$, which is not coupled to the perturbation, so that its energy is not shifted by the perturbation. We assume that there is a coherence $\tilde{\rho}_{ik}(t)$ between $|\varphi_i\rangle$ and $|\varphi_k\rangle$, and study how it is modified by the perturbation acting on $|\varphi_i\rangle$. The computation of $\tilde{\rho}_{ik}(t+\Delta t)$ is quite similar to that of $\tilde{\rho}_{if}(t+\Delta t)$, except that $b_k(t+\Delta t)$ remains equal to $b_k(t)$, up to any order in W since the state $|\varphi_k\rangle$ is not coupled to the perturbation W. Replacing f by k in (E-31), the only non-zero term is on the second line, equal to $\overline{b_i^{(2)}(t+\Delta t)b_k^*(t)}$. The computation proceeds as for $\overline{b_i^{(2)}(t+\Delta t)b_f^*(t)}$ and leads to the same result as (E-32) where f is replaced by k. Averaging over W yields the same result as (E-34) where f is again replaced by k; the factor 2 in front of J_1^* is no longer there, since the term of the first line of (E-31) no longer comes into play to double the value of the second line. We finally get:

$$\frac{\Delta\tilde{\rho}_{ik}(t)}{\Delta t} = \frac{\tilde{\rho}_{ik}(t+\Delta t) - \tilde{\rho}_{ik}(t)}{\Delta t} = -(\frac{\Gamma}{2} - i\,\delta_{if})\;\tilde{\rho}_{ik}(t) \tag{E-39}$$

The coherence between $|\varphi_i\rangle$ and $|\varphi_k\rangle$ is damped at a rate $\Gamma/2$ and its evolution frequency, equal to $\omega_{ik} = \omega_i - \omega_k$ in the absence of the perturbation, is changed by $-\delta_{if}$. Since the state $|\varphi_k\rangle$ is not coupled to the perturbation, the state $|\varphi_i\rangle$ must be shifted by $\delta_i = -\hbar\delta_{if}$. As for the state $|\varphi_f\rangle$, it is shifted by $\delta_f = +\hbar\delta_{if}$, since relation (E-38) indicates that the difference between the two shifts of $|\varphi_i\rangle$ and $|\varphi_f\rangle$ must be equal to $-2\hbar\delta_{if}$.

Let us focus on the sign of δ_{if}, given by relation (E-22). We saw in § D-1 that $\tilde{C}_{if}(\omega)$ is real; we assume this function to be positive in an angular frequency domain centered around $\omega = \omega_c$, and zero everywhere else. Relation (E-22) shows that:

$$\omega_c > \omega_{if} \quad \Rightarrow \quad \delta_{if} > 0 \quad ; \quad \omega_c < \omega_{if} \quad \Rightarrow \quad \delta_{if} < 0 \tag{E-40}$$

For $E_i > E_f$, we have $\omega_{if} > 0$; when $\omega_c > \omega_{if}$, relation (E-38) shows that the shift δ_{if} decreases the energy difference between the two states $|\varphi_i\rangle$ and $|\varphi_f\rangle$. For $E_i < E_f$, we have $\omega_{if} < 0$; it is now when $\omega_c < \omega_{if} < 0$ that the shift decreases the energy difference. In both cases, and if ω_c has the same sign as ω_{if}, the energy difference is decreased when $|\omega_c| > |\omega_{if}|$; in the opposite case, the energy levels get further from each other.

E-3. Broadband optical excitation of an atom

We now apply the previous results to the excitation of a two-level atom by broadband radiation. The radiation is described by an incoherent superposition of monochromatic fields with frequencies ω spreading over an interval of width $\Delta\omega$, and with random phases. Consequently, the coupling between the atom and the radiation is a random perturbation. The larger $\Delta\omega$, the shorter the perturbation correlation time, as we shall see below. We can directly use the results of §§ D and E, to obtain the absorption rate of the radiation by the atom, as well as the energy shifts due to the coupling between the atom and the radiation[10].

[10]This problem is also studied in § 3-b of Complement A_{XIII} using a different method. In that complement, we shall sum the transition probabilities associated with each of the monochromatic waves

E-3-a. Correlation functions of the interaction Hamiltonian

The matrix element $W_{fi}(t)$ of the perturbation W associated with the atom-radiation coupling can be written as:

$$W_{fi}(t) = -\langle\varphi_f| DE(t) |\varphi_i\rangle = -\mathcal{D}E(t) \tag{E-41}$$

where D is the electric dipole moment of the atom and $E(t)$ the electric field of the incident radiation[11]; we have set:

$$\mathcal{D} = \langle\varphi_f| D |\varphi_i\rangle \tag{E-42}$$

and will assume, for the sake of simplicity, that $\mathcal{D}$ is real (this can be obtained by a change of the relative phase of $|\varphi_i\rangle$ and $|\varphi_f\rangle$). The correlation function of the perturbation is then proportional to that of the electric field:

$$\overline{W_{fi}(t) W_{if}(t-\tau)} = \mathcal{D}^2 \, \overline{E(t) E(t-\tau)} \tag{E-43}$$

This field can be expanded on its Fourier components:

$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}\omega \; \mathrm{e}^{i\omega t} \; \tilde{E}(\omega) \tag{E-44}$$

with, since the field is real:

$$\tilde{E}(-\omega) = \tilde{E}^*(\omega) \tag{E-45}$$

We assume that the phases of the ω components are independent of each other, and completely random. This leads to:

$$\overline{\tilde{E}(\omega)\tilde{E}^*(\omega')} = I(\omega) \;\; \delta(\omega - \omega') \tag{E-46}$$

where $I(\omega)$describes the spectral distribution of the incident radiation; the function $I(\omega)$ is supposed to have a width $\Delta\omega$.

Taking (E-45) and (E-46) into account, we can now write the correlation function (E-43) as:

$$\begin{aligned} \overline{W_{fi}(t) W_{if}(t-\tau)} &= \frac{\mathcal{D}^2}{2\pi} \int \mathrm{d}\omega \int \mathrm{d}\omega' \; \mathrm{e}^{i\omega t} \; \mathrm{e}^{-i\omega'(t-\tau)} \; \overline{\tilde{E}(\omega) \tilde{E}^*(\omega')} \\ &= \frac{\mathcal{D}^2}{2\pi} \int \mathrm{d}\omega \; I(\omega) \; \mathrm{e}^{i\omega\tau} \end{aligned} \tag{E-47}$$

We first note that this function only depends on the difference of times: the perturbation is a stationary random function. Secondly, if the spectral distribution $I(\omega)$ of the incident radiation is a bell shaped curve of width $\Delta\omega$, the time correlation function of the perturbation decreases with a characteristic time τ_c inversely proportional to this width:

$$\tau_c \simeq \frac{1}{\Delta\omega} \tag{E-48}$$

This means that, if we assume that the atom is excited by a radiation with a broad enough spectrum, the correlation time τ_c will be short enough to fulfill the conditions necessary for applying the results of §§ D and E.

present in the incident radiation.

[11] To simplify the equations, we shall ignore the vectorial nature of $E(t)$ et D.

E-3-b. Absorption rates and light shifts

Relation (E-47) shows that:

$$C_{if}(\omega) = \frac{\mathcal{D}^2}{\sqrt{2\pi}} I(\omega) \tag{E-49}$$

Using (D-19) for the value of Γ, we get:

$$\Gamma = \frac{\mathcal{D}^2}{\hbar^2} I(\omega_{if}) \tag{E-50}$$

As we saw above, Γ yields the transition rate per unit time between the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$. This rate is proportional to $I(\omega_{if})$, i.e. to the Fourier transform of the time correlation function of the perturbation, calculated at the frequency ω_{if} of the transition. This result is entirely different from the result obtained with a monochromatic radiation. In this latter case, and at resonance, one expects a Rabi oscillation between the two states $|\varphi_i\rangle$ and $|\varphi_f\rangle$.

When considering the excitation probability of an atom in its ground state, the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are, respectively, the lower and upper states of the transition. This means that:

$$E_i - E_f = \hbar\omega_{if} < 0 \tag{E-51}$$

The angular frequency ω_{fi} appearing in (E-50) is then negative: in an absorption process, it is the Fourier components with negative frequencies ω that come into play (note however that for an electric field in $\cos\omega t$ or $\sin\omega t$, the positive and negative frequency components have the same intensity, and the distinction is no longer essential). Furthermore, the previous calculations are still valid when $E_i - E_f > 0$, a case that corresponds to "stimulated emission", or induced emission (see § C-2 of Chapter XX). During this process, the radiation stimulates the transition of the atom from an excited state to its ground state. This treatment justifies the introduction by Einstein (§ C-4 of Chapter XX) of the coefficients B_{fi} and B_{if} describing the absorption and stimulated emission in the presence of black body radiation (which is broadband).

We can also evaluate the atomic energy shifts due to the presence of the radiation. The results of § E-2-c show that the excitation of the atom by broadband radiation shifts the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ by the respective values $-\hbar\delta_{if}$ and $+\hbar\delta_{if}$. The "light shift" δ_{if} is proportional (with a positive proportionality constant) to the following integral over ω:

$$\delta_{if} \propto \mathcal{P}\int \mathrm{d}\omega\ I(\omega)\frac{1}{\omega - \omega_{if}} \tag{E-52}$$

These light shifts are proportional to the light intensity since they depend linearly on $I(\omega)$. Their sign depends on the detuning between the central frequency ω_c of the incident radiation and the frequency ω_{fi} of the atomic transition. As we have seen in § E-2-e, if $|\omega_c|$ is larger that $|\omega_{if}|$, meaning that the incident radiation is detuned towards the blue, the energies of the two levels get closer under the effect of the radiation. We get the opposite conclusion if the incident radiation is detuned towards the red. These light shifts will be further discussed using the "dressed atom approach" in Complement $\mathrm{C_{XX}}$. We will show that they also exist for an atom excited by monochromatic radiation, and that they are useful tools for manipulating atoms and their motion.

References and suggestions for further reading:

Perturbation series expansion of the evolution operator: Messiah (1.17), Chap. XVII, §§ 1 and 2.

Sudden or adiabatic modification of the Hamiltonian: Messiah (1.17), Chap. XVII, § II ; Schiff (1.18), Chap. 8, § 35.

Diagramatic representation of a perturbation series (Feynman diagrams) : Ziman (2.26), Chap. 3 ; Mandl (2.9), Chaps. 12 to 14 ; Bjorkén and Drell (2.10), Chaps. 16 and 17.

A_{XIII} : INTERACTION OF AN ATOM WITH AN ELECTROMAGNETIC WAVE

Illustration of the general considerations of § C-2 of Chapter XIII, using the important example of an atom interacting with a sinusoidal electromagnetic wave. Introduces fundamental concepts such as: spectral line selection rules, absorption and induced emission of radiation, oscillator strength... Although moderately difficult, can be recommended for a first reading, because of the importance of the concepts introduced.

B_{XIII} : LINEAR AND NON-LINEAR RESPONSE OF A TWO-LEVEL SYSTEM SUBJECTED TO A SINUSOIDAL PERTURBATION

A simple model for the study of some non-linear effects that appear in the interaction of an electromagnetic wave with an atomic system (saturation effects, multiple-quanta transitions, etc.). More difficult than A_{XIII} (graduate level); should therefore be reserved for a subsequent study.

C_{XIII} : OSCILLATIONS OF A SYSTEM BETWEEN TWO DISCRETE STATES UNDER THE EFFECT OF A RESONANT PERTURBATION

Study of the behavior, over a long time interval, of a system that has discrete energy levels, subjected to a resonant perturbation. Completes, in greater detail, the results of § C-2 of Chapter XIII, which are valid only for short times. Relatively simple.

D_{XIII} : DECAY OF A DISCRETE STATE RESONANTLY COUPLED TO A CONTINUUM OF FINAL STATES

Study of the behavior, over a long time interval, of a discrete state resonantly coupled to a continuum of final states. Completes the results obtained in § C-3 of Chapter XIII (Fermi golden rule), which were established only for short time intervals. Proves that the probability of finding the particle in the discrete level decreases exponentially, and justifies the concept of lifetimes introduced phenomenologically in Complement K_{III}. Important for its numerous physical applications; graduate level.

E_{XIII} : TIME-DEPENDENT RANDOM PERTURBATION, RELAXATION

This complement provides a more detailed and precise view of the study of §§ D and E-2 on the effects of a random perturbation. The "motional narrowing condition" is assumed to be valid, which means that the memory time of the perturbation is much shorter than the time it takes for the perturbation to have a significant effect. This complement first part, this complement provides de general equations of evolution of the density matrix. In a second part, the theory is applied to an ensemble of spins 1/2 coupled to a random isotropic perturbation. This complement is important because of its numerous applications: magnetic resonance, optics, etc.

F_{XIII} : EXERCISES

Exercise 10 can be done at the end of Complement A_{XIII}; it is a step by step study of the effects of the external degrees of freedom of a quantum mechanical system on the frequencies of the electromagnetic radiation it can absorb (Doppler effect, recoil energy, Mössbauer effect).

Some exercises (especially 8 and 9) are more difficult than others, but treat important phenomena.

Complement A_{XIII}

Interaction of an atom with an electromagnetic wave

1 The interaction Hamiltonian. Selection rules 1340

1-a Fields and potentials associated with a plane electromagnetic wave 1340

1-b The interaction Hamiltonian at the low-intensity limit 1341

1-c The electric dipole Hamiltonian 1342

1-d The magnetic dipole and electric quadrupole Hamiltonians . 1347

2 Non-resonant excitation. Comparison with the elastically bound electron model 1350

2-a Classical model of the elastically bound electron 1350

2-b Quantum mechanical calculation of the induced dipole moment 1351

2-c Discussion. Oscillator strength 1352

3 Resonant excitation. Absorption and induced emission . . 1353

3-a Transition probability associated with a monochromatic wave 1353

3-b Broad-line excitation. Transition probability per unit time . . 1354

In § C of Chapter XIII, we studied the special case of a sinusoidally time-dependent perturbation: $W(t) = W \sin \omega t$. We encountered the resonance phenomenon which occurs when ω is close to one of the Bohr angular frequencies $\omega_{fi} = (E_f - E_i)/\hbar$ of the physical system under consideration.

A particularly important application of this theory is the treatment of an atom interacting with a monochromatic wave. In this complement, we will use this example to illustrate the general considerations of Chapter XIII and to introduce certain fundamental concepts of atomic physics such as spectral line selection rules, absorption and induced emission of radiation, oscillator strength, etc...

As in Chapter XIII, we shall confine ourselves to first-order perturbation calculations. Some higher-order effects in the interaction of an atom with an electromagnetic wave ("non-linear" effects) will be taken up in Complement B_{XIII}.

We shall begin (§ 1) by analyzing the structure of the interaction Hamiltonian between an atom and the electromagnetic field. This will permit us to isolate the electric dipole, magnetic dipole and electric quadrupole terms, and to study the corresponding selection rules. Then we shall calculate the electric dipole moment induced by a non-resonant incident wave (§ 2) and compare the results obtained with those of the model of the elastically bound electron. Finally, we shall study (§ 3) the processes of absorption and induced emission of radiation which appear in the resonant excitation of an atom.

Comment:

In all complements of Chapter XIII the atom is treated quantum mechanically, but the electromagnetic field is treated classically, as a time-dependent perturbation acting on the atom. In Chapter XX and its complements, a more elaborate study will be given with a full quantum treatment of both the electromagnetic field and the atom; the interaction hamiltonian is

then time-independent. This permits the description of physical effects such as the spontaneous emission of photons by atoms in excited states, which does not appear when the field is treated classically.

1. The interaction Hamiltonian. Selection rules

1-a. Fields and potentials associated with a plane electromagnetic wave

Consider a plane electromagnetic wave[1], of wave vector $\mathbf{k}$ (parallel to Oy) and angular frequency $\omega = ck$. The electric field of the wave is parallel to Oz and the magnetic field, to Ox (Fig. 1).

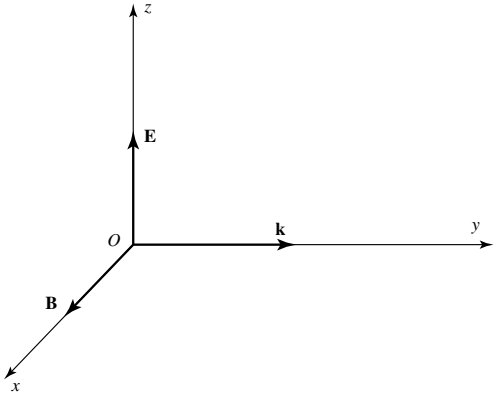


Figure 1: The electric field E and magnetic field B of a plane wave of wave vector $\mathbf{k}$.

For such a wave, it is always possible, with a suitable choice of gauge (*cf.* Appendix III, § 4-b-α), to make the scalar potential $U(\mathbf{r}, t)$ zero. The vector potential $\mathbf{A}(\mathbf{r}, t)$ is then given by the real expression:

$$\mathbf{A}(\mathbf{r}, t) = \mathcal{A}_0 \mathbf{e}_z \mathrm{e}^{i(ky-\omega t)} + \mathcal{A}_0^* \mathbf{e}_z \mathrm{e}^{-i(ky-\omega t)} \tag{1}$$

where $\mathcal{A}_0$ is a complex constant whose argument depends on the choice of the time origin. We then have:

$$\mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t}\mathbf{A}(\mathbf{r}, t) = i\omega \mathcal{A}_0 \mathbf{e}_z \mathrm{e}^{i(ky-\omega t)} - i\omega \mathcal{A}_0^* \mathbf{e}_z \mathrm{e}^{-i(ky-\omega t)} \tag{2}$$

$$\mathbf{B}(\mathbf{r}, t) = \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r}, t) = ik \mathcal{A}_0 \mathbf{e}_x \mathrm{e}^{i(ky-\omega t)} - ik \mathcal{A}_0^* \mathbf{e}_x \mathrm{e}^{-i(ky-\omega t)} \tag{3}$$

We shall choose the time origin such that the constant $\mathcal{A}_0$ is pure imaginary, and we set:

$$i\omega \mathcal{A}_0 = \frac{\mathcal{E}}{2} \tag{4a}$$

$$ik \mathcal{A}_0 = \frac{\mathcal{B}}{2} \tag{4b}$$

[1]For the sake of simplicity, we shall confine ourselves here to the case of a plane wave. The results obtained in this complement, however, can be generalized to an any electromagnetic field.

where $\mathcal{E}$ and $\mathcal{B}$ are two real quantities such that:

$$\frac{\mathcal{E}}{\mathcal{B}} = \frac{\omega}{k} = c \tag{5}$$

We then obtain:

$$\mathbf{E}(\mathbf{r}, t) = \mathcal{E}\mathbf{e}_z \cos(ky - \omega t) \tag{6}$$

$$\mathbf{B}(\mathbf{r}, t) = \mathcal{B}\mathbf{e}_x \cos(ky - \omega t) \tag{7}$$

$\mathcal{E}$ and $\mathcal{B}$ are therefore the amplitudes of the electric and magnetic fields of the plane wave considered.

Finally, we shall calculate the Poynting vector[2] $\mathbf{G}$ associated with this plane wave:

$$\mathbf{G} = \varepsilon_0 c^2 \mathbf{E} \times \mathbf{B} \tag{8}$$

Replacing $\mathbf{E}$ and $\mathbf{B}$ in (8) by their expressions (6) and (7), and taking the time-average value over a large number of periods, we obtain, using (5):

$$\overline{\mathbf{G}} = \varepsilon_0 c \frac{\mathcal{E}^2}{2} \mathbf{e}_y \tag{9}$$

1-b. The interaction Hamiltonian at the low-intensity limit

The preceding wave interacts with an atomic electron (of mass m and charge q) situated at a distance r from O and bound to this point O by a central potential $V(r)$ (created by a nucleus assumed to be motionless at O). The quantum mechanical Hamiltonian of this electron can be written:

$$H = \frac{1}{2m}[\mathbf{P} - q\mathbf{A}(\mathbf{R}, t)]^2 + V(R) - \frac{q}{m}\mathbf{S} \cdot \mathbf{B}(\mathbf{R}, t) \tag{10}$$

The last term of (10) represents the interaction of the spin magnetic moment of the electron with the oscillating magnetic field of the plane wave. $\mathbf{A}(\mathbf{R}, t)$ and $\mathbf{B}(\mathbf{R}, t)$ are the operators obtained by replacing, in the classical expressions (1) and (3), x, y, z by the observables X, Y, Z.

In expanding the square that appears on the right-hand side of (10), we should, in theory, remember that $\mathbf{P}$ does not generally commute with a function of $\mathbf{R}$. Such a precaution is, however, unnecessary in the present case, since, as $\mathbf{A}$ is parallel to Oz [formula (1)], only the P_z component enters into the double product; now P_z commutes with the Y component of $\mathbf{R}$, which is the only one to appear in expression (1) for $\mathbf{A}(\mathbf{R}, t)$. We can then take:

$$H = H_0 + W(t) \tag{11}$$

where:

$$H_0 = \frac{\mathbf{P}^2}{2m} + V(R) \tag{12}$$

[2] Recall that the energy flux across a surface element $\mathrm{d}S$ perpendicular to the unit vector $\mathbf{n}$ is equal to $\mathbf{G} \cdot \mathbf{n}\, \mathrm{d}S$.

is the atomic Hamiltonian, and:

$$W(t) = -\frac{q}{m}\mathbf{P}\cdot\mathbf{A}(\mathbf{R},t) - \frac{q}{m}\mathbf{S}\cdot\mathbf{B}(\mathbf{R},t) + \frac{q^2}{2m}[\mathbf{A}(\mathbf{R},t)]^2 \tag{13}$$

is the interaction Hamiltonian with the incident plane wave [the matrix elements of $W(t)$ approach zero when $\mathcal{A}_0$ approaches zero].

The first two terms on the right-hand side of (13) depend linearly on $\mathcal{A}_0$, and the third one depends on it quadratically. With ordinary light sources, the intensity is sufficiently low that the effect of the $\mathcal{A}_0^2$ term can be neglected compared to that of the $\mathcal{A}_0$ term. We shall therefore write:

$$W(t) \simeq W_I(t) + W_{II}(t) \tag{14}$$

with:

$$W_I(t) = -\frac{q}{m}\mathbf{P}\cdot\mathbf{A}(\mathbf{R},t) \tag{15}$$

$$W_{II}(t) = -\frac{q}{m}\mathbf{S}\cdot\mathbf{B}(\mathbf{R},t) \tag{16}$$

We shall evaluate the relative orders of magnitude of the matrix elements of $W_I(t)$ and $W_{II}(t)$ between two bound states of the electron. Those of $\mathbf{S}$ are of the order of $\hbar$, and $\mathbf{B}$ is of the order of $k\mathcal{A}_0$ [*cf.* formula (3)]. Thus:

$$\frac{W_{II}(t)}{W_I(t)} \simeq \frac{\frac{q}{m}\hbar k\mathcal{A}_0}{\frac{q}{m}p\mathcal{A}_0} = \frac{\hbar k}{p} \tag{17}$$

According to the uncertainty relations, $\hbar/p$ is, at most, of the order of atomic dimensions (characterized by the Bohr radius, $a_0 \simeq 0.5$ Å). k is equal to $2\pi/\lambda$, where λ is the wavelength associated with the incident wave. In the spectral domains used in atomic physics (the optical or Hertzian domains), λ is much greater than a_0, so that:

$$\frac{W_{II}(t)}{W_I(t)} \simeq \frac{a_0}{\lambda} \ll 1 \tag{18}$$

1-c. The electric dipole Hamiltonian

α. The electric dipole approximation. Interpretation

Using expression (1) for $\mathbf{A}(\mathbf{R},t)$, we can put $W_I(t)$ in the form:

$$W_I(t) = -\frac{q}{m}P_z[\mathcal{A}_0 e^{ikY}e^{-i\omega t} + \mathcal{A}_0^* e^{-ikY}e^{i\omega t}] \tag{19}$$

We now expand the exponential $e^{\pm ikY}$ in powers of kY:

$$e^{\pm ikY} = 1 \pm ikY - \frac{1}{2}k^2Y^2 + ... \tag{20}$$

Since Y is of the order of atomic dimensions, we have, as above:

$$kY \simeq \frac{a_0}{\lambda} \ll 1 \tag{21}$$

We therefore obtain a good approximation for W_I by retaining only the first term of expansion (20). Let W_{DE} be the operator obtained by replacing $e^{\pm ikY}$ by 1 on the right-hand side of (19). Using (4-a), we get:

$$W_{DE}(t) = \frac{q\mathcal{E}}{m\omega} P_z \sin \omega t \tag{22}$$

$W_{DE}(t)$ is called the "electric dipole Hamiltonian". The electric dipole approximation, which is based on conditions (18) and (21), therefore consists of neglecting $W_{II}(t)$ relative to $W_I(t)$ and identifying $W_I(t)$ with $W_{DE}(t)$:

$$W(t) \simeq W_{DE}(t) \tag{23}$$

Let us show that, if we replace $W(t)$ by $W_{DE}(t)$, the electron oscillates as if it were subjected to a *uniform* sinusoidal electric field $\mathcal{E}\mathbf{e}_z \cos \omega t$, whose amplitude is that of the electric field of the incident plane wave evaluated at the point O. Physically, this means that the wave function of the bound electron is too localized about O for the electron to "feel" the spatial variation of the electric field of the incident plane wave. We shall therefore calculate the evolution of $\langle \mathbf{R} \rangle(t)$. Ehrenfest's theorem (*cf.* Chap. III, § D-1-d) leads to:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{R} \rangle = \dfrac{1}{i\hbar}\langle [\mathbf{R}, H_0 + W_{DE}] \rangle = \dfrac{\langle \mathbf{P} \rangle}{m} + \dfrac{q\mathcal{E}}{m\omega}\mathbf{e}_z \sin \omega t \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{P} \rangle = \dfrac{1}{i\hbar}\langle [\mathbf{P}, H_0 + W_{DE}] \rangle = -\langle \boldsymbol{\nabla} V(R) \rangle \end{cases} \tag{24}$$

Eliminating $\langle \mathbf{P} \rangle$ from these two equations, we obtain, after a simple calculation:

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle \mathbf{R} \rangle = -\langle \boldsymbol{\nabla} V(R) \rangle + q\mathcal{E}\mathbf{e}_z \cos \omega t \tag{25}$$

which is indeed the predicted result: the center of the wave packet associated with the electron moves like a particle of mass m and charge q, subjected to both the central force of the atomic bond [the first term on the right-hand side of (25)] and the influence of a uniform electric field [the second term of (25)].

Comment:

Expression (22) for the electric dipole interaction Hamiltonian seems rather unusual for a particle of charge q interacting with a uniform electric field $\mathbf{E} = \mathcal{E}\mathbf{e}_z \cos \omega t$. We would tend to write the interaction Hamiltonian in the form:

$$W'_{DE}(t) = -\mathbf{D} \cdot \mathbf{E} = -q\mathcal{E}Z \cos \omega t \tag{26}$$

where $\mathbf{D} = q\mathbf{R}$ is the electric dipole moment associated with the electron.

Actually, expressions (22) and (26) are equivalent. We shall show that we can go from one to the other by a gauge transformation (which does not modify the physical content of quantum mechanics; *cf.* Complement $\mathrm{H_{III}}$). The gauge used to obtain (22) is:

$$\begin{cases} \mathbf{A}(\mathbf{r}, t) = \dfrac{\mathcal{E}}{\omega}\mathbf{e}_z \sin(ky - \omega t) & \text{(27a)} \\ U(\mathbf{r}, t) = 0 & \text{(27b)} \end{cases}$$

[to write (27a), we have replaced $\mathcal{A}_0$ by $\mathcal{E}/2i\omega$ in (1); *cf.* formula (4a)]. Now consider the gauge transformation associated with the function:

$$\chi(\mathbf{r},t) = z\frac{\mathcal{E}}{\omega}\sin\omega t \tag{28}$$

It leads to a new gauge $\{\mathbf{A}', U'\}$ characterized by:

$$\begin{cases} \mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\chi = \mathbf{e}_z\dfrac{\mathcal{E}}{\omega}[\sin(ky-\omega t)+\sin\omega t] & \text{(29a)} \\ U' = U - \dfrac{\partial\chi}{\partial t} = -z\mathcal{E}\cos\omega t & \text{(29b)} \end{cases}$$

The electric dipole approximation amounts to replacing ky by 0 everywhere. We then see that in this approximation:

$$\mathbf{A}' \simeq \mathbf{e}_z\frac{\mathcal{E}}{\omega}[\sin(-\omega t)+\sin\omega t] = 0 \tag{30}$$

If, in addition, we neglect, as we did above, the magnetic interaction terms related to the spin, we obtain, for the system's Hamiltonian:

$$\begin{aligned} H' &= \frac{1}{2m}(\mathbf{P}-q\mathbf{A}')^2 + V(R) + qU'(\mathbf{R},t) \\ &= \frac{\mathbf{P}^2}{2m} + V(R) + qU'(\mathbf{R},t) \\ &= H_0 + W'(t) \end{aligned} \tag{31}$$

where H_0 is the atomic Hamiltonian given by (12), and:

$$W'(t) = qU'(\mathbf{R},t) = -qZ\mathcal{E}\cos\omega t = W'_{DE}(t) \tag{32}$$

is the usual form (26) of the electric dipole interaction Hamiltonian.

Recall that the state of the system is no longer described by the same ket when we go from gauge (27) to gauge (29) (*cf.* Complement H_{III}). The replacement of $W_{DE}(t)$ by $W'_{DE}(t)$ must therefore be accompanied by a change of state vector, the physical content, of course, remaining the same.

In the rest of this complement, we shall continue to use gauge (27).

β. The matrix elements of the electric dipole Hamiltonian

Later, we shall need the expressions for the matrix elements of W_{DE} between $|\varphi_i\rangle$ and $|\varphi_f\rangle$, eigenstates of H_0 of eigenvalues E_i and E_f. According to (22), these matrix elements can be written:

$$\langle\varphi_f|W_{DE}(t)|\varphi_i\rangle = \frac{q\mathcal{E}}{m\omega}\sin\omega t\langle\varphi_f|P_z|\varphi_i\rangle \tag{33}$$

It is simple to replace the matrix elements of P_z by those of Z on the right-hand side of (33). Insofar as we are neglecting all magnetic effects in the atomic Hamiltonian [*cf.* expression (12) for H_0], we can write:

$$[Z, H_0] = i\hbar\frac{\partial H_0}{\partial P_z} = i\hbar\frac{P_z}{m} \tag{34}$$

which yields:

$$\langle\varphi_f|[Z,H_0]|\varphi_i\rangle = \langle\varphi_f|ZH_0 - H_0Z|\varphi_i\rangle$$
$$= -(E_f - E_i)\langle\varphi_f|Z|\varphi_i\rangle = \frac{i\hbar}{m}\langle\varphi_f|P_z|\varphi_i\rangle \tag{35}$$

Introducing the Bohr angular frequency $\omega_{fi} = (E_f - E_i)/\hbar$, we then get:

$$\langle\varphi_f|P_z|\varphi_i\rangle = im\omega_{fi}\langle\varphi_f|Z|\varphi_i\rangle \tag{36}$$

and, consequently:

$$\langle\varphi_f|W_{DE}(t)|\varphi_i\rangle = iq\frac{\omega_{fi}}{\omega}\mathcal{E}\sin\omega t\langle\varphi_f|Z|\varphi_i\rangle \tag{37}$$

The matrix elements of $W_{DE}(t)$ are therefore proportional to those of Z.

Comment:

It is the matrix element of Z which appears in (37) because we have chosen the electric field $\mathbf{E}(\mathbf{r},t)$ parallel to Oz. In practice, we may be led to choose a frame $Oxyz$ related, not to the light polarization, but to the symmetry of the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$. For example, if the atoms are placed in a uniform magnetic field $\mathbf{B}_0$, the most convenient quantization axis for the study of their stationary states $|\varphi_n\rangle$ is obviously parallel to $\mathbf{B}_0$. The polarization of the electric field $\mathbf{E}(\mathbf{r},t)$ can then be arbitrary relative to Oz. In this case, we must replace the matrix element of Z in (37) by that of a linear combination of X, Y and Z.

γ. Electric dipole transition selection rules

If the matrix element of W_{DE} between the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ is different from zero, that is, if $\langle\varphi_f|Z|\varphi_i\rangle$ is non-zero[3], the transition $|\varphi_i\rangle \longrightarrow |\varphi_f\rangle$ is said to be an electric dipole transition. To study the transitions induced between $|\varphi_i\rangle$ and $|\varphi_f\rangle$ by the incident wave, we can then replace $W(t)$ by $W_{DE}(t)$. If, on the other hand, the matrix element of $W_{DE}(t)$ between $|\varphi_i\rangle$ and $|\varphi_f\rangle$ is zero, we must pursue the expansion of $W(t)$ further, and the corresponding transition is either a magnetic dipole transition or an electric quadrupole transition, etc...[4] (see following sections). Since $W_{DE}(t)$ is much larger than the subsequent terms of the power series expansion of $W(t)$ in a_0/λ, electric dipole transitions will be, by far, the most intense. In fact, most optical lines emitted by atoms correspond to electric dipole transitions.

Let:

$$\begin{cases} \varphi_{n_i,l_i,m_i}(\mathbf{r}) = R_{n_i,l_i}(r)Y_{l_i}^{m_i}(\theta,\varphi) \\ \varphi_{n_f,l_f,m_f}(\mathbf{r}) = R_{n_f,l_f}(r)Y_{l_f}^{m_f}(\theta,\varphi) \end{cases} \tag{38}$$

[3] Actually, it suffices for one of the three numbers $\langle\varphi_f|Z|\varphi_i\rangle$, $\langle\varphi_f|X|\varphi_i\rangle$ or $\langle\varphi_f|Y|\varphi_i\rangle$ to be different from zero (*cf.* comment of § β above).

[4] It may happen that all the terms of the expansion have zero matrix elements. The transition is then said to be forbidden to all orders (it can be shown that this is always the case if $|\varphi_i\rangle$ and $|\varphi_f\rangle$ both have zero angular momenta).

be the wave functions associated with $|\varphi_i\rangle$ and $|\varphi_f\rangle$. Since:

$$z = r\cos\theta = \sqrt{\frac{4\pi}{3}}\; r\, Y_1^0(\theta) \tag{39}$$

the matrix element of Z between $|\varphi_i\rangle$ and $|\varphi_f\rangle$ is proportional to the angular integral:

$$\int \mathrm{d}\Omega\; Y_{l_f}^{m_f *}(\theta,\varphi)\; Y_1^0(\theta)\; Y_{l_i}^{m_i}(\theta,\varphi) \tag{40}$$

According to the results of Complement C_X, this integral is different from zero only if:

$$l_f = l_i \pm 1 \tag{41}$$

and:

$$m_f = m_i \tag{42}$$

Actually, it would suffice to choose another polarization of the electric field (for example, parallel to Ox or Oy; see comment of § β) to have:

$$m_f = m_i \pm 1 \tag{43}$$

From (41), (42) and (43), we obtain the electric dipole transition selection rules:

$$\Delta l = l_f - l_i = \pm 1 \tag{44a}$$
$$\Delta m = m_f - m_i = -1, 0, +1 \tag{44b}$$

Comments:

(*i*) Z is an odd operator. It can connect only two states of different parities. Since the parities of $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are those of l_i and l_f, $\Delta l = l_f - l_i$ must be odd, as is compatible with (44a).

(*ii*) If there exists a spin-orbit coupling $\xi(r)\mathbf{L}\cdot\mathbf{S}$ between $\mathbf{L}$ and $\mathbf{S}$ (*cf.* Chap. XII, § B-1-b-β), the stationary states of the electron are labeled by the quantum numbers l, s, J, m_J (with $\mathbf{J} = \mathbf{L} + \mathbf{S}$). The electric dipole transition selection rules can be obtained by looking for the non-zero matrix elements of $\mathbf{R}$ in the $\{|l, s, J, m_J\rangle\}$ basis. By using the expansions of these basis vectors on the kets $|l, m\rangle|s, m_S\rangle$ (*cf.* Complement A_X, § 2), we find, starting with (44a) and (44b), the selection rules:

$$\Delta J = 0, \pm 1 \tag{44a}$$
$$\Delta l = \pm 1 \tag{44b}$$
$$\Delta m_J = 0, \pm 1 \tag{44c}$$

Note that a $\Delta J = 0$ transition is not forbidden [unless $J_i = J_f = 0$; *cf.* note on the preceding page]. This is due to the fact that J is not related to the parity of the level.

Finally, we point out that selection rules (44a, 44b, 44c) can be generalized to many-electron atoms.

1-d. The magnetic dipole and electric quadrupole Hamiltonians

α. Higher-order terms in the interaction Hamiltonian

The interaction Hamiltonian given by (14) can be written in the form:

$$W(t) = W_I(t) + W_{II}(t) = W_{DE}(t) + [W_I(t) - W_{DE}(t)] + W_{II}(t) \tag{45}$$

Thus far, we have studied $W_{DE}(t)$. As we have seen, the ratio of $W_I(t) - W_{DE}(t)$ and $W_{II}(t)$ to $W_{DE}(t)$ is of the order of a_0/λ.

To calculate $W_I(t) - W_{DE}(t)$, we simply replace $\mathrm{e}^{\pm ikY}$ by $\mathrm{e}^{\pm ikY} - 1 \simeq \pm ikY + ...$ in (19), which yields:

$$W_I(t) - W_{DE}(t) = -\frac{q}{m}[ik\mathcal{A}_0\mathrm{e}^{-i\omega t} - ik\mathcal{A}_0^*\mathrm{e}^{i\omega t}]P_zY + ... \tag{46}$$

or, using (4b):

$$W_I(t) - W_{DE}(t) = -\frac{q}{m}\mathcal{B}\cos\omega t P_zY + ... \tag{47}$$

If we write P_zY in the form:

$$P_zY = \frac{1}{2}(P_zY - ZP_y) + \frac{1}{2}(P_zY + ZP_y) = \frac{1}{2}L_x + \frac{1}{2}(P_zY + ZP_y) \tag{48}$$

we obtain, finally:

$$W_I(t) - W_{DE}(t) = -\frac{q}{2m}L_x\mathcal{B}\cos\omega t - \frac{q}{2m}\mathcal{B}\cos\omega t[YP_z + ZP_y] + ... \tag{49}$$

In the expression for $W_{II}(t)$ [formulas (16) and (3)], it is entirely justified to replace $\mathrm{e}^{\pm ikY}$ by 1. We thus obtain a term of order a_0/λ relative to $W_I(t)$, that is, of the same order of magnitude as $W_I(t) - W_{DE}(t)$:

$$W_{II}(t) = -\frac{q}{m}S_x\mathcal{B}\cos\omega t + ... \tag{50}$$

Substituting (49) and (50) into (45) and grouping the terms differently, we obtain:

$$W(t) = W_{DE}(t) + W_{DM}(t) + W_{QE}(t) + ... \tag{51}$$

with:

$$W_{DM} = -\frac{q}{2m}(L_x + 2S_x)\mathcal{B}\cos\omega t \tag{52}$$

$$W_{QE} = -\frac{q}{2mc}(YP_z + ZP_y)\mathcal{E}\cos\omega t \tag{53}$$

[we have replaced $\mathcal{B}$ by $\mathcal{E}/c$ in (53)]. W_{DM} and W_{QE} (which are, *a priori*, of the same order of magnitude) are, respectively, the magnetic dipole and electric quadrupole Hamiltonians.

β. Magnetic dipole transitions

The transitions induced by W_{DM} are called magnetic dipole transitions. W_{DM} represents the interaction of the total magnetic moment of the electron with the oscillating magnetic field of the incident wave.

The magnetic dipole transition selection rules can be obtained by considering the conditions which must be met by $|\varphi_i\rangle$ and $|\varphi_f\rangle$ in order for W_{DM} to have a non-zero matrix element between these two states. Since neither L_x nor S_x changes the quantum number l, we must have, first of all, $\Delta l = 0$. L_x changes the eigenvalue m_L of L_z by ± 1, which gives $\Delta m_L = \pm 1$. S_x changes the eigenvalues m_S of S_z by ± 1, so that $\Delta m_S = \pm 1$. Note, furthermore, that if the magnetic field of the incident wave were parallel to Oz, we would have $\Delta m_L = 0$ and $\Delta m_S = 0$. Grouping these results, we obtain the magnetic dipole transition selection rules:

$$\begin{cases} \Delta l = 0 \\ \Delta m_L = \pm 1, 0 \\ \Delta m_S = \pm 1, 0 \end{cases} \tag{54}$$

Comment:

In the presence of a spin-orbit coupling, the eigenstates of H_0 are labeled by the quantum numbers l and J. Since L_x and S_x do not commute with $\mathbf{J}^2$, W_{DM} can connect states with the same l but different J. By using the addition formulas for an angular momentum l and an angular momentum $1/2$ (*cf.* Complement A_X, § 2), it can easily be shown that selection rules (54) become:

$$\begin{cases} \Delta l = 0 \\ \Delta J = \pm 1, 0 \\ \Delta m_J = \pm 1, 0 \end{cases} \tag{55}$$

Note that the hyperfine transition $F = 0 \leftrightarrow F = 1$ of the ground state of the hydrogen atom (*cf.* Chap. XII, § D) is a magnetic dipole transition, since the components of $\mathbf{S}$ have non-zero matrix elements between the $F = 1$ states and the $|F = 0,\ m_F = 0\rangle$ state.

γ. Electric quadrupole transitions

Using (34), we can write:

$$\begin{aligned} YP_z + ZP_y = YP_z + P_yZ &= \frac{m}{i\hbar}\{Y[Z, H_0] + [Y, H_0]Z\} \\ &= \frac{m}{i\hbar}(YZH_0 - H_0YZ) \end{aligned} \tag{56}$$

from which we obtain, as in (36):

$$\langle\varphi_f|W_{QE}(t)|\varphi_i\rangle = \frac{q}{2ic}\omega_{fi}\langle\varphi_f|YZ|\varphi_i\rangle\mathcal{E}\cos\omega t \tag{57}$$

The matrix element of $W_{QE}(t)$ is therefore proportional to that of YZ, which is a component of the electric quadrupole moment of the atom (*cf.* Complement E_X). In

addition, the following quantity appears in (57):

$$\frac{q\omega_{fi}}{c}\mathcal{E} = q\frac{\omega_{fi}}{\omega}\frac{\omega}{c}\mathcal{E} = q\frac{\omega_{fi}}{\omega}k\mathcal{E} \tag{58}$$

which, according to (2), is of the order of $q\,\partial\mathcal{E}_z/\partial y$. The operator $W_{QE}(t)$ can therefore be interpreted as the interaction of the electric quadrupole moment of the atom with the gradient[5] of the electric field of the plane wave.

To obtain the electric quadrupole transition selection rules, we simply note that, in the $\{|\mathbf{r}\rangle\}$ representation, YZ is a linear superposition of $r^2Y_2^1(\theta,\varphi)$ and $r^2Y_2^{-1}(\theta,\varphi)$. Therefore, in the matrix element $\langle\varphi_f|YZ|\varphi_i\rangle$ there appear angular integrals:

$$\int \mathrm{d}\Omega Y_{l_f}^{m_f *}(\theta,\varphi)Y_2^{\pm 1}(\theta,\varphi)Y_{l_i}^{m_i}(\theta,\varphi) \tag{59}$$

which, according to the results of Complement $\mathrm{C_X}$, are different from zero only if $\Delta l = 0$, ± 2 and $\Delta m = \pm 1$. This last relation becomes $\Delta m = \pm 2, \pm 1,\ 0$ when we consider an arbitrary polarization of the incident wave (*cf.* comment of § 1-c-β), and the electric quadrupole transition selection rules can be written, finally:

$$\begin{cases} \Delta l = 0, \pm 2 \\ \Delta m = 0, \pm 1, \pm 2 \end{cases} \tag{60}$$

Comments:

(*i*) W_{DM} and W_{QE} are even operators and can therefore connect only states of the same parity, which is compatible with (54) and (60). For a given transition, W_{DM} and W_{QE} are never in competition with W_{DE}. This facilitates the observation of magnetic dipole and electric quadrupole transitions.

Most of the transitions that occur in the microwave or radio-frequency domain – in particular, magnetic resonance transitions (*cf.* Complement $\mathrm{F_{IV}}$) – are magnetic dipole transitions.

(*ii*) For a $\Delta l = 0$, $\Delta m = 0, \pm 1$ transition, the two operators W_{DM} and W_{QE} simultaneously have non-zero matrix elements. However, it is possible to find experimental conditions under which only magnetic dipole transitions are induced. All we need to do is place the atom, not in the path of a plane wave, but inside a cavity or radiofrequency loops, at a point where $\mathbf{B}$ is large but the gradient of $\mathbf{E}$ is negligible.

(*iii*) For a $\Delta l = 2$ transition, W_{DM} cannot be in competition with W_{QE}, and we have a pure quadrupole transition. As an example of a quadrupole transition, we can mention the green line of atomic oxygen (5577 Å), which appears in the aurora borealis spectrum.

(*iv*) If we pursued the expansion of $e^{\pm ikY}$ further, we would find electric octupole and magnetic quadrupole terms, etc.

In the rest of this complement, we shall confine ourselves to electric dipole transitions. In the next Complement, $\mathrm{B_{XIII}}$, on the other hand, we shall consider a magnetic dipole transition.

[5] It is normal for the electric field gradient to appear, since $W_{QE}(t)$ was obtained by expanding the potentials in a Taylor series in the neighborhood of O.

2. Non-resonant excitation. Comparison with the elastically bound electron model

In this section, we shall assume that the atom, initially in the ground state $|\varphi_0\rangle$, is excited by a non-resonant plane wave: ω coincides with none of the Bohr angular frequencies associated with transitions from $|\varphi_0\rangle$.

Under the effect of this excitation, the atom acquires an electric dipole moment $|\mathbf{D}\rangle(t)$ which oscillates at the angular frequency ω (forced oscillation) and is proportional to $\mathcal{E}$ when $\mathcal{E}$ is small (linear response). We shall use perturbation theory to calculate this induced dipole moment, and we shall show that the results obtained are very close to those found with the classical model of the elastically bound electron.

This model has played a very important role in the study of the optical properties of materials. It enables us to calculate the polarization induced by the incident wave in a material. This polarization, which depends linearly on the field $\mathcal{E}$, behaves like a source term in Maxwell's equations. When we solve these equations, we find plane waves propagating in the material at a velocity different from c. This allows us to calculate the refractive index of the material in terms of various characteristics of elastically bound electrons (natural frequencies, number per unit volume, etc.). Thus, we see that it is very important to compare the predictions of this model (which we shall review in § a) with those of quantum mechanics.

2-a. Classical model of the elastically bound electron

α. Equation of motion

Consider an electron subjected to a restoring force directed towards the point O and proportional to the displacement. In the classical Hamiltonian corresponding to (12), we then have:

$$V(r) = \frac{1}{2} m\omega_0^2 r^2 \tag{61}$$

where ω_0 is the electron's natural angular frequency.

If we make the same approximations, using the classical interaction Hamiltonian, as those which enabled us to obtain expression (22) for $W_{DE}(t)$ (the electric dipole approximation) in quantum mechanics, a calculation similar to that of § 1-c-α [*cf.* equation (25)] yields the equation of motion:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} z + \omega_0^2 z = \frac{q\mathcal{E}}{m} \cos \omega t \tag{62}$$

This is the equation of a harmonic oscillator subject to a sinusoidal force.

β. General solution

The general solution of (62) can be written:

$$z = A \cos(\omega_0 t - \varphi) + \frac{q\mathcal{E}}{m(\omega_0^2 - \omega^2)} \cos \omega t \tag{63}$$

where A and φ are real constants which depend on the initial conditions. The first term of (63), $A\cos(\omega_0 t - \varphi)$, represents the general solution of the homogeneous equation (the

electron's free motion). The second term is a particular solution of the equation (forced motion of the electron).

We have not, thus far, taken damping into account. Without going into detailed calculations, we shall cite the effects of weak damping: after a certain time τ, it causes the natural motion to disappear and very slightly modifies the forced motion (provided that we are sufficiently far from resonance: $|\omega - \omega_0| \gg 1/\tau$). We shall therefore retain only the second term of (63):

$$z = \frac{q\mathcal{E}\cos\omega t}{m(\omega_0^2 - \omega^2)} \tag{64}$$

Comment:

Far from resonance, the exact damping mechanism is of little importance, provided that it is weak. We shall not, therefore, take up the problem of the exact description of this damping, either in quantum or in classical mechanics. We shall merely use the fact that it exists to ignore the free motion of the electron.

It would be different for a resonant excitation: the induced dipole moment would then depend critically on the exact damping mechanism (spontaneous emission, thermal relaxation, etc.). This is why we shall not try to calculate $\langle \mathbf{D} \rangle(t)$ in § 3 (the case of a resonant excitation). We shall be concerned only with calculating the transition probabilities.

In Complement B_{XIII}, we shall study a specific model of a system placed in an electromagnetic wave and at the same time subject to dissipative processes (Bloch equations of a system of spins). We shall then be able to calculate the induced dipole moment for any exciting frequency.

γ. *Susceptibility*

Let $\mathcal{D} = qz$ be the electric dipole moment of the system. According to (64), we have:

$$\mathcal{D} = qz = \frac{q^2}{m(\omega_0^2 - \omega^2)}\mathcal{E}\cos\omega t = \chi\mathcal{E}\cos\omega t \tag{65}$$

where the "susceptibility" χ is given by:

$$\chi = \frac{q^2}{m(\omega_0^2 - \omega^2)} \tag{66}$$

2-b. Quantum mechanical calculation of the induced dipole moment

We shall begin by calculating, to first order in $\mathcal{E}$, the state vector $|\psi(t)\rangle$ of the atom at time t. We shall choose for the interaction Hamiltonian, the electric dipole Hamiltonian W_{DE} given by (22). In addition, we shall assume that:

$$|\psi(t=0)\rangle = |\varphi_0\rangle \tag{67}$$

We apply the results of § C-1 of Chapter XIII, replacing W_{ni} by $\frac{q\mathcal{E}}{m\omega}\langle\varphi_n|P_z|\varphi_i\rangle$ and $|\varphi_i\rangle$ by $|\varphi_0\rangle$. This leads to[6]:

$$|\psi(t)\rangle = \mathrm{e}^{-iE_0t/\hbar}|\varphi_0\rangle + \sum_{n\neq 0} \lambda\ b_n^{(1)}(t)\ \mathrm{e}^{-iE_nt/\hbar}|\varphi_n\rangle \tag{68}$$

or, using (C-4) of Chapter XIII and multiplying $|\psi(t)\rangle$ by the global phase factor $\mathrm{e}^{iE_0t/\hbar}$, which has no physical importance:

$$|\psi(t)\rangle = |\varphi_0\rangle + \sum_{n\neq 0} \frac{q\mathcal{E}}{2im\hbar\omega}\langle\varphi_n|P_z|\varphi_0\rangle \left\{\frac{\mathrm{e}^{-i\omega_{n0}t} - \mathrm{e}^{i\omega t}}{\omega_{n0}+\omega} - \frac{\mathrm{e}^{-i\omega_{n0}t} - \mathrm{e}^{-i\omega t}}{\omega_{n0}-\omega}\right\}|\varphi_n\rangle \tag{69}$$

From this, we find $\langle\psi(t)|$ and $\langle D_z\rangle(t) = \langle\psi(t)|qZ|\psi(t)\rangle$. In the calculation of this average value, we retain only the terms linear in $\mathcal{E}$, and we neglect all those that oscillate at angular frequencies $\pm\omega_{n0}$ (the natural motion, which would disappear if we took weak damping into account). Finally, replacing $\langle\varphi_n|P_z|\varphi_0\rangle$ by its expression in terms of $\langle\varphi_n|Z|\varphi_0\rangle$ [*cf.* equation (36)], we find:

$$\langle D_z\rangle(t) = \frac{2q^2}{\hbar}\mathcal{E}\cos\omega t \sum_n \frac{\omega_{n0}\ |\langle\varphi_n|Z|\varphi_0\rangle|^2}{\omega_{n0}^2 - \omega^2} \tag{70}$$

2-c. Discussion. Oscillator strength

α. The concept of oscillator strength

We set:

$$f_{n0} = \frac{2m\ \omega_{n0}\ |\langle\varphi_n|Z|\varphi_0\rangle|^2}{\hbar} \tag{71}$$

f_{n0} is a real dimensionless number, characteristic of the $|\varphi_0\rangle \leftrightarrow |\varphi_n\rangle$ transition and called the oscillator strength[7] of this transition. If $|\varphi_0\rangle$ is the ground state, f_{n0} is positive, like ω_{n0}.

Oscillator strengths satisfy the following sum rule (the Thomas-Reiche-Kuhn sum rule):

$$\sum_n f_{n0} = 1 \tag{72}$$

This can be shown as follows. Using (36), we can write:

$$f_{n0} = \frac{1}{i\hbar}\langle\varphi_0|Z|\varphi_n\rangle\langle\varphi_n|P_z|\varphi_0\rangle - \frac{1}{i\hbar}\langle\varphi_0|P_z|\varphi_n\rangle\langle\varphi_n|Z|\varphi_0\rangle \tag{73}$$

The summation over n can be performed by using the closure relation relative to the $\{|\varphi_n\rangle\}$ basis, and we get:

$$\sum_n f_{n0} = \frac{1}{i\hbar}\langle\varphi_0|(ZP_z - P_zZ)|\varphi_0\rangle = \langle\varphi_0|\varphi_0\rangle = 1 \tag{74}$$

[6] Since W_{DE} is odd, $\langle\varphi_0|W_{DE}(t)|\varphi_0\rangle$ is zero, so $b_0^{(1)}(t) = 0$.

[7] The operator Z enters into (71) because the incident wave is linearly polarized along Oz. It would, however, be possible to give a general definition of the oscillator strength, independent of the polarization of the incident wave.

β. The quantum mechanical justification for the elastically bound electron model

We substitute definition (71) into (70) and multiply the expression so obtained by the number $\mathcal{N}$ of atoms contained in a volume whose linear dimensions are much smaller than the wavelength λ of the radiation. The total electric dipole moment induced in this volume can then be written:

$$\mathcal{N}\langle D_z\rangle(t) = \sum_n \mathcal{N} f_{n0}\ \frac{q^2}{m(\omega_{n0}^2 - \omega^2)}\ \mathcal{E}\cos\omega t \tag{75}$$

Comparing (75) and (65), we see that it is like having $\mathcal{N}$ classical oscillators [since $\sum_n \mathcal{N} f_{n0} = \mathcal{N}$ according to (72)] whose natural angular frequencies are not all the same since they are equal to the various Bohr angular frequencies of the atom associated with the transitions from $|\varphi_0\rangle$. According to (75), the proportion of oscillators with the angular frequency ω_{n0} is f_{n0}.

Thus, for a non-resonant wave, we have justified the classical model of the elastically bound electron. Quantum mechanics gives the frequencies of the various oscillators and the proportion of oscillators that have a given frequency. This result shows the importance of the concept of oscillator strength and enables us to understand *a posteriori* why the elastically bound electron model was so useful in the study of the optical properties of materials.

3. Resonant excitation. Absorption and induced emission

3-a. Transition probability associated with a monochromatic wave

Consider an atom initially in the state $|\varphi_i\rangle$ placed in an electromagnetic wave whose angular frequency is close to a Bohr angular frequency ω_{fi}.

The results of § C-1 of Chapter XIII (sinusoidal excitation) are directly applicable to the calculation of the transition probability $\mathscr{P}_{if}(t;\omega)$. We find, using expression (37) (thus making the electric dipole approximation):

$$\mathscr{P}_{if}(t;\omega) = \frac{q^2}{4\hbar^2}\left(\frac{\omega_{fi}}{\omega}\right)^2 |\langle\varphi_f|Z|\varphi_i\rangle|^2\ \mathcal{E}^2 F(t,\omega-\omega_{fi}) \tag{76}$$

where:

$$F(t,\omega-\omega_{fi}) = \left\{\frac{\sin[(\omega_{fi}-\omega)t/2]}{(\omega_{fi}-\omega)/2}\right\}^2 \tag{77}$$

We have already discussed the resonant nature of $\mathscr{P}_{if}(t;\omega)$ in Chapter XIII. At resonance, $\mathscr{P}_{if}(t;\omega)$ is proportional to $\mathcal{E}^2$, that is, to the incident flux of electromagnetic energy [*cf.* formula (9)].

Comments:

(*i*) If instead of using the gauge (27), leading to the matrix element (37), we had used the gauge (29) leading to the Hamiltonian (32), the factor $(\omega_{fi}/\omega)^2$ in (76) would be missing. The fact that the results are different is not at all surprising. The states $|\varphi_f\rangle$ and $|\varphi_i\rangle$, and consequently $\mathscr{P}_{if}(t;\omega)$, do not have the same physical meaning in the two gauges.

(*ii*) However, as $t \to \infty$, the diffraction function $F(t, \omega - \omega_{fi})$ tends towards $\delta(\omega - \omega_{fi})$, and the factor $(\omega_{fi}/\omega)^2$ approaches unity. This leads to the same probability density $\mathscr{P}_{if}(t;\omega)$ in the two gauges. This result can be easily understood if we consider the incident electromagnetic wave to be a quasi-monochromatic wave packet of very large but finite spatial extent, rather than a plane wave extending to infinity. When $t \to \pm\infty$ the **E** field "seen" by the atom tends towards zero and the gauge transformation associated with the function χ defined in (28) tends towards unity. Consequently $|\varphi_i\rangle$ and $|\varphi_f\rangle$ each represent the same physical states in the two gauges.

(*iii*) Obviously, it is also possible to consider the transition probability between two well-defined energy states of the atomic system for a finite time interval. In this case, the eigenstates $|\varphi_i\rangle$ and $|\varphi_f\rangle$ of the atomic Hamiltonian H_0 written in (12) only represent states of well-defined atomic energy (kinetic plus potential) in the gauge (29) where **A** is zero [see (30)] and $\mathbf{p}^2/2m$ represents the kinetic energy. The same physical states would be represented in gauge (27) by the states $\exp[-iq\chi(\mathbf{r},t)/\hbar]|\varphi_i\rangle$ and $\exp[-iq\chi(\mathbf{r},t)/\hbar]|\varphi_f\rangle$ respectively. For finite t, calculations are therefore simpler in the gauge (29). Since in the rest of this complement we replace $F(t, \omega - \omega_{fi})$ by $\delta(\omega - \omega_{fi})$ [see (79)], we will be considering the limit $t \to \infty$ for which the above difficulties disappear.

3-b. Broad-line excitation. Transition probability per unit time

In practice, the radiation which strikes the atom is very often non-monochromatic. We shall denote by $\mathcal{F}(\omega)\,\mathrm{d}\omega$ the incident flux of electromagnetic energy per unit surface within the interval $[\omega, \omega + \mathrm{d}\omega]$. The variation of $\mathcal{F}(\omega)$ with respect to ω is shown in Figure 2. Δ is the excitation line width. If Δ is infinite, we say that we are dealing with a "white spectrum".

The different monochromatic waves which constitute the incident radiation are generally incoherent: they have no well-defined phase relation. The total transition probability $\overline{\mathscr{P}}_{if}$ can therefore be obtained by summing the transition probabilities associated with each of these monochromatic waves. We must, consequently, replace $\mathcal{E}^2$ by $2\,\mathcal{F}(\omega)\,\mathrm{d}\omega/\varepsilon_0 c$ in (76) [formula (9)] and integrate over ω. This gives:

$$\overline{\mathscr{P}}_{if}(t) = \frac{q^2}{2\varepsilon_0 c\hbar^2}\,|\langle\varphi_f|Z|\varphi_i\rangle|^2 \int \mathrm{d}\omega \left(\frac{\omega_{fi}}{\omega}\right)^2 \mathcal{F}(\omega)\,F(t, \omega - \omega_{fi}) \tag{78}$$

We can then proceed as in § C-3 of Chapter XIII to evaluate the integral that appears in (78). Compared to a function of ω whose width is much larger than $4\pi/t$, the function $F(t, \omega - \omega_{fi})$ (see Figure 3 of Chapter XIII) behaves like $\delta(\omega - \omega_{fi})$. If t is large enough to make $4\pi/t \ll \Delta$ (Δ: excitation line width) while remaining small enough for the perturbation treatment to be valid, we can, in (78), assume that:

$$F(t, \omega - \omega_{fi}) \simeq 2\pi t\,\delta(\omega - \omega_{fi}) \tag{79}$$

which yields:

$$\overline{\mathscr{P}}_{if}(t) = \frac{\pi q^2}{\varepsilon_0 c\hbar^2}\,|\langle\varphi_f|Z|\varphi_i\rangle|^2\,\mathcal{F}(\omega_{fi})\,t \tag{80}$$

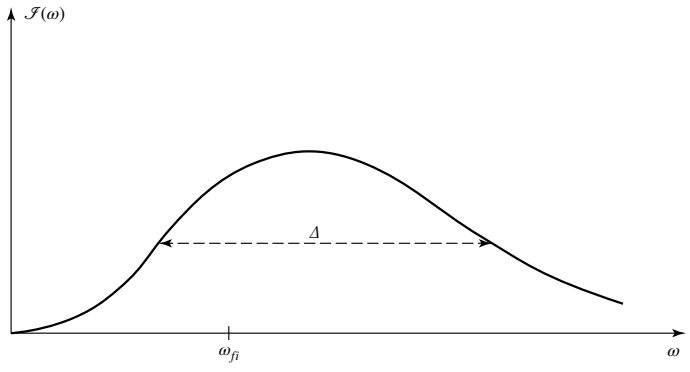


Figure 2: The spectral distribution of the incident flux of electromagnetic energy per unit surface. Δ is the width of this spectral distribution.

We can write (80) in the form:

$$\overline{\mathscr{P}}_{if}(t) = C_{if}\mathcal{F}(\omega_{fi})\,t \tag{81}$$

where:

$$C_{if} = \frac{4\pi^2}{\hbar}|\langle\varphi_f|Z|\varphi_i\rangle|^2\alpha \tag{82}$$

and α is the fine-structure constant:

$$\alpha = \frac{q^2}{4\pi\varepsilon_0}\frac{1}{\hbar c} = \frac{e^2}{\hbar c} \simeq \frac{1}{137} \tag{83}$$

This result shows that $\overline{\mathscr{P}}_{if}(t)$ increases linearly with time. The *transition probability per unit time* $\mathcal{W}_{if}$ is therefore equal to:

$$\mathcal{W}_{if} = C_{if}\mathcal{F}(\omega_{fi}) \tag{84}$$

$\mathcal{W}_{if}$ is proportional to the value of the incident intensity for the resonance frequency ω_{fi}, to the fine-structure constant α, and to the square of the modulus of the matrix element of Z, which is related [by (71)] to the oscillator strength of the $|\varphi_f\rangle \leftrightarrow |\varphi_i\rangle$ transition.

In this complement, we have considered the case of radiation propagating along a given direction with a well-defined polarization state. By averaging the coefficients C_{if} over all propagation directions and over all possible polarization states, we could introduce coefficients B_{if}, analogous to the coefficients C_{if}, defining the transition probabilities per unit time for an atom placed in isotropic radiation. The coefficients B_{if} (and B_{fi}) are none other than the coefficients introduced by Einstein to describe the absorption (and induced emission). Thus, we see how quantum mechanics enables us to calculate these coefficients.

Comment:

A third coefficient, A_{fi}, was introduced by Einstein to describe the spontaneous emission of a photon, which occurs when the atom falls back from the upper state $|\varphi_f\rangle$ to the lower state $|\varphi_i\rangle$. The theory presented in this complement does not explain spontaneous emission. In the absence of incident radiation, the interaction Hamiltonian is zero, and the eigenstates of H_0 (which is then the total Hamiltonian) are stationary states.

Actually, the preceding model is insufficient, since it treats asymmetrically the atomic system (which is quantized) and the electromagnetic field (which is considered classically). When we quantize both systems, we find, even in the absence of incident photons, that the coupling between the atom and the electromagnetic field continues to have observable effects (a simple interpretation of these effects is given in Complement K_V). The eigenstates of H_0 are no longer stationary states, since H_0 is no longer the Hamiltonian of the total system, and we can indeed calculate the probability per unit time of spontaneous emission of a photon (*cf.* Chap. XX, § C-3) . Quantum mechanics therefore also enables us to obtain the Einstein coefficient A_{fi}.

References and suggestions for further reading:

See, for example: Schiff (1.18), Chap. 11; Bethe and Jackiw (1.21), Part II, Chaps. 10 and 11; Bohm (5.1), Chap. 18, §§ 12 to 44.

For the elastically bound electron model: Berkeley 3 (7.1), supplementary topic 9; Feynman I (6.3), Chap. 31 and Feynman II (7.2), Chap. 32.

For Einstein coefficients: the original article (1.31), Cagnac and Pebay-Peyroula (11.2), Chap. III and Chap. XIX, § 4.

For the exact definition of oscillator strength: Sobel'man (11.12), Chap. 9, § 31.

For atomic multipole radiation and its selection rules: Sobel'man (11.12), Chap. 9, § 32.

Complement B_{XIII}

Linear and non-linear responses of a two-level system subject to a sinusoidal perturbation

1 Description of the model 1358

1-a Bloch equations for a system of spin 1/2's interacting with a radiofrequency field 1358

1-b Some exactly and approximately soluble cases 1359

1-c Response of the atomic system 1359

2 The approximate solution of the Bloch equations of the system 1360

2-a Perturbation equations 1360

2-b The Fourier series expansion of the solution 1362

2-c The general structure of the solution 1363

3 Discussion 1364

3-a Zeroth-order solution: competition between pumping and relaxation 1364

3-b First-order solution: the linear response 1364

3-c Second-order solution: absorption and induced emission . . . 1366

3-d Third-order solution: saturation effects and multiple-quanta transitions 1368

4 Exercises: applications of this complement 1372

In the preceding complement, we applied first-order time-dependent perturbation theory to the treatment of some effects produced by the interaction of an atomic system and an electromagnetic wave: appearance of an induced dipole moment, induced emission and absorption processes, etc.

We shall now consider a simple example, in which it is possible to pursue the perturbation calculations to higher orders without too many complications. This will allow us to demonstrate some interesting "non-linear" effects: saturation effects, non-linear susceptibility, the absorption and induced emission of several photons, etc. In addition, the model we shall describe takes into account (phenomenologically) the dissipative coupling of the atomic system with its surroundings (the relaxation process). This will enable us to complete the results related to the "linear response" obtained in the preceding complement. For example, we shall calculate the atom's induced dipole moment, not only far from resonance, but also at resonance.

Some of the effects we are going to describe are objects of a great deal of research. Their study necessitates very strong electromagnetic fields, which cab be obtained only with lasers. New branches of research have thus appeared with lasers: quantum electronics, non-linear optics, etc. The calculation methods described in this complement (for a very simple model) are applicable to these problems.

Comment:

The Comment at the end of the introduction of Complement A_{XIII} applies to the present complement as well: here, we limit ourselves to a semi-classical treatment, where the atomic system is treated quantum mechanically but the electromagnetic field classically. A full quantum treatment of both systems will be given in Chapter XX; see in particular Complement C_{XX}, which describes the "dressed atom" method and non-perturbative calculations.

1. Description of the model

1-a. Bloch equations for a system of spin 1/2's interacting with a radiofrequency field

We shall return to the system described in § 4-a of Complement F_{IV}: a system of spin 1/2's placed in a static field $\mathbf{B}_0$ parallel to Oz, interacting with an oscillating radiofrequency field and subject to "pumping" and "relaxation" processes.

If $\boldsymbol{\mathcal{M}}(t)$ is the total magnetization of the spin system contained in the cell (Fig. 6 of Complement F_{IV}), we showed in Complement F_{IV} that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{M}}(t) = n\boldsymbol{\mu}_0 - \frac{1}{T_R}\boldsymbol{\mathcal{M}}(t) + \gamma\boldsymbol{\mathcal{M}}(t) \times \mathbf{B}(t) \tag{1}$$

The first term on the right-hand side describes the preparation, or the "pumping" of the system: n spins enter the cell per unit time, each one with an elementary magnetization $\boldsymbol{\mu}_0$ parallel to Oz. The second term arises from relaxation processes, characterized by the average time T_R required for a spin either to leave the cell or have its direction changed by collision with the walls. Finally, the last term of (1) corresponds to the precession of the spins about the total magnetic field:

$$\mathbf{B}(t) = B_0\mathbf{e}_z + \mathbf{B}_1(t) \tag{2}$$

$\mathbf{B}(t)$ is the sum of a static field $B_0\mathbf{e}_z$ parallel to Oz and a radiofrequency field $\mathbf{B}_1(t)$ of angular frequency ω.

Comments:

(*i*) The transitions which we shall study in this complement (which connect the two states $|+\rangle$ and $|-\rangle$ of each spin 1/2) are magnetic dipole transitions.

(*ii*) One could question our using expression (1) relative to average values rather than the Schrödinger equation. We do so because we are studying a statistical ensemble of spins coupled to a thermal reservoir (via collisions with the cell walls). We cannot describe this ensemble in terms of a state vector: we must use a density operator (see Complement E_{III}). The equation of motion of this operator is called a "master equation" and we can show that it is exactly equivalent to (1) (see Complement F_{IV}, § 3 and 4, and Complement E_{IV}, where we show that the average value of the magnetization determines the density matrix of an ensemble of spin 1/2's).

It turns out that the master equation satisfied by the density operator and the Schrödinger equation studied in § C-1 of Chapter XIII have the same structure as (1): a linear differential equation, with constant or sinusoidally varying coefficients. The approximation methods we describe in this chapter are, therefore, applicable to any of these equations.

1-b. Some exactly and approximately soluble cases

If the radiofrequency field $\mathbf{B}_1(t)$ is rotating, that is, if:

$$\mathbf{B}_1(t) = B_1(\mathbf{e}_x \cos\omega t + \mathbf{e}_y \sin\omega t) \tag{3}$$

equation (1) can be solved exactly [changing to the frame which is rotating with $\mathbf{B}_1$ transforms (1) into a time-independent linear differential system]. The exact solution of (1) corresponding to such a situation is given in § 4-b of Complement $\mathrm{F_{IV}}$.

Here, we shall assume $\mathbf{B}_1$ to be linearly polarized along Ox:

$$\mathbf{B}_1(t) = B_1\mathbf{e}_x \cos\omega t \tag{4}$$

In this case, it is not possible[1] to find an exact analytic solution of equation (1) (there is no transformation equivalent to changing to the rotating frame). We shall see, however, that a solution can be found in the form of a power series expansion in $\mathbf{B}_1$.

Comment:

The calculations we shall present here for spin 1/2's can also be applied to other situations in which we can confine ourselves to two levels of the system and ignore all others. We know (*cf.* Complement $\mathrm{C_{IV}}$) that we can associate a fictitious spin 1/2 with any two-level system. The problem considered here is therefore that of an arbitrary two-level system subject to a sinusoidal perturbation.

1-c. Response of the atomic system

The set of terms which, in $\mathcal{M}_x$, $\mathcal{M}_y$, $\mathcal{M}_z$, depend on B_1 constitute the "response" of the atom to the electromagnetic perturbation. They represent the magnetic dipole moment induced in the spin system by the radiofrequency field. We shall see that such a dipole moment is not necessarily proportional to B_1; the terms in B_1 represent the linear response, and the others (terms in B_1^2, B_1^3, ...), the "non-linear response". In addition, we shall see that the induced dipole moment does not oscillate only at the angular frequency ω, but also at its various harmonics $p\omega$ ($p = 0$, 2, 3, 4, ...).

It is easy to see why we should be interested in calculating the response of the atomic system. Such a calculation is useful for the theory of the propagation of an electromagnetic wave in a material, or for the theory of atomic oscillators, "masers" or "lasers".

Consider an electromagnetic field. Because of the coupling between this field and the atomic system, a polarization appears in the material, due to the atomic dipole moments (arrow directed towards the right in Figure 1). This polarization acts like a source term in Maxwell's equations and contributes to the creation of the electromagnetic field (arrow directed towards the left in Figure 1). When we "close the loop", that is, when we take the field so created to be equal to the one with which we started, we obtain the wave propagation equations in the material (refractive index) or the oscillator equations (in the absence of external fields, an

[1] A linearly polarized field can be obtained as a superposition of a left and a right circular components. It would be possible to find an exact solution for each of these components taken separately. However, equation (1) is not linear, in the sense that a solution corresponding to (4) cannot be obtained by superposing two exact solutions, one of which corresponds to (3) and the other one to the field rotating in the opposite direction [in the term $\gamma\boldsymbol{\mathcal{M}} \times \mathbf{B}$ that appears on the right-hand side of (1), $\boldsymbol{\mathcal{M}}$ depends on $\mathbf{B}_1$].

electromagnetic field may appear in the material, if there is sufficient amplification: the system becomes unstable and can oscillate spontaneously). In this complement, we shall be concerned only with the first step of the calculation (the atomic response).

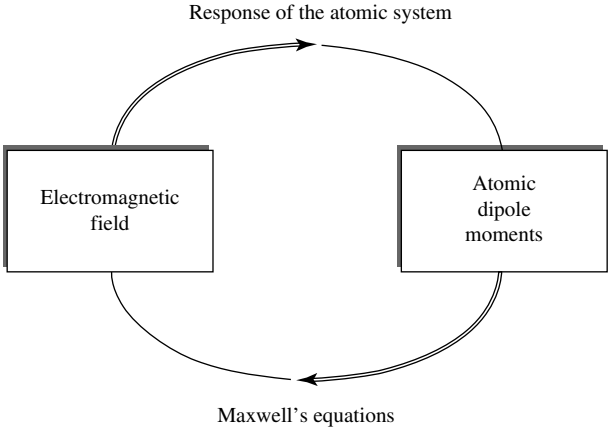


Figure 1: Schematic representation of the calculations to be performed in studying the propagation of an electromagnetic wave in a material (or the operation of an atomic oscillator, a laser or a maser). We begin by calculating the dipole moments induced in the material by a given electromagnetic field (the response of the atomic system). The corresponding polarization acts like a source term in Maxwell's equations and contributes to the creation of the electromagnetic field. We then take the field obtained to be equal to the one with which we started.

2. The approximate solution of the Bloch equations of the system

2-a. Perturbation equations

As in Complement F_{IV}, we set:

$$\begin{cases} \omega_0 = -\gamma B_0 & (5) \\ \omega_1 = -\gamma B_1 & (6) \end{cases}$$

$\hbar\omega_0$ represents the energy difference of the spin states $|+\rangle$ and $|-\rangle$ (Fig. 2). Substituting (4) into (2), and (2) into (1), we obtain, after a simple calculation:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}_z = n\mu_0 - \dfrac{\mathcal{M}_z}{T_R} + i\dfrac{\omega_1}{2}\cos\omega t(\mathcal{M}_- - \mathcal{M}_+) & (7a) \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}_\pm = -\dfrac{\mathcal{M}_\pm}{T_R} + i\omega_0\mathcal{M}_\pm \mp i\omega_1\cos\omega t\mathcal{M}_z & (7b) \end{cases}$$

with:

$$\mathcal{M}_\pm = \mathcal{M}_x \pm i\mathcal{M}y \quad (8)$$

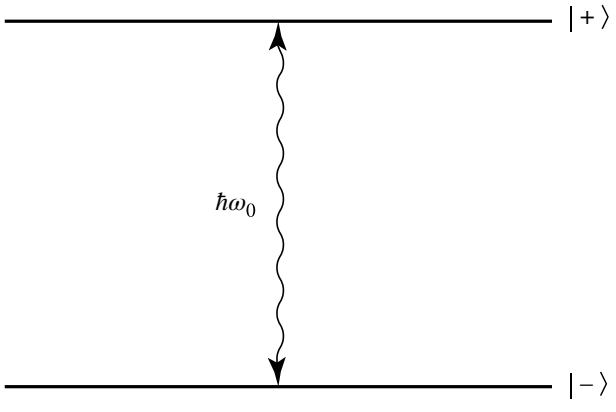


Figure 2: Energy levels of a spin 1/2 in a static magnetic field $\mathbf{B}_0$; ω_0 is the Larmor angular frequency in the field $\mathbf{B}_0$.

Note that the source term $n\mu_0$ exists only in the equation of motion of $\mathcal{M}_z$, since $\boldsymbol{\mu}_0$ is parallel to Oz, and the pumping is said to be longitudinal[2]. We also point out that the relaxation time can be different for the longitudinal components ($\mathcal{M}_z$) and the transverse components ($\mathcal{M}_\pm$) of the magnetization. For the sake of simplicity, we shall choose a single relaxation time here.

Equations (7a) and (7b), called the "Bloch equations", cannot be solved exactly. We shall therefore determine their solution in the form of a power series expansion in ω_1:

$$\begin{cases} \mathcal{M}_z = {}^{(0)}\mathcal{M}_z + \omega_1\,{}^{(1)}\mathcal{M}_z + \omega_1^2\,{}^{(2)}\mathcal{M}_z + ... + \omega_1^n\,{}^{(n)}\mathcal{M}_z + ... & \text{(9a)} \\ \mathcal{M}_\pm = {}^{(0)}\mathcal{M}_\pm + \omega_1\,{}^{(1)}\mathcal{M}_\pm + \omega_1^2\,{}^{(2)}\mathcal{M}_\pm + ... + \omega_1^n\,{}^{(n)}\mathcal{M}_\pm + ... & \text{(9b)} \end{cases}$$

Substituting (9a) and (9b) into (7a) and (7b), and setting equal the coefficients of terms in ω_1^n, we obtain the following perturbation equations:

n=0 :

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^{(0)}\mathcal{M}_z = n\mu_0 - \dfrac{1}{T_R}\,{}^{(0)}\mathcal{M}_z & \text{(10a)} \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^{(0)}\mathcal{M}_\pm = -\dfrac{1}{T_R}\,{}^{(0)}\mathcal{M}_\pm \pm i\omega_0\,{}^{(0)}\mathcal{M}_\pm & \text{(10b)} \end{cases}$$

$n \neq 0$:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^{(n)}\mathcal{M}_z = -\dfrac{1}{T_R}\,{}^{(n)}\mathcal{M}_z + \dfrac{i}{2}\cos\omega t[{}^{(n-1)}\mathcal{M}_- - {}^{(n-1)}\mathcal{M}_+] & \text{(11a)} \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^{(n)}\mathcal{M}_\pm = -\dfrac{1}{T_R}\,{}^{(n)}\mathcal{M}_\pm \pm i\omega_0\,{}^{(n)}\mathcal{M}_\pm \mp i\cos\omega t\,{}^{(n-1)}\mathcal{M}_z & \text{(11b)} \end{cases}$$

[2]In certain experiments, the pumping is "transverse" ($\boldsymbol{\mu}_0$ is perpendicular to $\mathbf{B}_0$). See exercise 1 at the end of this complement.

2-b. The Fourier series expansion of the solution

Since the only time-dependent terms on the right-hand side of (10) and (11) are sinusoidal, the steady-state solution of (10) and (11) is periodic, of period $2\pi/\omega$. We can expand it in a Fourier series:

$$\begin{cases} {}^{(n)}\mathcal{M}_z = \displaystyle\sum_{p=-\infty}^{+\infty} {}_{p}^{(n)}\mathcal{M}_z \, \mathrm{e}^{ip\omega t} & \text{(12a)} \\ {}^{(n)}\mathcal{M}_\pm = \displaystyle\sum_{p=-\infty}^{+\infty} {}_{p}^{(n)}\mathcal{M}_\pm \, \mathrm{e}^{ip\omega t} & \text{(12b)} \end{cases}$$

${}_{p}^{(n)}\mathcal{M}_z$ and ${}_{p}^{(n)}\mathcal{M}_\pm$ represent the $p\omega$ Fourier components of the nth-order solution.

By taking ${}^{(n)}\mathcal{M}_z$ real and ${}^{(n)}\mathcal{M}_+$ and ${}^{(n)}\mathcal{M}_-$ as complex conjugates of each other, we obtain the following reality conditions:

$$\begin{cases} {}_{p}^{(n)}\mathcal{M}_z = \left[{}_{-p}^{(n)}\mathcal{M}_z \right]^* & \text{(13a)} \\ {}_{p}^{(n)}\mathcal{M}_\pm = \left[{}_{-p}^{(n)}\mathcal{M}_\mp \right]^* & \text{(13b)} \end{cases}$$

Substituting (12a) and (12b) into (10) and (11), and setting equal to zero the coefficient of each exponential $\mathrm{e}^{ip\omega t}$, we find:

$n = 0$:

$$\begin{cases} {}_{0}^{(0)}\mathcal{M}_z = n\mu_0 T_R \\ {}_{p}^{(0)}\mathcal{M}_z = 0 \qquad \text{if } p \neq 0 \\ {}_{p}^{(0)}\mathcal{M}_\pm = 0 \qquad \text{for any } p \end{cases} \tag{14}$$

$n \neq 0$:

$$\begin{cases} \left(ip\omega + \dfrac{1}{T_R} \right) {}_{p}^{(n)}\mathcal{M}_z = \dfrac{i}{4} \left[{}_{p+1}^{(n-1)}\mathcal{M}_- + {}_{p-1}^{(n-1)}\mathcal{M}_- - {}_{p+1}^{(n-1)}\mathcal{M}_+ - {}_{p-1}^{(n-1)}\mathcal{M}_+ \right] & \text{(15a)} \\ \left[i(p\omega \mp \omega_0) + \dfrac{1}{T_R} \right] {}_{p}^{(n)}\mathcal{M}_\pm = \mp \dfrac{i}{2} \left[{}_{p+1}^{(n-1)}\mathcal{M}_z + {}_{p-1}^{(n-1)}\mathcal{M}_z \right] & \text{(15b)} \end{cases}$$

These algebraic equations can be solved immediately:

$$\begin{cases} {}_{p}^{(n)}\mathcal{M}_z = \dfrac{i}{4\left(ip\omega + \frac{1}{T_R} \right)} \left[{}_{p+1}^{(n-1)}\mathcal{M}_- + {}_{p-1}^{(n-1)}\mathcal{M}_- - {}_{p+1}^{(n-1)}\mathcal{M}_+ - {}_{p-1}^{(n-1)}\mathcal{M}_+ \right] & \text{(16a)} \\ {}_{p}^{(n)}\mathcal{M}_\pm = \mp \dfrac{i}{2\left[i(p\omega \mp \omega_0) + \frac{1}{T_R} \right]} \left[{}_{p+1}^{(n-1)}\mathcal{M}_z + {}_{p-1}^{(n-1)}\mathcal{M}_z \right] & \text{(16b)} \end{cases}$$

Thus, expressions (16) give the nth-order solution explicitly in terms of the $(n-1)$th-order solution. Since the zeroth-order solution is known [*cf.* equations (14)], the problem is, in theory, entirely solved.

2-c. The general structure of the solution

It is possible to arrange the various terms of the expansion of the solution in a double-entry table in which the perturbation order n labels the columns and the degree p of the harmonic $p\omega$ being considered labels the rows. To zeroth-order, only ${}^{(0)}_{0}\mathcal{M}_z$ is different from zero. By iteration, using (16), we can deduce the other non-zero higher-order terms (table I), thus obtaining a "tree-like structure". The following properties can be found directly by recurrence, using (16):

(*i*) At even perturbation orders, only the longitudinal magnetization is modified; at odd orders, only the transverse magnetization.

(*ii*) At even perturbation orders, only the even harmonics are involved; at odd orders, only the odd harmonics.

(*iii*) For each value of n, the values of p to be retained are n, $n-2$, ... $-n+2$, $-n$.

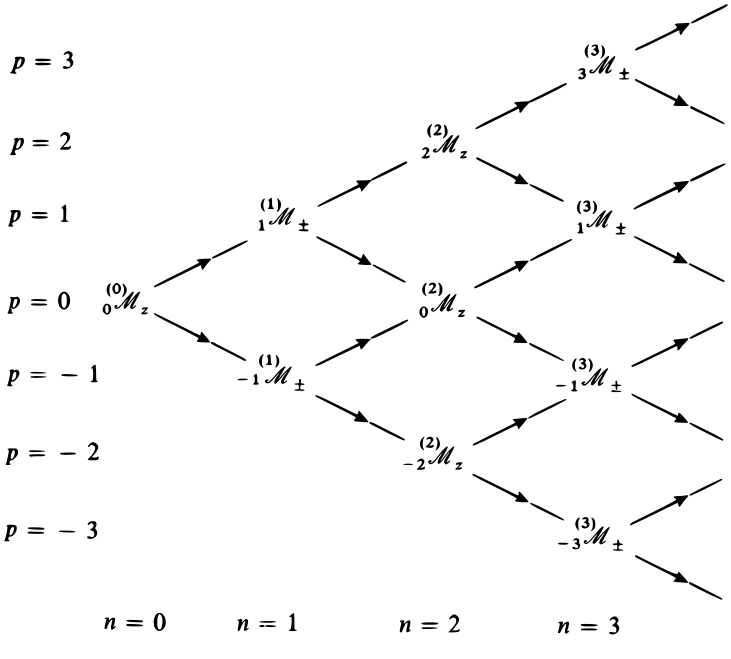


Table I: *Double-entry table indicating the $p\omega$ Fourier components of the magnetization that are non-zero to the nth perturbation order in ω_1.*

Comment:

This structure is valid only for a particular polarization of the radiofrequency field $\mathbf{B}_1(t)$ (perpendicular to $\mathbf{B}_0$). Analogous tables could be constructed for other radiofrequency polarizations.

3. Discussion

We shall now interpret the results of this calculation, through third order.

3-a. Zeroth-order solution: competition between pumping and relaxation

According to (14), the only non-zero zeroth-order component is:

$$ {}^{(0)}_{0}\mathcal{M}_z = n\mu_0 T_R \tag{17}$$

In the absence of radiofrequency fields, there is only a static longitudinal magnetization ($p = 0$). Since $\mathcal{M}_z$, is proportional to the population difference of the states $|+\rangle$ and $|-\rangle$ shown in Figure 2 (*cf.* Complement E_{IV}), it can also be said that the pumping populates these two states unequally.

The larger the number of particles entering the cell (the more efficient the pumping) and the longer T_R (the slower the relaxation), the larger ${}^{(0)}_{0}\mathcal{M}_z$. The zeroth-order solution (17) therefore describes the dynamic equilibrium resulting from competition between the pumping and relaxation processes.

From now on, in order to simplify the notation, we shall set:

$$\mathcal{M}_0 = {}^{(0)}_{0}\mathcal{M}_z \tag{18a}$$

$$\Gamma_R = \frac{1}{T_R} \tag{18b}$$

3-b. First-order solution: the linear response

To first order, only the transverse magnetization $\boldsymbol{\mathcal{M}}_\perp$ is different from zero. Since $\mathcal{M}_+ = \mathcal{M}_-^*$, it suffices to study $\mathcal{M}_+$.

α. Motion of the transverse magnetization

According to Table I, for $n = 1$, we have $p = \pm 1$. Setting $n = 1$ and $p = \pm 1$ in (16b), using (18), we get:

$$\begin{cases} {}^{(1)}_{1}\mathcal{M}_+ = \dfrac{\mathcal{M}_0}{2}\dfrac{1}{\omega_0 - \omega + i\Gamma_R} & \text{(19a)} \\ {}^{(1)}_{-1}\mathcal{M}_+ = \dfrac{\mathcal{M}_0}{2}\dfrac{1}{\omega_0 + \omega + i\Gamma_R} & \text{(19b)} \end{cases}$$

Substituting these expressions into (12b) and then into (9b), we obtain $\mathcal{M}_+$ to first order in ω_1:

$$\mathcal{M}_+ = \omega_1 \frac{\mathcal{M}_0}{2}\left[\frac{e^{i\omega t}}{\omega_0 - \omega + i\Gamma_R} + \frac{e^{-i\omega t}}{\omega_0 + \omega + i\Gamma_R}\right] \tag{20}$$

The point representing $\mathcal{M}_+$ describes the same motion in the complex plane as the projection $\mathcal{M}_\perp$ of $\mathcal{M}$ in the plane perpendicular to $\mathbf{B}_0$. According to (20), this motion results from the superposition of two circular motions with the same angular velocity, one of them right circular (the $e^{i\omega t}$ term) and the other left circular (the $e^{-i\omega t}$ term). The resulting motion, in the general case, is therefore elliptical.

β. Existence of two resonances

The right circular motion has a maximum amplitude when $\omega_0 = \omega$, and the left circular motion, when $\omega_0 = -\omega$. $\mathcal{M}_\perp$ therefore presents two resonances (while for a rotating field, there was a single resonance; see Complement F_{IV}). The interpretation of this phenomenon is as follows: the linear radiofrequency field can be broken down into a left and a right circular field, each of which induces a resonance; since the rotation directions are opposed, the static fields $\mathbf{B}_0$ for which these resonances appear are opposed.

γ. Linear susceptibility

Near a resonance ($\omega_0 \simeq \omega$, for example), we can neglect the non-resonant term in (20). We then get:

$$\mathcal{M}_+ \underset{\omega \simeq \omega_0}{\simeq} \omega_1 \frac{\mathcal{M}_0}{2} \frac{e^{i\omega t}}{\omega_0 - \omega + i\Gamma_R} \tag{21}$$

$\mathcal{M}_+$ is therefore proportional to the rotating radiofrequency field component in the direction corresponding to the resonance, $B_1\, e^{i\omega t}/2$ in this case.

The ratio of $\mathcal{M}_+$ to this component is called the linear susceptibility $\chi(\omega)$:

$$\chi(\omega) = -\gamma \mathcal{M}_0 \frac{1}{\omega_0 - \omega + i\Gamma_R} \tag{22}$$

$\chi(\omega)$ is a complex susceptibility because of the existence of a phase difference between $\boldsymbol{\mathcal{M}}_\perp$ and the rotating component of the radiofrequency field responsible for the resonance.

The square of the modulus of $\chi(\omega)$ has the classical resonant form in the neighborhood of $\omega = \omega_0$ (Fig. 3), over an interval of width:

$$\Delta\omega = 2\Gamma_R = \frac{2}{T_R} \tag{23}$$

The longer the relaxation time T_R, the sharper the resonance curve. From now on, we shall assume that the two resonances $\omega_0 = \omega$ and $\omega_0 = -\omega$ are completely separated, i.e. that:

$$\omega/\Gamma_R = \omega T_R \gg 1 \tag{24}$$

The phase difference varies from 0 to $\pm\pi$ when we pass through resonance. It is equal to $\pm\pi/2$ at resonance: it is when $\boldsymbol{\mathcal{M}}_\perp$ and the rotating component are out of phase by $\pi/2$ that the work of the couple exerted by the field on $\boldsymbol{\mathcal{M}}$ is maximal. The sign of this work depends on the sign of $\mathcal{M}_0$, that is, on that of μ_0: it depends on whether the spin states of the entering particles are $|+\rangle$ or $|-\rangle$. In one case (spins entering in the lower level), the work is furnished by the field, and energy is transferred from the field to the spins (absorption). In the opposite case (particles entering in the higher level), the work is negative, and energy is transferred from the spins to the field (induced emission). The latter situation occurs in atomic amplifiers and oscillators (masers and lasers).

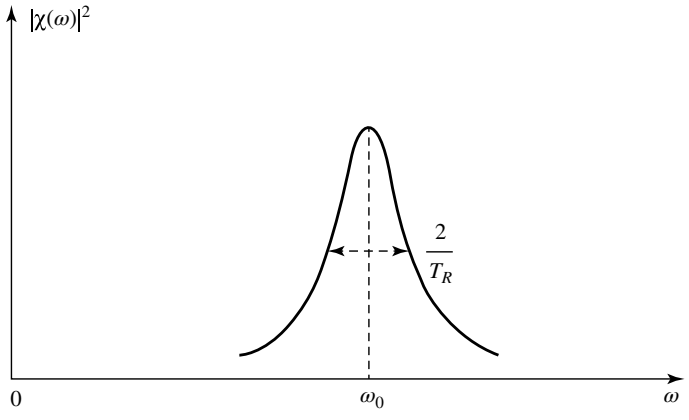


Figure 3: Variation of the square of the modulus $|\chi(\omega)|^2$ of the linear susceptibilite of the spin system, with respect to ω. A resonance appears, of width $2/T_R$, centered at $\omega = \omega_0$.

3-c. Second-order solution: absorption and induced emission

To second order, according to Table I, only ${}^{(2)}_{0}\mathcal{M}_z$ and ${}^{(2)}_{\pm 2}\mathcal{M}_z$ are non-zero. First, we shall study ${}^{(2)}_{0}\mathcal{M}_z$, that is, the static population difference of the states $|+\rangle$ and $|-\rangle$ to second order. We shall then consider ${}^{(2)}_{\pm 2}\mathcal{M}_z$, that is, the generation of the second harmonic.

α. Variation of the population difference of the two states of the system

${}^{(2)}_{0}\mathcal{M}_0$ corrects the zeroth-order result obtained for ${}^{(0)}_{0}\mathcal{M}_0$. According to (16a) and (13b):

$$
\begin{aligned}
{}^{(2)}_{0}\mathcal{M}_z &= \frac{i}{4\Gamma_R}\left[{}^{(1)}_{1}\mathcal{M}_- + {}^{(1)}_{-1}\mathcal{M}_- - {}^{(1)}_{1}\mathcal{M}_+ - {}^{(1)}_{-1}\mathcal{M}_+\right] \\
&= \frac{i}{4\Gamma_R}\left[{}^{(1)}_{-1}\mathcal{M}^*_+ + {}^{(1)}_{1}\mathcal{M}^*_+ - {}^{(1)}_{1}\mathcal{M}_+ - {}^{(1)}_{-1}\mathcal{M}_+\right] \qquad (25)
\end{aligned}
$$

which, according to first-order solutions (19a) and (19b), yields:

$$
{}^{(2)}_{0}\mathcal{M}_z = -\frac{\mathcal{M}_0}{4}\left[\frac{1}{(\omega-\omega_0)^2+\Gamma_R^2} + \frac{1}{(\omega+\omega_0)^2+\Gamma_R^2}\right] \qquad (26)
$$

Grouping the static terms ($p = 0$) through second order in (9a), we get:

$$
\mathcal{M}_z(\text{static}) = \mathcal{M}_0\left\{1 - \frac{\omega_1^2}{4}\left[\frac{1}{(\omega-\omega_0)^2+\Gamma_R^2} + \frac{1}{(\omega+\omega_0)^2+\Gamma_R^2}\right] + \ldots\right\} \qquad (27)
$$

Figure 4 represents this static longitudinal magnetization as a function of ω_0.

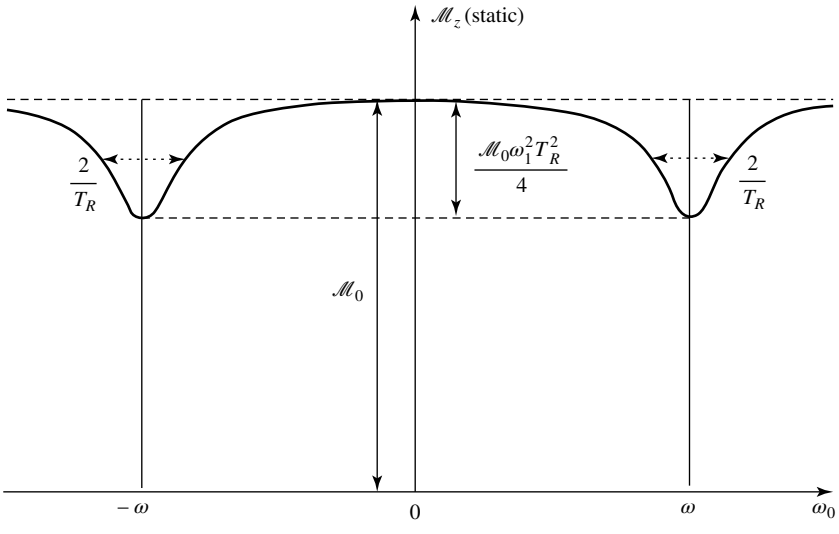


Figure 4: Variation of the static longitudinal magnetization with respect to ω_0. To second order in the perturbation treatment, there appear two resonances of width $2/T_R$, centered at $\omega_0 = \omega$ and $\omega_0 = -\omega$. The calculation is valid only if the relative intensity of the resonances is small, that is, if $\omega_1 T_R \ll 1$.

The population difference is therefore always *decreased*, to second order, relative to its value in the absence of radiofrequency, and the decrease is proportional to the *intensity* of the radiofrequency field. This is simple to understand: under the effect of the incident field, transitions are induced from $|+\rangle$ to $|-\rangle$ (induced emission) or from $|-\rangle$ to $|+\rangle$ (absorption); whatever the sign of the initial population difference, the transitions from the more populated state are the more numerous, so that they decrease the population difference.

Comment:

The maximum value of $\omega_1^2 \,|_0^{(2)}\mathcal{M}_z|$ is $\mathcal{M}_0\omega_1^2/4\Gamma_R^2 = \mathcal{M}_0\omega_1^2T_R^2/4$ (the resonance amplitude which appears as a dip in Figure 4). For the perturbation expansion to make sense, it is therefore necessary that:

$$\omega_1 T_R \ll 1 \tag{28}$$

β. Generation of the second harmonic

According to (16a), (13b), (19a) and (19b):

$$
\begin{aligned}
{}^{(2)}_{2}\mathcal{M}_z &= \frac{1}{4(2\omega - i\Gamma_R)} \left[{}^{(1)}_{-1}\mathcal{M}^*_+ - {}^{(1)}_{1}\mathcal{M}_+ \right] \\
&= \frac{\mathcal{M}_0}{8(2\omega - i\Gamma_R)} \left[\frac{1}{\omega_0 + \omega - i\Gamma_R} - \frac{1}{\omega_0 - \omega + i\Gamma_R} \right]
\end{aligned} \tag{29}
$$

${}^{(2)}_{2}\mathcal{M}_z$ describes a vibration of the magnetic dipole along Oz at the angular frequency 2ω. The system can therefore radiate a wave of angular frequency 2ω, polarized (as far as the magnetic field is concerned) linearly along Oz.

Thus, we see that an atomic system is not generally a linear system. It can double the excitation frequency, triple it (as we shall see later), etc. The same type of phenomenon exists in optics for very high intensities ("non-linear optics"): a red laser beam (produced, for example, by a ruby laser) falling on a material such as a quartz crystal can give rise to an ultraviolet light beam (doubled frequency).

Comment:

It will prove useful to compare $|{}^{(2)}_{0}\mathcal{M}_z|$ and $|{}^{(2)}_{2}\mathcal{M}_z|$ in the neighborhood of $\omega_0 = \omega$. According to (29), for $\omega \simeq \omega_0$, we have:

$$
|{}^{(2)}_{2}\mathcal{M}_z| \simeq \frac{\mathcal{M}_0}{16\omega_0\Gamma_R} \tag{30}
$$

Similarly, (26) indicates that:

$$
|{}^{(2)}_{0}\mathcal{M}_z| \simeq \frac{\mathcal{M}_0}{4\Gamma_R^2} \tag{31}
$$

Therefore, for $\omega \simeq \omega_0$:

$$
\frac{|{}^{(2)}_{2}\mathcal{M}_z|}{|{}^{(2)}_{0}\mathcal{M}_z|} \simeq \frac{\Gamma_R}{4\omega_0} = \frac{1}{4\omega_0 T_R} \ll 1 \tag{32}
$$

according to (24).

3-d. Third-order solution: saturation effects and multiple-quanta transitions

To third order, Table I shows that only ${}^{(3)}_{\pm 1}\mathcal{M}_\pm$ and ${}^{(3)}_{\pm 3}\mathcal{M}_\pm$ are non-zero; it suffices to study ${}^{(3)}\mathcal{M}_+$.

${}^{(3)}_{1}\mathcal{M}_+$ corrects to third order the right circular motion of $\boldsymbol{\mathcal{M}}_\perp$, found to first order and analyzed in § 3-b above. We shall see that ${}^{(3)}_{1}\mathcal{M}_+$ corresponds to a saturation effect in the susceptibility of the system.

${}^{(3)}_{3}\mathcal{M}_+$ represents a new component of the motion of $\boldsymbol{\mathcal{M}}_\perp$, of angular frequency 3ω of the motion of $\boldsymbol{\mathcal{M}}_\perp$ (generation of the third harmonic). Moreover, the resonant nature of ${}^{(3)}_{3}\mathcal{M}_+$ in the neighborhood of $\omega_0 = 3\omega$ can be interpreted as resulting from the simultaneous absorption of three radiofrequency photons, a process which conserves both the total energy and the total angular momentum.

α. Saturation of the susceptibility of the system

According to (16b):

$$ {}^{(3)}_{1}\mathcal{M}_{+} = \frac{1}{2}\frac{1}{\omega_0 - \omega + i\Gamma_R}\left[{}^{(2)}_{2}\mathcal{M}_z + {}^{(2)}_{0}\mathcal{M}_z\right] \tag{33} $$

Since we are interested in the correction to the right circular motion discussed in § 3.b, which is resonant at $\omega = \omega_0$, we shall place ourselves in the neighborhood of $\omega_0 = \omega$. We can then, according to the comment in the preceding section [*cf.* formula (32], neglect ${}^{(2)}_{2}\mathcal{M}_z$ compared to ${}^{(2)}_{0}\mathcal{M}_z$. Thus we obtain, using expression (26) for ${}^{(2)}_{0}\boldsymbol{\mathcal{M}}_z$ (neglecting the term whose resonance peak is at $\omega_0 = -\omega$):

$$ {}^{(3)}_{1}\mathcal{M}_{+} \simeq -\frac{\mathcal{M}_0}{8}\ \frac{1}{\omega_0 - \omega + i\Gamma_R}\ \frac{1}{(\omega - \omega_0)^2 + \Gamma_R^2} \tag{34} $$

If we regroup results (34) and (19a), we find the expression for the right circular motion of $\mathcal{M}_+$ at the frequency $\omega/2\pi$, to third order in ω_1:

$$ \mathcal{M}_{+}(\text{right circular}) = \omega_1 \frac{\mathcal{M}_0}{2}\frac{\mathrm{e}^{i\omega t}}{\omega_0 - \omega + i\Gamma_R}\left[1 - \frac{\omega_1^2}{4}\frac{1}{(\omega - \omega_0)^2 + \Gamma_R^2}\right] \tag{35} $$

Comparing (35) and (21), we see that the susceptibility of the system goes from value (22) to the value:

$$ \chi(\omega) = -\gamma\mathcal{M}_0\frac{1}{\omega_0 - \omega + i\Gamma_R}\left[1 - \frac{\omega_1^2}{4}\frac{1}{(\omega_0 - \omega)^2 + \Gamma_R^2}\right] \tag{36} $$

It is therefore multiplied by a factor smaller than one; the greater the intensity of the radiofrequency field and the nearer we are to resonance, the smaller the factor. The system is then said to be "saturated". The ω_1^2 term of (36) is called the "non-linear susceptibility".

The physical meaning of this saturation is very clear. A weak electromagnetic field induces in the atomic system a dipole moment which is proportional to it. If the field amplitude is increased, the dipole cannot continue to increase proportionally to the field. The absorption and emission transitions induced by the field decrease the population difference of the atomic states involved. Consequently, the atomic system responds less and less to the field. Furthermore, we see that the term in brackets in (36) is none other than the term that expresses the decrease in the population difference to second order [*cf.* formula (27), in which the term resonant at $\omega_0 = -\omega$ was neglected].

Comment:

The saturation terms play a very important role in all maser or laser theories. Consider Figure 1 again. If we keep only the linear response term in the first step of the calculation (arrow directed to the right), the induced dipole moment is proportional to the field. If the material amplifies (and if the losses of the electromagnetic cavity are sufficiently small), the reaction of the dipole on the field (arrow directed to the left) tends to increase the

field by a quantity proportional to it. Thus, we obtain for the field a linear differential equation which leads to a solution which increases linearly with time.

It is the saturation terms that prevent this unlimited increase. They lead to an equation whose solution remains bounded and approaches a limit which is the steady-state laser field in the cavity. Physically, these saturation terms express the fact that the atomic system cannot furnish the field with an energy greater than that corresponding to the population difference initially introduced by the pumping.

β. Three-photon transitions

According to (16b) and (29):

$$
\begin{aligned}
{}^{(3)}_{3}\mathcal{M}_{+} &= \frac{1}{2}\frac{1}{\omega_0 - 3\omega + i\Gamma_R}\,{}^{(2)}_{2}\mathcal{M}_{z} \\
&= \frac{\mathcal{M}_0}{16}\ \frac{1}{\omega_0 - 3\omega + i\Gamma_R}\ \frac{1}{2\omega - i\Gamma_R}\left[\frac{1}{\omega_0 + \omega - i\Gamma_R} - \frac{1}{\omega_0 - \omega + i\Gamma_R}\right]
\end{aligned}
\tag{37}
$$

With respect to the term ${}^{(3)}_{3}\mathcal{M}_{+}$, we could make the same comment as for ${}^{(2)}_{2}\mathcal{M}_{z}$: the atomic system produces harmonics of the excitation frequency (here, the third harmonic).

The difference with the discussion of the preceding section relative to ${}^{(2)}_{2}\mathcal{M}_{z}$ is the appearance of a resonance centered at $\omega_0 = 3\omega$ [due to the first resonant denominator of (37)].

We can give a particle interpretation of the $\omega_0 = \omega$ resonance discussed in the preceding sections: the spin goes from the state $|-\rangle$ to the state $|+\rangle$ by absorbing a photon (or emitting it, depending on the relative positions of the $|+\rangle$ and $|-\rangle$ states). There is resonance when the energy $\hbar\omega$ of the photon is equal to the energy $\hbar\omega_0$ of the atomic transition. We could give an analogous particle interpretation of the $\omega_0 = 3\omega$ resonance. Since $\hbar\omega_0 = 3\hbar\omega$, the transition necessarily involves three photons, since the total energy must be conserved.

We may wonder why no resonance has appeared to second order for $\hbar\omega_0 = 2\hbar\omega$ (two-photon transition). The reason is that the total angular momentum must also be conserved during the transition. The linear radiofrequency field is, as we have already said, a superposition of two fields rotating in opposite directions. With each of these rotating fields are associated photons of a different type. For the right circular field, it is σ^+ photons, transporting an angular momentum $+\hbar$ relative to Oz. For the left circular field, it is σ^- photons, transporting an angular momentum $-\hbar$. To go from the $|-\rangle$ state to the $|+\rangle$ state, the spin must absorb an angular momentum $+\hbar$ relative to Oz (the difference between the two eigenvalues of S_z). It can do so by absorbing a σ^+ photon; if $\omega_0 = \omega$, there is also conservation of the total energy, which explains the appearance of the $\omega_0 = \omega$ resonance. The system can also acquire an angular momentum $+\hbar$ by absorbing three photons (Fig. 5): two σ^+ photons and one σ^- photon. Therefore, if $\omega_0 = 3\omega$, both energy and total angular momentum can be conserved, which explains the $\omega_0 = 3\omega$ resonance. On the other hand, two photons can never give the atom an angular momentum $+\hbar$: either both photons are σ^+ and they carry $2\hbar$, or they are both σ^- and they carry $-2\hbar$, or one is σ^+ and one is σ^- and they carry no total angular momentum.

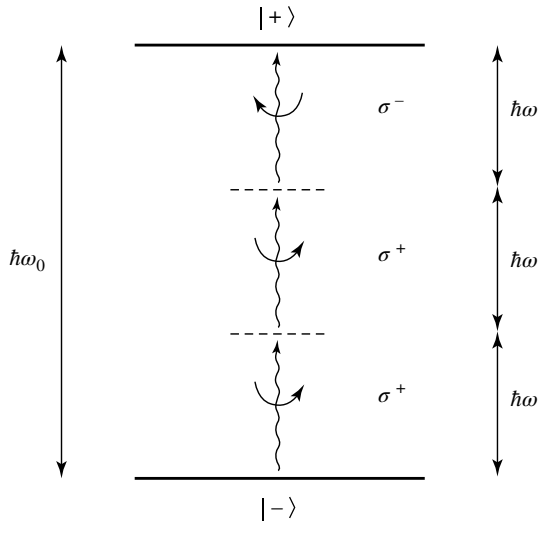


Figure 5: The spin can go from the $|-\rangle$ state to the $|+\rangle$ state by absorbing three photons of energy $\hbar\omega$. The total energy is conserved if $\hbar\omega_0 = 3\hbar\omega$. The angular momentum is conserved if two photons have a σ^+ polarization (each carries an angular momentum $+\hbar$ relative to Oz) and the third has a σ^- polarization (it carries an angular momentum $-\hbar$).

These arguments can easily be generalized and enable us to show that resonances appear when $\omega_0 = \omega$, 3ω, 5ω, 7ω, ..., $(2n+1)\omega$, ..., corresponding to the absorption of an odd number of photons. Furthermore, we see from formula (16b) that ${}^{(2n+1)}_{2n+1}\mathcal{M}_+$ gives rise to a resonance peak for $\omega_0 = (2n+1)\omega$. Nothing analogous occurs at even orders since, according to Table I, we must then use equation (16a).

Comments:

(*i*) If the field $\mathbf{B}_1$ is rotating, there is only one type of photon, σ^+ or σ^-. The same argument shows that a *single resonance* can then occur, at $\omega_0 = \omega$ if the photons are σ^+ and at $\omega_0 = -\omega$ if they are σ^-. This enables us to understand why the calculations are much simpler for a rotating field and lead to an exact solution. It is instructive to apply the method of this complement to the case of a rotating field and to show that the perturbation series can be summed to give the solution found directly in Complement $\mathrm{F_{IV}}$.

(*ii*) Consider a system having two levels of different parities, subject to the influence of an oscillating electric field. The interaction Hamiltonian then has the same structure as the one we are studying in this complement: S_x has only non-diagonal elements. Similarly, the electric dipole Hamiltonian, since it is odd, can have no diagonal elements. In the second case, the calculations

are very similar to the preceding ones and lead to analogous conclusions: resonances are found for $\omega_0 = \omega$, 3ω, 5ω, ... The interpretation of the "odd" nature of the spectrum is then as follows: the electric dipole photons have a negative parity, and the system must absorb an odd number of them in order to move from one level to another of different parity.

(*iii*) For the spin 1/2 case, assume that the linear radiofrequency field is neither parallel nor perpendicular to $\mathbf{B}_0$ (Fig. 6). $\mathbf{B}_1$ can then be broken down into a component parallel to $\mathbf{B}_0$, $\mathbf{B}_{1\parallel}$, with which are associated π photons (with zero angular momentum relative to Oz), and a component $\mathbf{B}_{1\perp}$, with which, as we have seen, σ^+ and σ^- photons are associated. In this case, the atom can increase its angular momentum relative to Oz by $+\hbar$, and move from $|-\rangle$ to $|+\rangle$, by absorbing two photons, one σ^+ and the other π. It can be shown, by applying the method of this complement, that for this polarization of the radiofrequency, a complete (even and odd) spectrum of resonances appears: $\omega_0 = \omega$, 2ω, 3ω, 4ω, ...

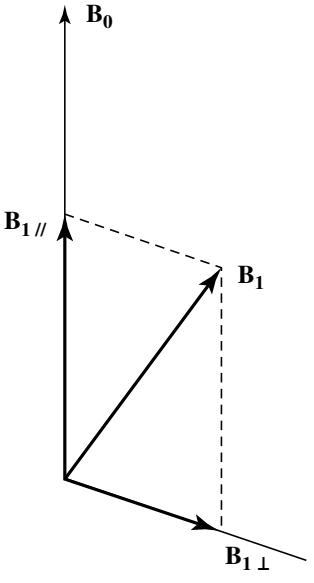


Figure 6: The static magnetic field $\mathbf{B}_0$ *and the radiofrequency field* $\mathbf{B}_1$*, in the case in which* $\mathbf{B}_1$ *is neither parallel nor perpendicular to* $\mathbf{B}_0$*.* $\mathbf{B}_{1\parallel}$ *and* $\mathbf{B}_{1\perp}$ *are the components of* $\mathbf{B}_1$ *parallel and perpendicular to* $\mathbf{B}_0$*.*

4. Exercises: applications of this complement

EXERCISE 1

In equations (1), set $\omega_1 = 0$ (no radiofrequency) and choose $\boldsymbol{\mu}_0$ parallel to Ox (transverse pumping).

Calculate the steady-state values of $\mathcal{M}_x$, $\mathcal{M}_y$ and $\mathcal{M}_z$. Show that $\mathcal{M}_x$ and $\mathcal{M}_y$ undergo resonant variations when the static field is swept about zero (the Hanle effect). Give a physical interpretation of these resonances (pumping in competition with Larmor precession) and show that they permit the measurement of the product γT_R.

EXERCISE 2

Consider a spin system subjected to the same static field $\mathbf{B}_0$ and to the same pumping and relaxation processes as in this complement. These spins are also subjected to two linear radiofrequency fields, the first one of angular frequency ω and amplitude B_1, parallel to Oz, and the second one of angular frequency ω' and amplitude B_1', parallel to Ox.

Using the general methods described in this complement, calculate the magnetization $\boldsymbol{\mathcal{M}}$ of the spin system to second order in $\omega_1 = -\gamma B_1$ and $\omega_1' = -\gamma B_1'$ (terms in ω_1^2, $\omega_1'^2$, $\omega_1\omega_1'$). We fix $\omega_0 = -\gamma B_0$ and ω_1. Assume $\omega_0 > \omega$, and let ω' vary. Show that, to this perturbation order, two resonances appear, one at $\omega' = \omega_0 - \omega$ and the other at $\omega' = \omega_0 + \omega$.

Give a physical interpretation of these two resonances (the first one corresponds to a two-photon absorption, and the second, to a Raman effect).

References and suggestions for further reading:

See section 15 of the bibliography.

Semiclassical theories of masers and lasers: Lamb (15.4) and (15.2), Sargent et al. (15.5), Chap. VIII, IX and X.

Non-linear optics: Baldwin (15.19), Bloembergen (15.21), Giordmaine (15.22).

Iterative solution of the master equation: Bloembergen (15.21), Chap. 2, §§ 3, 4 and 5 and Appendix III.

Multiphoton processes in R. F. range, Hanle effect: Brossel's lectures in (15.2).

Complement C_{XIII}

Oscillations of a system between two discrete states under the effect of a sinusoidal resonant perturbation

1 The method: secular approximation 1374
2 Solution of the system of equations 1375
3 Discussion 1376

The approximation method used to calculate the effect of a resonant perturbation in Chapter XIII is not valid over long periods of time. We have seen [*cf.* condition (C-18) of this chapter] that t must satisfy:

$$t \ll \frac{\hbar}{|W_{fi}|} \tag{1}$$

Suppose we want to study the behavior of a system subjected to a resonant perturbation over a considerable time [for which condition (1) is not satisfied]. Since the first-order solution is then insufficient, we could try to calculate a certain number of higher-order terms to obtain a better expression for $\mathscr{P}_{if}(t;\omega)$:

$$\mathscr{P}_{if}(t;\omega) = \left|\lambda b_f^{(1)}(t) + \lambda^2 b_f^{(2)}(t) + \lambda^3 b_f^{(3)}(t) + ...\right|^2 \tag{2}$$

Such a method would lead to unnecessarily long calculations.

We shall see here that it is possible to solve the problem more elegantly and rapidly by fitting the approximation method to the resonant nature of the perturbation. The resonance condition $\omega \simeq \omega_{fi}$ implies that only the two discrete states $|\varphi_i\rangle$ and $|\varphi_f\rangle$ are effectively coupled by $W(t)$. Since the system, at the initial instant, is in the state $|\varphi_i\rangle$ [$b_i(0) = 1$], the probability amplitude $b_f(t)$ of finding it in the state $|\varphi_f\rangle$ at time t can be appreciable. On the other hand, all the coefficients $b_n(t)$ (with $n \neq i,\ f$) remain much smaller than 1 since they do not satisfy the resonance condition. This is the basis of the method we shall use.

1. The method: secular approximation

In Chapter XIII, we replaced all the components $b_k(t)$ on the right-hand side of (B-11) by their values $b_k(0)$ at time $t = 0$. Here, we shall do the same thing for the components for which $k \neq i,\ f$. However, we shall explicitly keep $b_i(t)$ and $b_f(t)$. Thus, in order to determine $b_i(t)$ and $b_f(t)$, we are led to the system of equations [the perturbation having the form (C-1a) of Chap. XIII]:

$$\begin{aligned}
i\hbar\frac{\mathrm{d}}{\mathrm{d}t}b_i(t) &= \frac{1}{2i}\left\{\left[\mathrm{e}^{i\omega t} - \mathrm{e}^{-i\omega t}\right] W_{ii} b_i(t) + \left[\mathrm{e}^{i(\omega-\omega_{fi})t} - \mathrm{e}^{-i(\omega+\omega_{fi})t}\right] W_{if} b_f(t)\right\} \\
i\hbar\frac{\mathrm{d}}{\mathrm{d}t}b_f(t) &= \frac{1}{2i}\left\{\left[\mathrm{e}^{i(\omega+\omega_{fi})t} - \mathrm{e}^{-i(\omega-\omega_{fi})t}\right] W_{fi} b_i(t) + \left[\mathrm{e}^{i\omega t} - \mathrm{e}^{-i\omega t}\right] W_{ff} b_f(t)\right\}
\end{aligned} \tag{3}$$

On the right-hand side of these equations, certain coefficients of $b_i(t)$ and $b_f(t)$ are proportional to $\mathrm{e}^{\pm i(\omega-\omega_{fi})t}$, so they oscillate slowly in time when $\omega \simeq \omega_{fi}$.

On the other hand, the coefficients proportional either to $\mathrm{e}^{\pm i\omega t}$ or to $\mathrm{e}^{\pm i(\omega+\omega_{fi})t}$ oscillate much more rapidly. Here, we shall use the *secular approximation*, which consists in neglecting the second type of terms. The remaining ones, called "secular terms", are then those whose coefficients reduce to constants for $\omega = \omega_{fi}$. When integrated over time, they make significant contributions to the variations of the components $b_i(t)$ and $b_f(t)$. On the other hand, the contribution of the other terms is negligible, since their variation is too rapid (the integration of $\mathrm{e}^{i\Omega t}$ causes a factor $1/\Omega$ to appear, and the average value of $\mathrm{e}^{i\Omega t}$ over a large number of periods is practically zero).

Comment:

For the preceding argument to be valid, it is necessary for the temporal variation of a term $\mathrm{e}^{i\omega t}b_{i,\,f}(t)$ to be due principally to the exponential, and not to the component $b_{i,\,f}(t)$. Since ω is very close to ω_{fi}, this means that $b_{i,\,f}(t)$ must not significantly vary over a time interval of the order of $1/|\omega_{fi}|$. This is indeed true with the assumptions we have made, that is, with $W \ll H_0$. The variations of $b_i(t)$ and $b_f(t)$ (which are constants if $W = 0$) are due to the presence of the perturbation W, and are appreciable for times of the order of $\hbar/|W_{if}|$ [this can be verified directly from formulas (8), obtained below]. Since by hypothesis $|W_{if}| \ll \hbar|\omega_{fi}|$, this time is much greater than $1/|\omega_{fi}|$.

In conclusion, the secular approximation leads to the system of equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}b_i(t) = -\frac{1}{2\hbar}\mathrm{e}^{i(\omega-\omega_{fi})t}\,W_{if}b_f(t) \tag{4a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}b_f(t) = \frac{1}{2\hbar}\mathrm{e}^{-i(\omega-\omega_{fi})t}\,W_{fi}b_i(t) \tag{4b}$$

whose solution, very close to that of system (3), is easier to calculate, as we shall see in the next section.

2. Solution of the system of equations

We shall begin by considering the case for which $\omega = \omega_{fi}$. Differentiating (4a) and substituting (4b) into the result, we obtain:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}b_i(t) = -\frac{1}{4\hbar^2}|W_{fi}|^2 b_i(t) \tag{5}$$

Since the system is in the state $|\varphi_i\rangle$ at time $t = 0$, the initial conditions are:

$$b_i(0) = 1 \tag{6a}$$

$$b_f(0) = 0 \tag{6b}$$

which, according to (4), gives:

$$\frac{\mathrm{d}b_i}{\mathrm{d}t}(0) = 0 \tag{7a}$$

$$\frac{\mathrm{d}b_f}{\mathrm{d}t}(0) = \frac{W_{fi}}{2\hbar} \tag{7b}$$

The solution of (5) that satisfies (6a) and (7a) can be written:

$$b_i(t) = \cos\left(\frac{|W_{fi}|t}{2\hbar}\right) \tag{8a}$$

We can then calculate b_f from (4a):

$$b_f(t) = \mathrm{e}^{i\alpha_{fi}} \sin\left(\frac{|W_{fi}|t}{2\hbar}\right) \tag{8b}$$

where α_{fi} is the argument of W_{fi}. The probability $\mathscr{P}_{if}(t;\omega=\omega_{fi})$ of finding the system in the state $|\varphi_f\rangle$ at time t is therefore, in this case, equal to:

$$\mathscr{P}_{if}(t;\omega=\omega_{fi}) = \sin^2\left(\frac{|W_{fi}|t}{2\hbar}\right) \tag{9}$$

When ω is different from ω_{fi} (while remaining close to the resonance value), the differential system (4) is still exactly soluble. In fact, it is completely analogous to the one we obtained in Complement F_{IV} [*cf.* equation (15)] in studying the magnetic resonance of a spin 1/2. The same type of calculation as in that complement leads to the analogue of relation (27) (Rabi's formula), which can be written here:

$$\mathscr{P}_{if}(t;\omega) = \frac{|W_{if}|^2}{|W_{if}|^2 + \hbar^2(\omega-\omega_{fi})^2} \sin^2\left[\sqrt{\frac{|W_{fi}|^2}{\hbar^2} + (\omega-\omega_{fi})^2}\,\frac{t}{2}\right] \tag{10}$$

[when $\omega = \omega_{fi}$, this expression does reduce to (9)].

3. Discussion

The discussion of the result obtained in (10) is the same as that of the magnetic resonance of a spin 1/2 (*cf.* Complement F_{IV}, § 2-c). The probability $\mathscr{P}_{if}(t;\omega)$ is an oscillating function of time; for certain values of t, $\mathscr{P}_{if}(t;\omega) = 0$, and the system has gone back into the initial state $|\varphi_i\rangle$.

Furthermore, equation (10) measures the magnitude of the resonance phenomenon. When $\omega = \omega_{fi}$, however small the perturbation is, it can cause the system to move completely from the state $|\varphi_i\rangle$ to the state $|\varphi_f\rangle$[1]. On the other hand, if the perturbation is not resonant, the probability $\mathscr{P}_{if}(t;\omega)$ always remains less than 1.

Finally, it is interesting to compare the result obtained in this complement with the one obtained using the first-order theory in Chapter XIII. First of all, note that,

[1] The magnitude of the perturbation, characterized by $|W_{fi}|$, enters, at resonance, only into the time taken by the system to move from $|\varphi_i\rangle$ to $|\varphi_f\rangle$. The smaller $|W_{fi}|$, the longer the time.

for all values of t, the probability $\mathscr{P}_{if}(t;\omega)$ obtained in (10) is included between 0 and 1. The approximation method used here therefore enables us to avoid the difficulties encountered in Chapter XIII (*cf.* § C-2-c-β). When we let t approach zero in (9), we get (C-17) of this chapter. Thus, first-order perturbation theory is indeed valid for t sufficiently small (*cf.* comment of § B-3-b). It amounts to replacing the sinusoid which represents $\mathscr{P}_{if}(t;\omega)$ as a function of time by a parabola.

Complement D_{XIII}

Decay of a discrete state resonantly coupled to a continuum of final states

1 Statement of the problem 1378
2 Description of the model 1379
2-a Assumptions about the unperturbed Hamiltonian H_0 1379
2-b Assumptions about the coupling W 1380
2-c Results of first-order perturbation theory 1380
2-d Integrodifferential equation equivalent to the Schrödinger equation 1382
3 Short-time approximation. Relation to first-order perturbation theory 1383
4 Another approximate method for solving the Schrödinger equation 1385
5 Discussion 1386
5-a Lifetime of the discrete state 1386
5-b Shift of the discrete state due to the coupling with the continuum 1387
5-c Energy distribution of the final states 1388

1. Statement of the problem

In § C-3 of Chapter XIII, we showed that the coupling induced by a constant perturbation between an initial discrete state of energy E_i and a continuum of final states (some of which have an energy equal to E_i) causes the system to go from the initial state to this continuum of final states. More precisely, the probability of finding the system in a well-defined group of states of the continuum at time t increases linearly with time. Consequently, the probability $\mathscr{P}_{ii}(t)$ of finding the system in the initial state $|\varphi_i\rangle$ at time t must decrease linearly over time from the value $\mathscr{P}_{ii}(0) = 1$. It is clear that this result is valid only over short times, since extrapolation of the linear decrease of $\mathscr{P}_{ii}(t)$ to long times would lead to negative values of $\mathscr{P}_{ii}(t)$, which would be absurd for a probability. This raises the problem of determining the long-time behavior of the system.

We encountered an analogous problem when we studied the resonant transitions induced by a sinusoidal perturbation between two discrete states $|\varphi_i\rangle$ and $|\varphi_f\rangle$. First-order perturbation theory predicts a decrease proportional to t^2 of $\mathscr{P}_{ii}(t)$ from the initial value $\mathscr{P}_{ii}(0) = 1$. The method presented in Complement C_{XIII} shows that the system actually oscillates between the states $|\varphi_i\rangle$ and $|\varphi_f\rangle$. The decrease with t^2 found in § C of Chapter XIII merely represents the "beginning" of the corresponding sinusoid.

We might expect an analogous result in the problem with which we are concerned here (oscillations of the system between the discrete state and the continuum). We shall show that this is not the case: *the physical system leaves the state* $|\varphi_i\rangle$ *irreversibly.* We find an exponential decrease $e^{-\Gamma t}$ for $\mathscr{P}_{ii}(t)$ (for which the perturbation treatment gives

only the short-time behavior $1 - \Gamma t$). Thus, the *continuous* nature of the set of final states causes the reversibility found in Complement C_{XIII} to disappear; it is responsible for a *decay* of the initial state, which thus acquires a *finite lifetime* (unstable state; *cf.* Complement K_{III}).

The situation envisaged in the present complement is very frequently encountered in physics. For example, a system, initially in a discrete state, can split, under the effect of an internal coupling (described, consequently, by a time-independent Hamiltonian W), into two distinct parts whose energies (kinetic in the case of material particles and electromagnetic in the case of photons) can have, theoretically, any value; this gives the set of final states a continuous nature. Thus, in *α-decay*, a nucleus initially in a discrete state is transformed (via the tunnel effect) into a system composed of an α-particle and another nucleus. A many-electron atom A initially in a configuration (*cf.* Complements A_{XIV} and B_{XIV}) in which several electrons are excited can, under the effect of electrostatic interactions between electrons, give rise to a system formed of an ion A^+ and a free electron (the energy of the initial configuration must, of course, be greater than the simple ionization limit of A): this is the "*autoionization*" phenomenon. We can also cite the *spontaneous emission* of a photon by an excited atomic (or nuclear) state: the interaction of the atom with the quantized electromagnetic field couples the discrete initial state (the excited atom in the absence of photons) with a continuum of final states (the atom in a lower state in the presence of a photon of arbitrary direction, polarization and energy). Finally, we can mention the photoelectric effect, in which a perturbation, now sinusoidal, couples a discrete state of an atom A to a continuum of final states (the ion A^+ and the photoelectron e^-).

These few examples of unstable states taken from various domains of physics are sufficient to underline the importance of the problem we are treating in this complement.

2. Description of the model

2-a. Assumptions about the unperturbed Hamiltonian H_0

To simplify the calculations as much as possible, we shall make the following assumptions about the spectrum of the unperturbed Hamiltonian H_0. This spectrum includes:

(i) a discrete state $|\varphi_i\rangle$ of (non-degenerate) energy E_i:

$$H_0|\varphi_i\rangle = E_i|\varphi_i\rangle \tag{1}$$

(ii) a set of states $|\alpha\rangle$ forming a continuum:

$$H_0|\alpha\rangle = E|\alpha\rangle \tag{2}$$

E can take on a continuous infinity of values, distributed over a portion of the real axis *including* E_i. We shall assume, for example, that E varies from 0 to $+\infty$:

$$E \geq 0 \tag{3}$$

Each state $|\alpha\rangle$ is characterized by its energy E and a set of other parameters which we shall denote by β (as in § C-3-a-β of Chapter XIII). $|\alpha\rangle$ can therefore also be written

in the form $|\beta, E\rangle$. We have [*cf.* formula (C-28) of Chap. XIII]:

$$\mathrm{d}\alpha = \rho(\beta, E)\,\mathrm{d}\beta\,\mathrm{d}E \tag{4}$$

where $\rho(\beta, E)$ is the density of final states.

The eigenstates of H_0 satisfy the following relations (orthogonality and closure relations):

$$\begin{cases} \langle\varphi_i|\varphi_i\rangle = 1 & \text{(5a)} \\ \langle\varphi_i|\alpha\rangle = 0 & \text{(5b)} \\ \langle\alpha|\alpha'\rangle = \delta(\alpha - \alpha') & \text{(5c)} \end{cases}$$

$$|\varphi_i\rangle\langle\varphi_i| + \int \mathrm{d}\alpha\ |\alpha\rangle\langle\alpha| = 1 \tag{6}$$

2-b. Assumptions about the coupling W

We shall assume that W is not explicitly time-dependent and has no diagonal elements:

$$\begin{aligned} &\langle\varphi_i|W|\varphi_i\rangle = 0 \\ &\langle\alpha|W|\alpha\rangle = 0 \end{aligned} \tag{7}$$

(if these diagonal elements were not zero, we could always add them to those of H_0, which would simply amount to changing the unperturbed energies). Similarly, we shall assume that W cannot couple two states of the continuum:

$$\langle\alpha|W|\alpha'\rangle = 0 \tag{8}$$

The only non-zero matrix elements of W are then those connecting the discrete state $|\varphi_i\rangle$ with the states of the continuum. It is these matrix elements, $\langle\alpha|W|\varphi_i\rangle$, that are responsible for the decay of the state $|\varphi_i\rangle$.

The preceding assumptions are not too restrictive. In particular, condition (8) is very often satisfied in the physical problems alluded to at the end of § 1. The advantage of this model is that it enables us to investigate the physics of the decay phenomenon without too many complicated calculations. The essential physical conclusions would not be modified by using a more elaborate model.

Before taking up the new method for solving the Schrödinger equation which we are describing in this complement, we shall indicate the results of the first-order perturbation theory of Chapter XIII as they apply to this model.

2-c. Results of first-order perturbation theory

The discussion of § C-3 of Chapter XIII enables us to calculate [using, in particular, formula (C-36)] the probability of finding the physical system at time t (initially in the state $|\varphi_i\rangle$) in a final state of arbitrary energy belonging to a group of final states characterized by the interval $\delta\beta_f$ around the value β_f.

Here, we shall concern ourselves with the probability of finding the system in any of the final states $|\alpha\rangle$: neither E nor β is specified. We must therefore integrate expression (C-36) of Chapter XIII with respect to β, which gives the probability density [the integration over the energy was already performed in (C-36)]. Thus, we introduce the constant:

$$\Gamma = \frac{2\pi}{\hbar} \int \mathrm{d}\beta \;\; |\langle \beta, E = E_i | W | \varphi_i \rangle|^2 \;\; \rho(\beta, E = E_i) \tag{9}$$

The desired probability is then equal to Γt. With the assumptions of § 2-a, it represents the probability of the system having left the state $|\varphi_i\rangle$ at time t. If we call $\mathscr{P}_{ii}(t)$ the probability that the system is still in this state at time t, we have:

$$\mathscr{P}_{ii}(t) = 1 - \Gamma t \tag{10}$$

In the discussion of the following sections, it is important to recall the validity conditions for (10):

(*i*) Expression (10) results from a first-order perturbation theory which is valid only if $\mathscr{P}_{ii}(t)$ differs only slightly from its initial value $\mathscr{P}_{ii}(0) = 1$. We then must have:

$$t \ll \frac{1}{\Gamma} \tag{11}$$

(*ii*) Furthermore, (10) is valid only for sufficiently long times t.

To state the second condition more precisely, and to see, in particular, if it is compatible with (11), we return to expression (C-31) of Chapter XIII (β and E are no longer constrained to vary only inside the intervals $\delta\beta_f$ and δE_f). Instead of proceeding as we did in Chapter XIII, we shall integrate the probability density appearing in (C-31), first over β and then over E. The following integral then appears:

$$\frac{1}{\hbar^2} \int_0^\infty \mathrm{d}E \, F\left(t, \frac{E - E_i}{\hbar}\right) K(E) \tag{12}$$

where $K(E)$, which results from the first integration over β, is given by:

$$K(E) = \int \mathrm{d}\beta \; |\langle \beta, E | W | \varphi_i \rangle|^2 \, \rho(\beta, E) \tag{13}$$

$F\left(t, \frac{E-E_i}{\hbar}\right)$ is the diffraction function defined by (C-7) of Chapter XIII, centered at $E = E_i$ and of width $4\pi\hbar/t$.

Let $\hbar\Delta$ be the "width" of $K(E)$: $\hbar\Delta$ represents the order of magnitude of the E variation needed for $K(E)$ to change significantly (*cf.* Fig. 1). As soon as t is sufficiently large that:

$$t \gg \frac{1}{\Delta} \tag{14}$$

$F\left(t, \frac{E-E_i}{\hbar}\right)$ behaves like a "delta function" with respect to $K(E)$. Using relation (C-32)

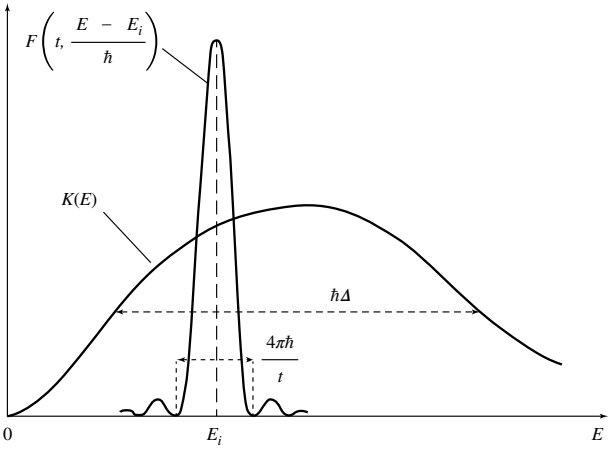


Figure 1: Variation of the functions $K(E)$ and $F\left(t, \frac{E-E_i}{\hbar}\right)$ with respect to E. The respective "widths" of the two curves are of the order of $\hbar\Delta$ and $4\pi\hbar/t$. For sufficiently large t, $F\left(t, \frac{E-E_i}{\hbar}\right)$ behaves like a "delta function" with respect to $K(E)$.

of Chapter XIII, we can then write (12) in the form:

$$\frac{2\pi}{\hbar}t\int \mathrm{d}E\ \delta(E-E_i)\ K(E) = \frac{2\pi t}{\hbar}\ K(E=E_i) = \Gamma t \tag{15}$$

since by comparing (9) and (13), it can easily be seen that:

$$\frac{2\pi}{\hbar}\ K(E=E_i) = \Gamma \tag{16}$$

Again we find that the linear decrease appearing in (10) is valid only if t is large enough to satisfy (14).

Conditions (11) and (14), obviously, are compatible only if:

$$\Delta \gg \Gamma \tag{17}$$

We have thus given a quantitative form to the condition stated in the note of Chap. XIII on page 1318. In the rest of this complement, we shall assume that inequality (17) is satisfied.

2-d. Integrodifferential equation equivalent to the Schrödinger equation

It is easy to adapt expressions (B-11) of Chapter XIII to the case we are studying here.

The state of the system at time t can be expanded on the $\{|\varphi_i\rangle, |\alpha\rangle\}$ basis:

$$|\psi(t)\rangle = b_i(t)\,\mathrm{e}^{-iE_it/\hbar}|\varphi_i\rangle + \int \mathrm{d}\alpha\, b(\alpha,t)\,\mathrm{e}^{-iEt/\hbar}|\alpha\rangle \tag{18}$$

When we substitute state vector (18) into the Schrödinger equation, using the assumptions stated in §§ 2-a and 2-b, we obtain, after a calculation that is analogous to the one in § B-1 of Chapter XIII, the following equations of motion:

$$\begin{cases} i\hbar \dfrac{\mathrm{d}}{\mathrm{d}t}\, b_i(t) = \displaystyle\int \mathrm{d}\alpha\ \mathrm{e}^{i(E_i-E)t/\hbar}\ \langle \varphi_i|W|\alpha\rangle\ b(\alpha,t) & (19) \\ i\hbar \dfrac{\mathrm{d}}{\mathrm{d}t}\, b(\alpha,t) = \ \mathrm{e}^{i(E-E_i)t/\hbar}\ \langle \alpha|W|\varphi_i\rangle\, b_i(t) & (20) \end{cases}$$

The problem consists of using these rigorous equations to predict the behavior of the system after a long time, taking into account the initial conditions:

$$\begin{cases} b_i(0) = 1 & (21a) \\ b(\alpha,0) = 0 & (21b) \end{cases}$$

The simplifying assumptions which we made for W imply that $\dfrac{\mathrm{d}}{\mathrm{d}t}\, b_i(t)$ depends only on $b(\alpha,t)$, and $\dfrac{\mathrm{d}}{\mathrm{d}t}\, b(\alpha,t)$, only on $b_i(t)$. Consequently, we can integrate equation (20), taking initial condition (21b) into account. Substituting the value obtained in this way for $b(\alpha,t)$ into (19), we obtain the following equation describing the evolution of $b_i(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\, b_i(t) = -\frac{1}{\hbar^2}\int \mathrm{d}\alpha \int_0^t \mathrm{d}t'\, \mathrm{e}^{i(E_i-E)(t-t')/\hbar}\ |\langle \alpha|W|\varphi_i\rangle|^2\, b_i(t') \tag{22}$$

By using (4) and performing the integration over β, we obtain, according to (13):

$$\frac{\mathrm{d}}{\mathrm{d}t}\, b_i(t) = -\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}E \int_0^t \mathrm{d}t'\, K(E)\ \mathrm{e}^{i(E_i-E)(t-t')/\hbar}\ b_i(t') \tag{23}$$

Thus, we have been able to obtain an equation involving only b_i. However, it must be noted that this equation is no longer a differential equation, but an integrodifferential equation: the time derivative $\dfrac{\mathrm{d}}{\mathrm{d}t}\, b_i(t)$ depends on the entire "history of the system" between the times 0 and t.

Equation (23) is rigorously equivalent to the Schrödinger equation, but we do not know how to solve it exactly. In the following sections, we shall describe two approximate methods for solving this equation. One of them (§ 3) is equivalent to the first-order theory of Chapter XIII; the other one (§ 4) enables us to study the long-time behavior of the system more satisfactorily.

3. Short-time approximation. Relation to first-order perturbation theory

If t is not too large, that is, if $b_i(t)$ is not too different from $b_i(0) = 1$, we can replace $b_i(t')$ by $b_i(0) = 1$ on the right-hand side of (23). This right-hand side then reduces to a double integral, over E and t', whose integration presents no difficulties:

$$-\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}E \int_0^t \mathrm{d}t'\, K(E)\, \mathrm{e}^{i(E_i-E)(t-t')/\hbar} \tag{24}$$

We shall perform this calculation explicitly, since it allows us to introduce two constants [one of which is Γ, defined by (9)] which play an important role in the more elaborate method described in § 4.

We shall begin by integrating over t' in (24). According to relation (47) of Appendix II, the limit of this integral for $t \to \infty$ is the Fourier transform of the Heaviside step function. More precisely:

$$\lim_{t\to\infty} \int_0^t \mathrm{e}^{i(E_i-E)\tau/\hbar}\, \mathrm{d}\tau = \hbar \left[\pi\delta(E_i - E) + i\mathcal{P}\left(\frac{1}{E_i - E}\right)\right] \tag{25}$$

(we have set $t - t' = \tau$).

Actually, it is not necessary to let t approach infinity in order to use (25) in the calculation of (24). It suffices for $\hbar/t$ to be much smaller than the "width" $\hbar\Delta$ of $K(E)$, that is, for t to be much greater than $1/\Delta$. We again find the validity condition (14). If this condition is satisfied, we can use (25) to write (24) in the form:

$$-\frac{\pi}{\hbar}K(E = E_i) - \frac{i}{\hbar}\mathcal{P}\int_0^\infty \frac{K(E)}{E_i - E}\, \mathrm{d}E \tag{26}$$

The first term of (26) is, according to (16), simply $-\Gamma/2$. We shall set:

$$\delta E = \mathcal{P}\int_0^\infty \frac{K(E)}{E_i - E}\mathrm{d}E \tag{27}$$

Therefore, the double integral (24) is equal to:

$$-\frac{\Gamma}{2} - i\frac{\delta E}{\hbar} \tag{28}$$

When $b_i(t')$ is replaced by $b_i(0) = 1$ in (23), this equation then becomes [as soon as (14) is satisfied]:

$$\frac{\mathrm{d}}{\mathrm{d}t}b_i(t) = -\frac{\Gamma}{2} - i\frac{\delta E}{\hbar} \tag{29}$$

The solution of (29), using the initial condition (21a), is very simple:

$$b_i(t) = 1 - \left(\frac{\Gamma}{2} + i\frac{\delta E}{\hbar}\right)t \tag{30}$$

Obviously, this result is valid only if $|b_i(t)|$ differs slightly from 1, that is if:

$$t \ll \frac{1}{\Gamma} \quad , \quad t \ll \frac{\hbar}{\delta E} \tag{31}$$

This is the other validity condition, (11), for first-order perturbation theory.

Using (30), we can easily calculate the probability $\mathscr{P}_{ii}(t) = |b_i(t)|^2$ that the system is still in the state $|\varphi_i\rangle$ at time t. If we neglect terms in Γ^2 and δE^2, we obtain:

$$\mathscr{P}_{ii}(t) = 1 - \Gamma t \tag{32}$$

All the results obtained in Chapter XIII can then be deduced from equation (23) when $b_i(t')$ is replaced by $b_i(0)$. This equation has also enabled us to introduce the parameter δE, whose physical significance will be discussed later [note that δE does not appear in the treatment of Chapter XIII because we were concerned only with the calculation of the probability $|b_i(t)|^2$, and not with that of the probability amplitude $b_i(t)$].

4. Another approximate method for solving the Schrödinger equation

A better approximation consists of replacing $b_i(t')$ by $b_i(t)$ rather than by $b_i(0)$ in (23). To see this, we shall begin by performing the integral over E which appears on the right-hand side of the rigorous equation, (23). We obtain a function of E_i and $t-t'$:

$$g(E_i, t-t') = -\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}E\, K(E)\, \mathrm{e}^{i(E_i-E)(t-t')/\hbar} \tag{33}$$

which is clearly different from zero only if $t-t'$ is very small. In (33), we are integrating over E the product of $K(E)$, which varies slowly with E (*cf.* Fig. 1), and an exponential whose period with respect to the variable E is $2\pi\hbar/(t-t')$. If we choose values of t and t' such that this period is much smaller than the width $\hbar\Delta$ of $K(E)$, the product of these two functions undergoes numerous oscillations when E is varied, and its integral over E is negligible. Consequently, the modulus of $g(E_i, t-t')$ is large for $t-t' \simeq 0$ and becomes negligible as soon as $t-t' \gg 1/\Delta$. This property means that, for all t, the only values of $b_i(t')$ to enter significantly into the right-hand side of (23) are those which correspond to t' very close to t (more precisely, $t-t' \underset{\sim}{<} 1/\Delta$). Indeed, once the integration over E has been performed, this right-hand side becomes:

$$\int_0^t g(E_i, t-t')\, b_i(t')\, \mathrm{d}t' \tag{34}$$

and we see that the presence of $g(E_i, t-t')$ practically eliminates the contribution of $b_i(t')$ as soon as $t-t' \gg 1/\Delta$.

Thus, the derivative $\dfrac{\mathrm{d}}{\mathrm{d}t}b_i(t)$ has only a very short memory of the previous values of $b_i(t)$ between 0 and t. Actually, it depends only on the values of b_i at times immediately before t, and *this is true for all* t. This property enables us to transform the integrodifferential equation (23) into a differential equation. If $b_i(t)$ varies very little over a time interval of the order of $1/\Delta$, we make only a small error by replacing $b_i(t')$ by $b_i(t)$ in (34). This yields:

$$b_i(t)\int_0^t g(E_i, t-t')\mathrm{d}t' = -\left(\frac{\Gamma}{2} + i\frac{\delta E}{\hbar}\right) b_i(t) \tag{35}$$

[to write the right-hand side of (35), we used the fact that the integral over t' of $g(E_i, t-t')$ is simply, according to (33), the double integral (24) evaluated in § 3 above].

Now, according to the results of § 3 (and as we shall see later), the time scale characteristic of the evolution of $b_i(t)$ is of the order of $1/\Gamma$ or $\hbar/\delta E$. The validity conditions for (35) are then:

$$\Gamma \ll \Delta \quad , \quad \frac{\delta E}{\hbar} \ll \Delta \tag{36}$$

which we have already assumed to be fulfilled [*cf.* (17)].

To a good approximation, and for all t, equation (23) can therefore be written:

$$\frac{\mathrm{d}}{\mathrm{d}t}\, b_i(t) = -\left(\frac{\Gamma}{2} + i\frac{\delta E}{\hbar}\right) b_i(t) \tag{37}$$

whose solution, using (21a), is obvious:

$$b_i(t) = \mathrm{e}^{-\Gamma t/2}\mathrm{e}^{-i\delta E t/h} \tag{38}$$

It can easily be shown that the limited expansion of (38) gives (30) to first order in Γ and δE.

Comment:

No upper bound has been imposed on the time t. On the other hand, the integral $\int_0^t g(E_i, t-t')\mathrm{d}t'$ which appears in (35) is equal to $-(\Gamma/2+i\delta E/\hbar)$ only if $t \gg 1/\Delta$, as we saw in § 3 above. For very short times, the theory presented here suffers from the same limitations as perturbation theory; however, it has the great advantage of being valid for long times.

If we now substitute expression (38) for $b_i(t)$ into equation (20), we obtain a very simple equation which enables us to determine the probability amplitude $b(\alpha, t)$ associated with the state $|\alpha\rangle$:

$$b(\alpha, t) = \frac{1}{i\hbar}\,\langle\alpha|W|\varphi_i\rangle \int_0^t \mathrm{e}^{-\Gamma t'/2}\,\mathrm{e}^{i(E-E_i-\delta E)t'/\hbar}\mathrm{d}t' \tag{39}$$

that is:

$$b(\alpha, t) = \frac{\langle\alpha|W|\varphi_i\rangle}{\hbar}\;\; \frac{1-\mathrm{e}^{-\Gamma t/2}\,\mathrm{e}^{i(E-E_i-\delta E)t/\hbar}}{\frac{1}{\hbar}(E-E_i-\delta E)+i\frac{\Gamma}{2}} \tag{40}$$

Equations (38) and (40), respectively, describe the decay of the initial state and the "filling" of the final states $|\alpha\rangle$. Now let us study in greater detail the physical content of these two equations.

5. Discussion

5-a. Lifetime of the discrete state

According to (38), we have:

$$\mathscr{P}_{ii}(t) = |b_i(t)|^2 = \mathrm{e}^{-\Gamma t} \tag{41}$$

$\mathscr{P}_{ii}(t)$ therefore decreases *irreversibly* from $\mathscr{P}_{ii}(0) = 1$ and approaches zero as $t \to \infty$ (Fig. 2). The discrete initial state is said to have a *finite lifetime* τ, where τ is the time constant of the exponential of Figure 2 :

$$\tau = \frac{1}{\Gamma} \tag{42}$$

This irreversible behavior contrasts sharply with the oscillations of the system (Rabi's formula) between two discrete states when it is subject to a resonant perturbation coupling these two states.

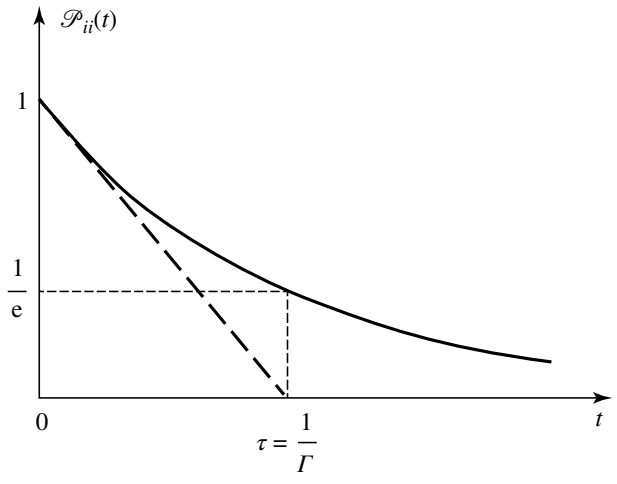


Figure 2: Variation with respect to time of the probability of finding the system in the discrete state $|\varphi_i\rangle$ at time t. We obtain an exponential decrease, $e^{-t/\tau}$, for which Fermi's golden rule gives the tangent at the origin (this tangent is represented by a dashed line).

5-b. Shift of the discrete state due to the coupling with the continuum

If we go from $b_i(t)$ to $c_i(t)$ [*cf.* formula (B-8) of Chapter XIII], we obtain, from (38):

$$c_i(t) = e^{-\Gamma t/2}\, e^{-i(E_i+\delta E)t/\hbar} \tag{43}$$

Recall that, in the absence of the coupling W, we would have:

$$c_i(t) = e^{-iE_i t/\hbar} \tag{44}$$

In addition to the exponential decrease, $e^{-\Gamma t/2}$, the coupling with the continuum is therefore responsible for a shift in the discrete state energy, which goes from E_i to $E_i + \delta E$. This is the interpretation of the quantity δE introduced in § 3.

Let us analyze expression (27) for δE more closely. Substituting definition (13) of $K(E)$ into (27), we get:

$$\delta E = \mathcal{P} \int_0^\infty \frac{dE}{E_i - E} \int d\beta\, \rho(\beta, E)\, |\langle \beta, E|W|\varphi_i\rangle|^2 \tag{45}$$

or, if we use (4) and replace $\langle \beta, E|$ by $\langle \alpha|$:

$$\delta E = \mathcal{P} \int d\alpha \frac{|\langle \alpha|W|\varphi_i\rangle|^2}{E_i - E} \tag{46}$$

The contribution to this integral of a particular state $|\alpha\rangle$ of the continuum, for which $E \neq E_i$, is:

$$\frac{|\langle \alpha|W|\varphi_i\rangle|^2}{E_i - E} \tag{47}$$

We recognize (47) as a familiar expression in stationary perturbation theory [*cf.* formula (B-14) of Chapter XI]. (47) represents the energy shift of the state $|\varphi_i\rangle$ due to the coupling with the state $|\alpha\rangle$, to second order in W. δE is simply the sum of the shifts due to the various states $|\alpha\rangle$ of the continuum. We might imagine that a problem would appear for the states $|\alpha\rangle$ for which $E = E_i$. Actually, the presence in (46) of the principal part $\mathcal{P}$ implies that the contribution of the states $|\alpha\rangle$ situated immediately above $|\varphi_i\rangle$ compensates that of the states situated immediately below.

Summing up:

(*i*) The coupling of $|\varphi_i\rangle$ with the states $|\alpha\rangle$ of the same energy is responsible for the finite lifetime of $|\varphi_i\rangle$ [the function $\delta(E_i - E)$ of formula (25) enters into the expression for Γ].

(*ii*) The coupling of $|\varphi_i\rangle$ with the states $|\alpha\rangle$ of different energies is responsible for an energy shift of the state $|\varphi_i\rangle$. This shift can be calculated by stationary perturbation theory (this was not obvious in advance).

Comment:

In the particular case of the spontaneous emission of a photon by an atom, δE represents the shift of the atomic level under study due to the coupling with the continuum of final states (an atom in another discrete state, in the presence of a photon). The difference between the shifts δE of the $2s_{1/2}$ and $2p_{1/2}$ states of the hydrogen atom is the "Lamb shift" [*cf.* Complement K_V, § 3-d-δ and Chapter XII, § C-3-b, comment (*iv*)].

5-c. Energy distribution of the final states

Once the discrete state has decayed, that is, when $t \gg 1/\Gamma$, the final state of the system belongs to the continuum of states $|\alpha\rangle$. It is interesting to study the energy distribution of the possible final states. For example, in the spontaneous emission of a photon by an atom, this energy distribution is that of the photon emitted when the atom falls back from the excited level to a lower level (the natural width of spectral lines).

When $t \gg 1/\Gamma$, the exponential which appears in the numerator of (40) is practically zero. We then have:

$$|b(\alpha,t)|^2 \underset{t\gg 1/\Gamma}{\sim} |\langle\alpha|W|\varphi_i\rangle|^2 \frac{1}{(E - E_i - \delta E)^2 + \hbar^2\Gamma^2/4} \tag{48}$$

$|b(\alpha,t)|^2$ actually represents a probability density. The probability of finding the system, after the decay, in a group of final states characterized by the intervals $\mathrm{d}\beta_f$ and $\mathrm{d}E_f$ about β_f and E_f can be calculated directly from (48):

$$\mathrm{d}\mathscr{P}(\beta_f, E_f, t) = |\langle\beta_f, E_f|W|\varphi_i\rangle|^2 \rho(\beta_f, E_f) \frac{1}{(E_f - E_i - \delta E)^2 + \hbar^2\Gamma^2/4}\, \mathrm{d}\beta_f\, \mathrm{d}E_f \tag{49}$$

Let us examine the E_f-dependence of the probability density:

$$\frac{\mathrm{d}\mathscr{P}(\beta_f, E_f, t)}{\mathrm{d}\beta_f\, \mathrm{d}E_f}$$

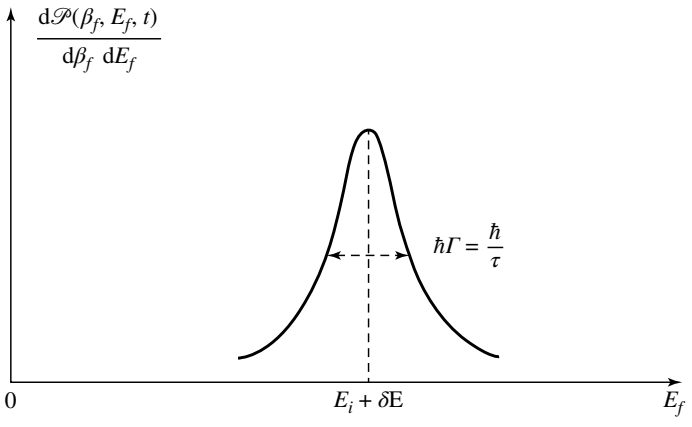


Figure 3: Form of the energy distribution of the final states attained by the system after the decay of the discrete state. We obtain a Lorentzian distribution centered at $E_i + \delta E$ (the energy of the discrete state corrected by the shift δE due to the coupling with the continuum). The shorter the lifetime τ of the discrete state, the wider the distribution (time-energy uncertainty relation).

Since $|\langle \beta_f, E_f|W|\varphi_i\rangle|^2 \rho(\beta_f, E_f)$ remains practically constant when E_f varies over an interval of the order of $\hbar\Gamma$, the variation of the probability density with respect to E_f is essentially determined by the function:

$$\frac{1}{(E_f - E_i - \delta E)^2 + \hbar^2\Gamma^2/4} \tag{50}$$

and has, consequently, the form shown in Figure 3. The energy distribution of the final states has a maximum for $E_f = E_i + \delta E$, that is, when the final state energy is equal to that of the initial state $|\varphi_i\rangle$, corrected by the shift δE. The form of the distribution is that of a Lorentz curve of width $\hbar\Gamma$, called the "natural width" of the state $|\varphi_i\rangle$. An energy dispersion of the final states therefore appears. The larger $\hbar\Gamma$ (that is, the shorter the lifetime $\tau = 1/\Gamma$ of the discrete state), the greater the dispersion. More precisely:

$$\Delta E_f = \hbar\Gamma = \frac{\hbar}{\tau} \tag{51}$$

Note again the analogy between (51) and the time-energy uncertainty relation. In the presence of the coupling W, the state $|\varphi_i\rangle$ can be observed only during a finite time, of the order of its lifetime τ. When we want to determine its energy by measuring that of the final state of the system, the uncertainty ΔE of the result cannot be much less than $\hbar/\tau$.

References:

The original article: Weisskopf and Wigner (2.33).

Complement E_{XIII}

Time-dependent random perturbation, relaxation

1 Evolution of the density operator 1391
1-a Coupling Hamiltonian, correlation times 1392
1-b Evolution of a single system 1393
1-c Evolution of the ensemble of systems 1396
1-d General equations for the relaxation 1397
2 Relaxation of an ensemble of spin 1/2's 1398
2-a Characterization of the operators, isotropy of the perturbation 1399
2-b Longitudinal relaxation 1400
2-c Transverse relaxation 1403
3 Conclusion 1408

This complement examines the problem studied in § D of Chapter XIII, both in a more precise and general way. Rather than studying a single system, we shall study an ensemble of individual quantum systems subjected to an external random perturbation. This type of situation often occurs in magnetic resonance experiments where one measures the global magnetization of an ensemble of spins each carrying a small magnetic moment, as for example the nuclear spins of atoms in a gas. As the atoms move, they undergo collisions with impurities contained in the gas or on the walls of the container. As mentioned in Chapter XIII, if these impurities carry a magnetic moment, such collisions may change the directions of the nuclear spins of the colliding atoms. The corresponding perturbation lasts for a very short time (the collision time), and is of a random nature since the magnetic moment of the impurities can have any direction. The gas of atomic spins is thus subjected to a sum of random perturbations that rapidly change their values (and signs), hence having a very short correlation time. Another classic example of random perturbation is an experiment where an ensemble of atoms is illuminated by a light source. Several reasons give the interaction between the atoms and the incident electromagnetic field a random character. First of all, most light sources produce fields that have rapid frequency and phase fluctuations. This means that the field itself must be characterized in a stochastic manner, with a short coherence time. Furthermore, even if the light source is an almost perfectly monochromatic laser, the atoms' motion is random. Because of the Doppler effect, the atoms will be coupled, in their own reference frame, to a field having a random frequency. Studying the propagation of a light beam in an atomic gas thus involves the study of a large number of individual atoms, each subjected to a different and random perturbation. Many examples exist of similar situations involving rapidly fluctuating perturbations.

This complement examines how the effect of such a random perturbation on an ensemble of individual systems must be treated using quantum mechanics. In a more general framework than the one used in § D of Chapter XIII, we will show that the coupling with the random perturbation produces a so-called "relaxation" phenomenon in the global system, very different from the evolution in the absence of the random

event. We saw that the coupling with a constant or perfectly sinusoidal perturbation produces oscillations in the physical system, at Bohr frequencies that are modified by the interaction. This complement will describe a totally different behavior, an exponential evolution with real exponents, leading to irreversible evolutions. An example of such behavior could be the relaxation of a system towards thermal equilibrium, very different than an oscillation.

We first study (§ 1) the evolution of the density operator characterizing the ensemble of systems. This leads to a general relaxation equation valid whenever the correlation times are very short. In the following section (§ 2) we apply this general equation to an important specific case, an ensemble of spin 1/2's coupled to statistically isotropic perturbations. This will enable us to explain the important concepts of "longitudinal relaxation" and "transverse relaxation", which play a central role in many magnetic resonance experiments.

1. Evolution of the density operator

Consider an ensemble of N individual systems labeled by the index $q = 1, 2, ...,N$. Each system is described by a density operator (Complements $\mathrm{E_{III}}$ and $\mathrm{E_{IV}}$) noted $\rho_q(t)$. Statistically, the ensemble of the N systems is described by the following density operator $\rho_N(t)$:

$$\rho_N(t) = \frac{1}{N}\sum_{q=1}^{N}\rho_q(t) \tag{1}$$

Each individual system evolves under the effect of an operator $H_q(t)$, the sum of two Hamiltonians:

$$H_q(t) = H_0 + W_q(t) \tag{2}$$

The first, H_0, is the Hamiltonian common to all the individual systems, corresponding for example to the coupling of their spins with an external static magnetic field. We assume that this Hamiltonian does not depend on time. The second, $W_q(t)$, is the coupling Hamiltonian with the random perturbation. It depends not only on the time but also on the index q of the individual system.

We note $|\varphi_n\rangle$ the eigenvectors of H_0:

$$H_0|\varphi_n\rangle = e_n|\varphi_n\rangle \tag{3}$$

having the energies $e_n = \hbar\omega_n$; we set:

$$\omega_{nk} = \omega_n - \omega_k \tag{4}$$

The matrix elements of the coupling Hamiltonian are written, in this basis:

$$W_{nk}^{(q)}(t) = \langle\varphi_n|W_q(t)|\varphi_k\rangle \tag{5}$$

1-a. Coupling Hamiltonian, correlation times

Consider the ensemble of the N functions $W_{nk}^{(q)}(t)$ obtained when q varies from 1 to N: they are N different realizations of the same matrix element. Choosing q randomly[1] introduces a random function of time, $W_{nk}(t)$. Changing the values of n and k, and calling P the dimension of the state space, we can define P^2 random functions. These random functions can be considered as the matrix elements of an operator $W(t)$ that is also a random function of time. In other words, $W(t)$ is the operator obtained by choosing randomly the value of q labeling the N operators $W_q(t)$. The statistical correlation properties of this random operator (or of its matrix elements) play an essential role in what follows.

The ensemble formed by the N systems can be considered as an ensemble of different possible realizations of the same individual system, called in statistical mechanics the "Gibbs ensemble". It is equivalent to take an average over this ensemble at a given time or over a single system taken at a large number of different times (ergodic hypothesis). This average will be symbolized by placing a horizontal line over the letter W.

We first assume that the average value of the perturbation is zero:

$$\overline{W}(t) = 0 \qquad \text{that is:} \quad \overline{W}_{nk}(t) = \overline{W}_{kn}(t) = 0 \ \text{ for any } k,\, n \tag{6}$$

This hypothesis is not restrictive since, if the average value of $\overline{W}(t)$ is any constant operator h, that operator can be added to the Hamiltonian H_0, hence keeping the average value of the perturbation equal to zero. As $W_{kn}(t)$ and $W_{nk}(t)$ are complex conjugates, their product is always a positive number whose average is, a priori, not zero. The amplitude of the perturbation is then defined by the average value of the products:

$$\overline{W_{kn}(t)\, W_{nk}(t)} = C_{nk}^0 \tag{7}$$

As we assume the random function to be stationary, this mean square is independent of time. In a more general way, one can define a series of cross-correlation coefficients, also time-independent:

$$\overline{W_{lp}(t)\, W_{nk}(t)} = C_{nk,lp}^0 \tag{8}$$

To characterize a function random in time, we also need to consider averages of products taken at different times. As we did in § D of Chap. XIII, we introduce the correlation functions between time t and $t+\tau$:

$$\overline{W_{kn}(t+\tau)\, W_{nk}(t)} = C_{nk}(\tau) \tag{9}$$

Since the random functions are stationary by hypothesis, the function C_{nk} only depends on the difference between the times $t+\tau$ and t, and we can also write:

$$\overline{W_{kn}(\tau)\, W_{nk}(0)} = C_{nk}(\tau) \tag{10}$$

We know that, by definition, the function $C_{nk}(\tau)$ starts from a positive value for $\tau = 0$. As the delay τ starts increasing, the correlation between $W_{kn}(\tau)$ and $W_{nk}(0)$

[1] We assume the number N to be very large. As an example, a millimeter cube of gas, at standard temperature and pressure, contains roughly 10^{16} atoms.

rapidly decreases. The function $C_{nk}(\tau)$ tends towards zero, with a characteristic time called the "correlation time" and noted τ_c:

$$\overline{W_{kn}(t+\tau)\,W_{nk}(t)} \to 0 \qquad \text{if} \qquad \tau \gg \tau_c \tag{11}$$

For instance, if the perturbation is induced by the collision between an atom and an impurity, it will clearly lose any memory of its value between one collision and the next, or even right after a single collision. As collision times are often very short, there are many examples where τ_c is a very short time. This analysis can be generalized to cross-correlation coefficients like the one described in (8).

One often uses a model where $C_{nk}(\tau)$ is a decreasing exponential of the delay τ:

$$\overline{W_{kn}(t+\tau)\,W_{nk}(t)} = C_{nk}^{0}\; e^{-\tau/\tau_c} \tag{12}$$

To simplify the notation, we only took into account a single correlation time τ_c, independent of k and n; the generalization to several correlation times is straightforward. Similar relations as the ones we just wrote can be obtained for cross-correlation coefficients like the one described in (8). This leads to a whole series of correlation times depending on numerous indices. In the collision example discussed above, all these times are of the same order of magnitude as the very short collision time; a natural approximation is to assume that they are comparable, and to call τ_c the longest amongst all these times.

1-b. Evolution of a single system

We now study the evolution of a single quantum system (the value of q is fixed). It will be treated in the interaction picture (exercise 15 of Complement $\mathrm{L_{III}}$), which we now briefly review.

α. *Interaction picture*

The evolution of each density operator $\rho_q(t)$ obeys the usual von Neumann equation, with a commutator on the right-hand side:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\rho_q(t) = [H_0 + W_q(t)\,,\rho_q(t)] \tag{13}$$

In this right-hand side, the term containing H_0 may lead to a rapid evolution. The term containing $W_q(t)$ is assumed to be smaller, hence leading to a slower evolution that can be treated using approximations.

(i) To start with, let us assume that the coupling Hamiltonian $W_q(t)$ is zero. The evolution of $\rho_q(t)$ is only due to H_0. It is useful to express it as a function of the evolution operator $U_0(t,t_i)$ between the times t_i and t (Complement $\mathrm{F_{III}}$) associated with the non-perturbed Hamiltonian H_0:

$$U_0(t,t_i) = e^{-iH_0(t-t_i)/\hbar} \tag{14}$$

This is a unitary operator:

$$U_0^{\dagger}(t,t_i)\,U_0(t,t_i) = 1 \qquad \text{for any } t_i \text{ or } t \tag{15}$$

When $W_q(t)$ is zero, we simply have:

$$\rho_q(t) = U_0(t,t_i)\ \ \rho_q(t_i)\ \ U_0^\dagger(t,t_i) \tag{16}$$

as we now show. Taking the derivative of this relation with respect to t, the derivation of $U_0(t,t_i)$ introduces a term $-iH_0\rho_q(t)/\hbar$; the derivation of $U_0^\dagger(t,t_i)$ introduces a term $+i\rho_q(t)H_0/\hbar$. Both terms together reconstruct the H_0 term of the commutator on the right-hand side of (13), which shows that equation (13) is verified by solution (16). As for the initial condition for $t = t_i$, it is also verified since the unitary operator $U_0(t,t_i)$ then becomes the unit operator.

We can also use $U_0(t,t_i)$ to perform an inverse unitary transformation on $\rho_q(t)$ and define the modified density operator $\widetilde{\rho}_q(t)$ as:

$$\widetilde{\rho}_q(t) = U_0^\dagger(t,t_i)\,\rho_q(t)\,U_0(t,t_i) \tag{17}$$

Inserting relation (16) in this definition, we obtain relation (15) twice, both on the left and on the right of $\rho_q(t_i)$. All the evolution operators thus disappear, and we get:

$$\widetilde{\rho}_q(t) = \rho_q(t_i) \tag{18}$$

This shows that $\widetilde{\rho}_q(t)$ does not depend on time as long as the coupling $W_q(t)$ remains equal to zero.

(*ii*) Even when $W_q(t)$ is no longer zero, it is still useful to apply the unitary transformation (17) to the density operator. This operation is generally referred to as the "passage to the interaction picture". In this picture, the evolution of the density operator $\widetilde{\rho}_q(t)$ is only due to the presence of the interaction $W_q(t)$. According to our assumptions, this evolution is much slower than the evolution $\rho_q(t)$, which is also governed by H_0. This property considerably facilitates the use of approximations and will be used in this complement.

Let us take the time derivative of (17), starting with the derivative of the two unitary operators on the left and on the right of $\rho_q(t)$, followed by the derivative of the operator itself. This yields:

$$\begin{aligned} i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\rho}_q(t) &= -H_0\widetilde{\rho}_q(t) + \widetilde{\rho}_q(t)H_0 + U_0^\dagger(t,t_i)\ \ i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\rho_q(t)\ \ U_0(t,t_i) \\ &= -\left[H_0,\widetilde{\rho}_q(t)\right] + U_0^\dagger(t,t_i)\left[H_0 + W_q(t),\rho_q(t)\right]U_0(t,t_i) \end{aligned} \tag{19}$$

Now for any operator A, the fact that U_0 is a unitary operator allows transforming the following commutator according to:

$$U_0^\dagger(t,t_i)\left[A,\rho_q(t)\right]U_0(t,t_i) = \left[U_0^\dagger(t,t_i)\,AU_0(t,t_i),U_0^\dagger(t,t_i)\,\rho_q(t)\,U_0(t,t_i)\right] \tag{20}$$

(to check this, one can simply expand both commutators and use the relation $U_0U_0^\dagger = 1$). The right-hand side now contains $\widetilde{\rho}_q(t)$. If $A = H_0 + W_q(t)$, since H_0 commutes with U_0, we get:

$$\begin{aligned} &U_0^\dagger(t,t_i)\left[H_0 + W_q(t),\rho_q(t)\right]U_0(t,t_i) \\ &\qquad = \left[H_0,\widetilde{\rho}_q(t)\right] + \left[U_0^\dagger(t,t_i)\,W_q(t)\,U_0(t,t_i),\widetilde{\rho}_q(t)\right] \end{aligned} \tag{21}$$

The right-hand side contains the unitary transform $\widetilde{W}_q(t)$ of $W_q(t)$, obtained by the same unitary transformation that led from $\rho_q(t)$ to $\widetilde{\rho}_q(t)$:

$$\widetilde{W}_q(t) = U_0^\dagger(t,t_i)\, W_q(t)\, U_0(t,t_i) \tag{22}$$

Inserting this expression in the right-hand side of (19), the commutators containing H_0 on both sides cancel out. We finally get the simple relation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\rho}_q(t) = \left[\widetilde{W}_q(t), \widetilde{\rho}_q(t)\right] \tag{23}$$

This evolution equation only contains operators in the interaction picture. The hamiltonian H_0 is no longer explicitly present (but is implicitly contained in the unitary transformation that leads to the interaction picture). It is easy to verify that $\widetilde{\rho}_q(t)$ does not evolve in the absence of perturbation.

β. Approximate calculation of the evolution

Integrating over time equation (23) yields:

$$\widetilde{\rho}_q(t) = \rho_q(t_i) + \frac{1}{i\hbar}\int_{t_i}^{t} \mathrm{d}t' \ \left[\widetilde{W}_q(t'), \widetilde{\rho}_q(t')\right] \tag{24}$$

which, inserted in the same equation, leads to:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\rho}_q(t) = \left[\widetilde{W}_q(t), \widetilde{\rho}_q(t_i)\right] + \frac{1}{i\hbar}\int_{t_i}^{t} \mathrm{d}t' \ \left[\widetilde{W}_q(t), \left[\widetilde{W}_q(t'), \widetilde{\rho}_q(t')\right]\right] \tag{25}$$

This evolution equation for $\widetilde{\rho}_q(t)$, which now contains a double commutator, is exact. As $\widetilde{\rho}_q(t')$ appears in the integral, it is an integro-differential equation.

This equation can be transformed into a simple differential equation, using the following approximation. If the effect of the perturbation W_q remains limited during the time interval from t_i to t, a good approximation of the evolution of $\widetilde{\rho}_q(t)$ in that interval is to replace $\widetilde{\rho}_q(t')$ by its value for any time chosen in that interval. Choosing for example the time t_i, yields the differential equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\rho}_q(t) = \left[\widetilde{W}_q(t), \widetilde{\rho}_q(t_i)\right] + \frac{1}{i\hbar}\int_{t_i}^{t} \mathrm{d}t' \ \left[\widetilde{W}_q(t), \left[\widetilde{W}_q(t'), \widetilde{\rho}_q(t_i)\right]\right] \tag{26}$$

The delay τ can be introduced explicitly by performing the change of integration variable:

$$\tau = t - t' \tag{27}$$

Dividing both sides by $i\hbar$, we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\rho}_q(t) = \frac{1}{i\hbar}\left[\widetilde{W}_q(t), \widetilde{\rho}_q(t_i)\right] - \frac{1}{\hbar^2}\int_0^{t-t_i} \mathrm{d}\tau \ \left[\widetilde{W}_q(t), \left[\widetilde{W}_q(t-\tau), \widetilde{\rho}_q(t_i)\right]\right] \tag{28}$$

1-c. Evolution of the ensemble of systems

To obtain the evolution equation for the density operator $\rho_N(t)$ describing the ensemble of the N systems, we first transform that operator as in (17) to use the interaction picture:

$$\widetilde{\rho}_N(t) = U_0^\dagger(t,t_i)\,\rho_N(t)\,U_0(t,t_i) \tag{29}$$

The initial density operator can easily be retrieved using the inverse unitary transformation. Definition (1) of $\widetilde{\rho}_N(t)$ shows that its evolution is obtained by summing relation (28) over the index q, and dividing the result by N, the total number of systems. In other words, it means that we have to take the ensemble average of both sides of (28). This operation is difficult to carry out without making some hypotheses about the characteristics of the random functions that come into play. We shall assume that the evolution of $\widetilde{\rho}_N(t)$ occurs with time constants that are much longer than the correlation time τ_c. We shall explain below (§ 1-d-β) what this implies in terms of the parameters defining the interactions, hence verifying that the computation is consistent.

With each time t we can associate a previous time t_i such that $t - t_i$ is very large compared to the correlation time τ_c, while remaining small compared to the characteristic evolution time of the density operator in the interaction picture. We shall then use relations (9) and (11) that characterize the random perturbation. The first commutator on the right-hand side of the evolution equation (28) involves two operators at different times and which are therefore not correlated: $\widetilde{\rho}_q(t_i)$ depends on values of the perturbation $\widetilde{W}_q$ at times earlier than t_i, whereas $\widetilde{W}_q(t)$ is the value of the perturbation at a time later than t_i by a time larger than τ_c. Taking the average over all the values of q then shows that this first term cancels out since we assumed in (6) that the average values of the matrix elements of the perturbation are zero. As for the following integral, it contains the average value over q of the product $\widetilde{W}_q(t)\,\widetilde{W}_q(t-\tau)\,\widetilde{\rho}_q(t_i)$, in that order or any other order. Contributions to this integral only come from values of the delay τ of the order of the correlation time τ_c; if $\tau \gg \tau_c$, $\widetilde{W}_q(t)$ depends neither on $\widetilde{W}_q(t-\tau)$ nor on $\widetilde{\rho}_q(t_i)$, which allows factoring an average value that is equal to zero. This has two consequences. First $\widetilde{\rho}_q(t_i)$ is not correlated with the two terms in $\widetilde{W}_q$, so that we can compute its average separately and replace $\widetilde{\rho}_q(t_i)$ by $\widetilde{\rho}_N(t_i)$. The second consequence is that we can replace the integral upper bound $t-t_i$ by infinity without changing significantly its value. This leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\rho}_N(t) = -\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau\left[\overline{\widetilde{W}(t),[\widetilde{W}(t-\tau),\widetilde{\rho}_N(t_i)]}\right] \tag{30}$$

where, as before, the bar on top of the operators stands for the ensemble average (this average only concerns the perturbation W, not the density operator).

An additional simplification comes from the fact that we assumed $t - t_i$ to be short compared to the evolution time of $\widetilde{\rho}_N(t)$. It is therefore a good approximation to replace $\widetilde{\rho}_N(t_i)$ by $\widetilde{\rho}_N(t)$ in the right-hand side of this equation. This finally leads to the relaxation equation of the density operator in the interaction representation:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\rho}_N(t) = -\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau\left[\overline{\widetilde{W}(t),[\widetilde{W}(t-\tau),\widetilde{\rho}_N(t)]}\right]} \tag{31}$$

Using (29), we obtain the corresponding equation in the usual representation.

1-d. General equations for the relaxation

We now use the previous results to compute the evolution of the density matrix.

α. Evolution of the matrix elements of the density operator

Relation (31) can be written in the basis of the eigenvectors $|\varphi_n\rangle$ of the Hamiltonian H_0, in order to directly obtain the coupled evolution equations of the different matrix elements of the density operator. Since:

$$\langle\varphi_n|\,\widetilde{W}_q(t)\,|\varphi_k\rangle = \mathrm{e}^{i\omega_{nk}(t-t_i)}\ \langle\varphi_n|\,W_q(t)\,|\varphi_k\rangle \tag{32}$$

where ω_{nk} is defined in (4), we get:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\varphi_n|\,\widetilde{\rho}_N(t)\,|\varphi_k\rangle = -\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau \\ \sum_{p,r}\Big\{&\mathrm{e}^{i\omega_{nr}(t-t_i)}\mathrm{e}^{i\omega_{rp}\tau}\ \overline{\langle\varphi_n|\,W(t)\,|\varphi_p\rangle\langle\varphi_p|\,W(t-\tau)\,|\varphi_r\rangle}\ \langle\varphi_r|\,\widetilde{\rho}_N(t)\,|\varphi_k\rangle \\ -\ &\mathrm{e}^{i(\omega_{np}-\omega_{kr})(t-t_i)}\mathrm{e}^{i\omega_{kr}\tau}\ \overline{\langle\varphi_n|\,W(t)\,|\varphi_p\rangle\langle\varphi_r|\,W(t-\tau)\,|\varphi_k\rangle}\ \langle\varphi_p|\,\widetilde{\rho}_N(t)\,|\varphi_r\rangle \\ -\ &\mathrm{e}^{i(\omega_{np}-\omega_{kr})(t-t_i)}\mathrm{e}^{i\omega_{pn}\tau}\ \overline{\langle\varphi_n|\,W(t-\tau)\,|\varphi_p\rangle\langle\varphi_r|\,W(t)\,|\varphi_k\rangle}\ \langle\varphi_p|\,\widetilde{\rho}_N(t)\,|\varphi_r\rangle \\ +\ &\mathrm{e}^{i\omega_{pk}(t-t_i)}\mathrm{e}^{i\omega_{rp}\tau}\ \overline{\langle\varphi_p|\,W(t-\tau)\,|\varphi_r\rangle\langle\varphi_r|\,W(t)\,|\varphi_k\rangle}\ \langle\varphi_n|\,\widetilde{\rho}_N(t)\,|\varphi_p\rangle\Big\}\end{aligned} \tag{33}$$

The random functions associated with W are stationary, as seen from relation (9). This allows adding an arbitrary time to the two variables they contain, in the right-hand side of the previous equation. We can thus replace t by τ and $t-\tau$ by 0.

We now leave the interaction picture and come back to the usual picture (laboratory picture) using the unitary transformation (17), written in the $|\varphi_n\rangle$ basis:

$$\langle\varphi_n|\,\rho_N(t)\,|\varphi_k\rangle = \mathrm{e}^{-i\omega_{nk}(t-t_i)}\,\langle\varphi_n|\,\widetilde{\rho}_N(t)\,|\varphi_k\rangle \tag{34}$$

This relation leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\varphi_n|\,\rho_N(t)\,|\varphi_k\rangle = i(\omega_k-\omega_n)\langle\varphi_n|\,\rho_N(t)\,|\varphi_k\rangle + \mathrm{e}^{-i\omega_{nk}(t-t_i)}\frac{\mathrm{d}}{\mathrm{d}t}\langle\varphi_n|\,\widetilde{\rho}_N(t)\,|\varphi_k\rangle \tag{35}$$

The general relaxation equations are then written:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\ \langle\varphi_n|\,\rho_N(t)\,|\varphi_k\rangle = i(\omega_k-\omega_n)\langle\varphi_n|\,\rho_N(t)\,|\varphi_k\rangle \\ -\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau\sum_{p,r}\Big\{&\mathrm{e}^{i\omega_{rp}\tau}\ \overline{\langle\varphi_n|\,W(\tau)\,|\varphi_p\rangle\langle\varphi_p|\,W(0)\,|\varphi_r\rangle}\ \langle\varphi_r|\,\rho_N(t)\,|\varphi_k\rangle \\ &-\mathrm{e}^{i\omega_{kr}\tau}\ \overline{\langle\varphi_n|\,W(\tau)\,|\varphi_p\rangle\langle\varphi_r|\,W(0)\,|\varphi_k\rangle}\ \langle\varphi_p|\,\rho_N(t)\,|\varphi_r\rangle \\ &-\mathrm{e}^{i\omega_{pn}\tau}\ \overline{\langle\varphi_n|\,W(0)\,|\varphi_p\rangle\langle\varphi_r|\,W(\tau)\,|\varphi_k\rangle}\ \langle\varphi_p|\,\rho_N(t)\,|\varphi_r\rangle \\ &+\mathrm{e}^{i\omega_{rp}\tau}\,\langle\varphi_n|\,\rho_N(t)\,|\varphi_p\rangle\ \overline{\langle\varphi_p|\,W(0)\,|\varphi_r\rangle\langle\varphi_r|\,W(\tau)\,|\varphi_k\rangle}\ \Big\}\end{aligned} \tag{36}$$

Noting P the dimension of the state space, the previous relations (33) or (36) yield P^2 differential equations that govern the time evolution of the matrix elements of $\widetilde{\rho}_N(t)$ or of $\rho_N(t)$. These differential equations are coupled with each other; their coefficients are time integrals of correlation functions of the perturbation, which are supposedly known for a given physical problem.

β. Short memory approximation

In view of the approximations we used, let us find under which conditions our calculations are consistent. The general validity condition is that there exists, for each time t, a previous time t_i such that the interval $t - t_i$ obeys two conditions: it must be simultaneously very long compared to the correlation time and very short compared to the evolution time in the interaction picture.

We can evaluate this evolution time by using an approximate expression of relation (33). We introduced in (7) the mean square of the matrix element $W_{nk}(t)$. Let us call H^2 the order of magnitude of such a mean square for the various values of n and k. The coefficients that multiply $\widetilde{\rho}_N(t)$ on the right-hand side of (33) can be replaced by this factor H^2, integrated over $\mathrm{d}\tau$. Taking (12) into account, this integral introduces a factor τ_c. The coefficients can thus be approximated by:

$$-\frac{H^2\tau_c}{\hbar^2} \tag{37}$$

With this approximation, the evolution equation (33) yields an evolution time of the order of $\hbar^2/H^2\tau_c$. Our computations are consistent if this time is much larger than τ_c, that is if:

$$H^2(\tau_c)^2 \ll \hbar^2 \tag{38}$$

In other words, our computations are valid if the correlation (memory) time of the perturbation is short compared to the characteristic time of its intensity, $\hbar/H$. This means that the perturbation will frequently change its value (and sign) before it can significantly change the system. This validity condition is often called the "motional narrowing condition", for a reason explained in § 2-c-δ.

Relations (33) or (36) are sets of first order differential equations. They describe the exponential relaxation of all the populations towards a situation where they all become equal. Carrying out calculations with these equations is not particularly difficult. However, it leads to the writing of complicated equations, in particular due to the large number of indices involved. In the general case, the populations $\langle\varphi_n|\,\rho_N(t)\,|\varphi_n\rangle$ are not only coupled to each other, but also to non-diagonal elements $\langle\varphi_n|\,\rho_N(t)\,|\varphi_k\rangle$ with $n \neq k$. All the matrix elements of the density operator can a priori be coupled to each other. One must then use additional approximations to select the terms essential for determining the relaxation properties.

In this complement, we shall only consider a simple particular case, that still allows us to develop a large number of physical concepts: the study of an ensemble of spin 1/2's undergoing an statistically isotropic perturbation.

2. Relaxation of an ensemble of spin 1/2's

Consider an ensemble of spin 1/2's, contained for example in a sample measured in a magnetic resonance experiment, such as the one mentioned in the introduction. The evolution equations then take on a simple form, easy to interpret. There are only two levels $|\varphi_i\rangle$, which will be noted $|+\rangle$ and $|-\rangle$. Their energy difference is:

$$\hbar(\omega_+ - \omega_-) = \hbar\omega_0 \tag{39}$$

It is useful to characterize the density operator of the spins by the average value of their angular momentum, which amounts to expanding this density operator on Pauli matrices (§ 5 of Complement $\mathrm{E_{IV}}$).

2-a. Characterization of the operators, isotropy of the perturbation

All the operators appearing in the previous equations now act in a 2-dimensional space. They are represented by matrices that can be expanded on the three Pauli matrices σ_x, σ_y and σ_z, as well as on the identity matrix, as shown in relation (22) of Complement $\mathrm{A_{IV}}$.

α. *Transformation of the operators*

We set:

$$\rho_N(t) = \frac{1}{2}\left[1 + \mathbf{M}(t)\cdot\sigma\right] \tag{40}$$

where σ stands for the vector operator whose components are the three Pauli matrices. The components of the vector $\mathbf{M}(t)$ are three real numbers that play the role of parameters defining $\rho_N(t)$. As we now show, the vector $\mathbf{M}(t)$ is simply the mean value of σ over the whole sample, whose total magnetization is thus proportional to $\mathbf{M}(t)$. Relation (11) of Complement $\mathrm{A_{IV}}$ indicates that:

$$\sigma_j\sigma_k = i\sum_l \varepsilon_{jkl}\,\sigma_l + \delta_{jk}\mathbb{1} \tag{41}$$

where ε_{jkl} is equal to zero if two indices are equal, equal to $+1$ if the series of indices jkl is an even permutation of the three axes Ox, Oy and Oz, and equal to -1 if the permutation is odd. It follows that the trace of a product of Pauli matrices is zero, unless two of the matrices are identical (in which case the trace equals 2). Consequently, the average value of the operator σ_j is given by:

$$\begin{aligned}\langle\sigma_j\rangle(t) &= \mathrm{Tr}\left\{\sigma_j\ \rho_N(t)\right\} = \frac{1}{2}\mathrm{Tr}\left\{[\sigma_j]^2\, M_j(t)\right\}\\ &= M_j(t)\end{aligned} \tag{42}$$

The operator H_0 is written in a form[2] similar to that of relation (12) in Complement $\mathrm{F_{IV}}$, which studies a magnetic resonance experiment:

$$H_0 = \frac{\hbar\omega_0}{2}\sigma_z \tag{43}$$

This operator corresponds to the effect of a magnetic field parallel to the Oz axis, which induces a rotation of the spins around that axis at the angular frequency ω_0. The unitary operator $U_0(t,t_i)$ defined in (14) is now a rotation operator of the spins (Complement $\mathrm{A_{IV}}$) through an angle $\omega_0(t-t_i)$; as for the adjoint operator $U_0^\dagger(t,t_i)$, it is also a rotation operator, but through an opposite angle.

[2] We did not give the operators H_0 or $W_n(t)$ a component on the identity operator, since it would not alter the commutators where these operators come into play.

The interaction operator can be written in a similar way:

$$W(t) = \frac{1}{2}\mathbf{h}(t) \cdot \sigma \tag{44}$$

where $\mathbf{h}(t)$ is a random vector function that characterizes the perturbation acting on the spins. In the same way as $W(t)$ was defined (§ 1-a) as a random choice amongst the N possible outcomes of the individual system labeled by the index q, $\mathbf{h}(t)$ characterizes the statistical properties of the three components of the local field acting on each spin at time t. Contrary to the field associated with H_0, this local field is random and can point in any direction, not necessarily parallel to the Oz axis. The three components $k = 1$, 2, 3 of this vector are noted $h_k(t)$. For ensemble averages, we assume that $\overline{h_k(t)} = 0$ and, as in (9) and (12), we shall write the correlation functions in the following way:

$$\overline{h_k(t+\tau)\, h_l(t)} = Q_{kl}(\tau) \tag{45}$$

where the $Q_{kl}(\tau)$ are rapidly decreasing functions of τ over times of the order of τ_c. The coefficients $Q_{kk}(\tau)$ are auto-correlation functions of the various components of $\mathbf{h}(t)$, the $Q_{kl}(\tau)$ for $k \neq l$ are cross-correlation functions pertaining to two different components.

β. Isotropy

We introduce an additional hypothesis, and assume that the perturbation affecting the spins is statistically isotropic: the correlation functions of the components of $\mathbf{h}(t)$ have no preferred direction. This means that the correlation functions $Q_{kk}(\tau)$ are identical for the three axes Ox, Oy and Oz (corresponding to $k = 1$, 2 and 3 respectively), whatever the value of τ. In other words, the ensemble of the $Q_{kl}(\tau)$ form a 3×3 matrix, which is rotation invariant, hence necessarily proportional to the unit matrix. Consequently, not only the auto-correlation coefficients are equal to each other, but also the cross-correlation coefficients Q_{kl} for $k \neq l$ must be equal to zero. Added to the stationarity of the perturbation, this hypothesis leads to:

$$\overline{h_k(t+\tau)\, h_l(t)} = \delta_{kl}\; Q(\tau) \tag{46}$$

One frequently models the decrease of $Q(\tau)$ with τ by a simple exponential, with a time constant τ_c. This leads to:

$$\overline{h_k(t+\tau)\, h_l(t)} = \delta_{kl}\; \overline{[h_k(0)]^2}\; e^{-\tau/\tau_c} \tag{47}$$

2-b. Longitudinal relaxation

When $|\varphi_n\rangle = |\varphi_k\rangle = |+\rangle$, the first term on the right-hand side of equation (36) in $(\omega_k - \omega_n)$ cancels out (no evolution of the populations due to H_0). In the following terms, we must replace $W(\tau)$ and $W(0)$ by their expression given by (44). This leads to products of matrix elements of two Pauli matrices σ_k and σ_l, multiplied by the statistical average (46) of the perturbation. Because of the factor δ_{kl} we must choose the same Pauli matrix in both operators $W(\tau)$ and $W(0)$.

α. Calculation of the relaxation time

Let us start by choosing twice the matrix σ_z. Since this matrix does not couple the states $|+\rangle$ and $|-\rangle$, we must have $|\varphi_p\rangle = |\varphi_r\rangle = |+\rangle$ in relation (36). As a result, the sum of the last 4 terms on the right-hand side is zero.

We then choose twice the matrix σ_x, which only has non-diagonal elements, all equal to 1. In the second line of the right-hand side of (36), when $n = k = +$, we must necessarily have $|\varphi_p\rangle = |-\rangle$ and $|\varphi_r\rangle = |+\rangle$, whereas in the fifth line, the opposite is true ($|\varphi_p\rangle = |+\rangle$ and $|\varphi_r\rangle = |-\rangle$); for the third and fourth lines we have $|\varphi_p\rangle = |\varphi_r\rangle = |-\rangle$. This yields the following term:

$$-\frac{1}{4\hbar^2}\int_0^\infty \mathrm{d}\tau\ \overline{h_x(t)\,h_x(t+\tau)}\left\{e^{i\omega_0\tau}\ \langle+|\,\rho_N(t)\,|+\rangle - e^{-i\omega_0\tau}\,\langle-|\,\rho_N(t)\,|-\rangle\right.$$
$$\left. -e^{i\omega_0\tau}\,\langle-|\,\rho_N(t)\,|-\rangle + e^{-i\omega_0\tau}\ \langle+|\,\rho_N(t)\,|+\rangle\right\} \tag{48}$$

or:

$$\frac{1}{2\hbar^2}\int_0^\infty \mathrm{d}\tau\ \overline{h_x(t)\,h_x(t+\tau)}\ \cos\omega_0\tau\ \left[\langle-|\,\rho_N(t)\,|-\rangle - \langle+|\,\rho_N(t)\,|+\rangle\right] \tag{49}$$

We finally choose twice the matrix σ_y, which has the same structure, with two matrix elements equal to $+i$ and $-i$, so that their product is also equal to unity. This term is the same as the σ_x term, except that we must replace h_x by h_y.

We finally obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle+|\,\rho_N(t)\,|+\rangle = \frac{1}{2T_1}\Big[\langle-|\,\rho_N(t)\,|-\rangle - \langle+|\,\rho_N(t)\,|+\rangle\Big] \tag{50}$$

with:

$$\frac{1}{T_1} = \frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau\ \left[\overline{h_x(t)\,h_x(t+\tau)} + \overline{h_y(t)\,h_y(t+\tau)}\right]\ \cos\omega_0\tau \tag{51}$$

The time T_1 is called the "longitudinal relaxation time". Its properties are discussed below.

The calculation of the evolution of the other diagonal element $\langle-|\,\rho_N(t)\,|-\rangle$ is practically the same and yields:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle-|\,\rho_N(t)\,|-\rangle = \frac{1}{2T_1}\Big[\langle+|\,\rho_N(t)\,|+\rangle - \langle-|\,\rho_N(t)\,|-\rangle\Big] \tag{52}$$

Now using (42) we can write the evolution of the Oz component of the magnetization:

$$\frac{\mathrm{d}}{\mathrm{d}t}M_z(t) = \frac{\mathrm{d}}{\mathrm{d}t}Tr\left\{\sigma_z\ \rho_N(t)\right\} = \frac{\mathrm{d}}{\mathrm{d}t}\Big[\langle+|\,\rho_N(t)\,|+\rangle - \langle-|\,\rho_N(t)\,|-\rangle\Big] \tag{53}$$

Taking the difference between (50) and (52) leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}M_z(t) = -\frac{1}{T_1}M_z(t) \tag{54}$$

This equation shows that the longitudinal (parallel to the static magnetic field) component of $\mathbf{M}(t)$ decreases exponentially with a time constant T_1, and tends toward zero when $t \to \infty$. The relaxation rate $1/T_1$ depends on the sum of correlation functions of both transverse (perpendicular to Oz) components of the perturbation. This was to be expected since it is the operators σ_x and σ_y that can induce transitions of the spins between their levels $|+\rangle$ and $|-\rangle$.

β. Role of the spectral density

The dependence in ω_0 of the relaxation probability $1/T_1$ can be interpreted in view of the results of Chapter XIII (§ C-2), where we studied a sinusoidal perturbation coupling two levels $|+\rangle$ and $|-\rangle$. We showed that the closer the perturbation frequency ω is to the Bohr frequency ω_0 associated with the energy difference between the two levels, the more effective the perturbation. In our present case, the perturbation is not a sinusoid but a random function. We thus expect the probability amplitude of the transition to involve the ω_0 Fourier component of the perturbation that acts between the instants 0 and $t \gg \tau_c$.

To further examine this idea, we introduce the Fourier transform $I_{kk}(\omega)$ of the correlation function, called the "spectral density":

$$\begin{aligned} I_{kk}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}\tau \; e^{-i\omega\tau} \; \overline{h_k(t)\, h_k(t-\tau)} \\ \overline{h_k(t)\, h_k(t-\tau)} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}\omega \; e^{i\omega\tau} \; I_{kk}(\omega) \end{aligned} \tag{55}$$

As the random function is stationary, we have:

$$\overline{h_k(t)\, h_k(t-\tau)} = \overline{h_k(t+\tau)\, h_k(t)} = \overline{h_k(t)\, h_k(t+\tau)} \tag{56}$$

This shows that the correlation function is an even function. As the Fourier transform of an even and real function is also even and real (Appendix I), we can write:

$$I_{kk}(-\omega) = I_{kk}(\omega) \tag{57}$$

Relation (51) can be rewritten as:

$$\frac{1}{T_1} = \frac{1}{2\sqrt{2\pi}\hbar^2} \int_0^{\infty} \mathrm{d}\tau \int_{-\infty}^{+\infty} \mathrm{d}\omega \; e^{i\omega\tau} \left[I_{xx}(\omega) + I_{yy}(\omega) \right] \; \left[e^{i\omega_0\tau} + e^{-i\omega_0\tau} \right] \tag{58}$$

Taking (57) into account, changing simultaneously the signs of the two integration variables τ and ω does not change the function to be integrated. It merely transforms the first integral over $\mathrm{d}\tau$ to an integral between 0 and $-\infty$, while the limits of the second become $+\infty$ and $-\infty$. Since two sign changes cancel each other, we can reverse the limits in each of the two integrals. This leads to:

$$\begin{aligned} \frac{1}{T_1} &= \frac{1}{2\sqrt{2\pi}\hbar^2} \int_{-\infty}^{0} \mathrm{d}\tau \int_{-\infty}^{+\infty} \mathrm{d}\omega \; e^{i\omega\tau} \left[I_{xx}(\omega) + I_{yy}(\omega) \right] \; \left[e^{i\omega_0\tau} + e^{-i\omega_0\tau} \right] \\ &= \frac{1}{4\sqrt{2\pi}\hbar^2} \int_{-\infty}^{\infty} \mathrm{d}\tau \int_{-\infty}^{+\infty} \mathrm{d}\omega \; e^{i\omega\tau} \left[I_{xx}(\omega) + I_{yy}(\omega) \right] \; \left[e^{i\omega_0\tau} + e^{-i\omega_0\tau} \right] \end{aligned} \tag{59}$$

Note that the final form (second line) of this relation was obtained by adding the first line of this relation to relation (58), and dividing by two. The integral over $\mathrm{d}\tau$ leads to:

$$\int_{-\infty}^{\infty} \mathrm{d}\tau \left[e^{i(\omega+\omega_0)\tau} + e^{i(\omega-\omega_0)\tau} \right] = 2\pi \left[\delta(\omega+\omega_0) + \delta(\omega-\omega_0) \right] \tag{60}$$

Taking again (57) into account, the two terms in $\delta(\omega+\omega_0)$ and $\delta(\omega-\omega_0)$ yield the same contribution, and we obtain:

$$\frac{1}{T_1} = \sqrt{\frac{\pi}{2}} \left[\frac{I_{xx}(\omega_0) + I_{yy}(\omega_0)}{\hbar^2} \right] \tag{61}$$

The transition probability is proportional to the sum of the spectral densities of the two perturbations responsible for the transitions. As expected, it is the resonant components of the perturbation that induce the transitions between the states $|+\rangle$ and $|-\rangle$.

γ. Exponential correlation function

The correlation functions of these components are often modeled by a simple exponential, as in (47). In that case, the integral over $\mathrm{d}\tau$ of (51) is easy to compute:

$$\begin{aligned} \int_0^\infty \mathrm{d}\tau \; \overline{h_x(t)\, h_x(t+\tau)} \left[e^{i\omega_0\tau} + e^{-i\omega_0\tau} \right] &= \overline{[h_x(0)]^2} \int_0^\infty \mathrm{d}\tau \; e^{-\tau/\tau_c} \left[e^{i\omega_0\tau} + e^{-i\omega_0\tau} \right] \\ &= -[h_x(0)]^2 \left[\frac{1}{i\omega_0 - 1/\tau_c} + \frac{1}{-i\omega_0 - 1/\tau_c} \right] \\ &= [h_x(0)]^2 \, \frac{2\tau_c}{1+(\omega_0\tau_c)^2} \end{aligned} \tag{62}$$

Adding the term corresponding to the effect of the h_y component, we get:

$$\frac{1}{T_1} = \frac{1}{\hbar^2} \left[\overline{[h_x(0)]^2} + \overline{[h_y(0)]^2} \right] \frac{\tau_c}{1+(\omega_0\tau_c)^2} \tag{63}$$

The longitudinal relaxation rate varies as a Lorentzian function of ω_0, plotted in Fig. 1.

The relaxation rate is maximum when $\omega_0 = 0$ (zero static field), and is equal to:

$$\frac{1}{T_1(\omega_0=0)} = \frac{1}{\hbar^2} \left[\overline{[h_x(0)]^2} + \overline{[h_y(0)]^2} \right] \tau_c \tag{64}$$

With our present notation, the motional narrowing condition (38) is written:

$$\overline{[h^{x,y}(0)]^2} \; (\tau_c)^2 \ll \hbar^2 \tag{65}$$

This leads to:

$$\frac{1}{T_1} \ll \frac{1}{\tau_c} \tag{66}$$

We thus verify that T_1, the characteristic evolution time of $M_z(t)$, is very long compared to the correlation time τ_c, hence proving the consistency of the approximations we have used.

2-c. Transverse relaxation

We now study the evolution of the non-diagonal elements of the density matrix between the states $|+\rangle$ and $|-\rangle$. Let us first show that the matrix element $\langle +|\, \rho_N(t)\, |-\rangle$, or its complex conjugate $\langle -|\, \rho_N(t)\, |+\rangle$, characterizes the transverse components $M_x(t)$

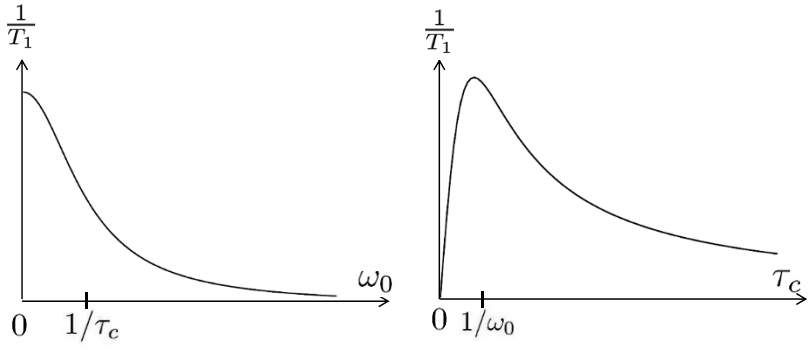


Figure 1: Plot of the longitudinal relaxation rate $1/T_1$ as a function of the energy difference $\hbar\omega_0$ between the energy levels (left-hand side of the figure), or as a function of the correlation time (right-hand side of the figure). This rate is proportional to the power spectrum of the perturbation at the frequency ω_0, and hence follows a Lorentzian function when plotted as a function of ω_0 – cf. relation (63). In the regime where $\omega_0 \gg 1/\tau_c$, the power spectrum of the perturbation decreases as $1/\omega_0^2$: the relaxation rate can be greatly reduced by increasing ω_0. If we now keep ω_0 fixed and increase the correlation time τ_c, we first get a linear variation of the relaxation probability, proportional to the time τ_c during which the perturbation acts in a coherent way. The probability then reaches a maximum for $\omega_0 = 1/\tau_c$, followed by a decrease in $1/\tau_c$ as the Fourier components of the perturbation at the frequency ω_0 become weaker and weaker.

and $M_y(t)$ of $\mathbf{M}(t)$. Using the expressions of the Pauli matrices – *cf.* for example relations (2) of Complement A_{IV} – we can compute the difference $\sigma_x - i\sigma_y$:

$$\sigma_x - i\sigma_y = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \tag{67}$$

which leads to:

$$\langle \sigma_x - i\sigma_y \rangle = Tr\left\{(\sigma_x - i\sigma_y)\,\rho_N(t)\right\} = 2\,\langle + |\,\rho_N(t)\,| - \rangle \tag{68}$$

We saw in § 2-a that the three components of $\mathbf{M}(t)$ were equal to the average values of the corresponding Pauli matrices. It thus follows:

$$\frac{M_x(t) - iM_y(t)}{2} = \langle + |\,\rho_N(t)\,| - \rangle \tag{69}$$

The real part of the non-diagonal matrix element $\langle + |\,\rho_N(t)\,| - \rangle$ directly yields $M_x(t)/2$, whereas its imaginary part yields the opposite of $M_y(t)/2$.

To avoid taking into account the evolution of this matrix element due to H_0, which introduces the first term in $(\omega_k - \omega_n)$ on the right-hand side of (36), we shall use the interaction picture. The evolution of $\langle + |\,\widetilde{\rho}_N(t)\,| - \rangle$ is then given by (33). In this relation, we replace the interaction operators $W(t)$ and $W(0)$ by their expression (44). As before, the statistical isotropy of the perturbation leads us to only keep the terms where the same component of $\mathbf{h}(t)$ appears in the two operators. We shall first examine the case where this component is either $h_x(t)$, or $h_y(t)$; the case where this component is $h_z(t)$ will be examined later.

α. Effect of the transverse components of the perturbation

We successively take into account the components on the Ox and Oy axes.

(*i*) The Ox component of the perturbation introduces in (33) the matrix elements of σ_x, which are all non-diagonal and equal to one. Each matrix element of W changes a ket $|+\rangle$ into a bra $\langle -|$, or vice versa. The Pauli matrix simply introduces a coefficient equal to 1.

In the first term on the right-hand side of (33), when $n = +$, we have $p = -$ and hence $r = +$; this term couples $\langle +| \widetilde{\rho}_N(t) |-\rangle$ to itself. As for the exponential, it introduces $e^{i\omega_0\tau}$. The same result is obtained for the fourth term: since $k = -$, we have $r = +$ and $p = -$, and the exponential again introduces the term $e^{i\omega_0\tau}$. The sum of these two terms yields:

$$-\frac{1}{2\hbar^2}\int_0^\infty \mathrm{d}\tau\ e^{i\omega_0\tau}\left[\overline{h_x(t)\,h_x(t-\tau)}\right] \times \langle +|\,\widetilde{\rho}_N(t)\,|-\rangle \tag{70}$$

The second and third terms on the right-hand side of (33) are different, since if $n = +$ we have $p = -$, whereas if $k = -$ we have $r = +$, which introduces on the right-hand side the matrix element $\langle -|\,\widetilde{\rho}_N(t)\,|+\rangle$. This means there is a term that couples two complex conjugate matrix elements:

$$-\frac{e^{2i\omega_0(t-t_i)}}{2\hbar^2}\int_0^\infty \mathrm{d}\tau\ e^{-i\omega_0\tau}\ \left[\overline{h_x(t)\,h_x(t-\tau)}\right]\ \times \langle -|\,\widetilde{\rho}_N(t)\,|+\rangle \tag{71}$$

(*ii*) The Oy component of the perturbation introduces in (33) the matrix elements of σ_y, also non-diagonal but now equal to $\pm i$. For the first and fourth term on the right-hand side of (33), this introduces in the previous calculation a factor $(i)(-i) = 1$, which does not change anything. As for the second and third term, the factor equals $(\pm i)^2 = -1$, which changes the sign of the result. Since the isotropy requires the correlation functions of $h_x(t)$ and of $h_y(t)$ to be equal, we obtain the opposite of (71), and both terms cancel out.

(*iii*) We finally get:

$$-\frac{1}{2\hbar^2}\int_0^\infty \mathrm{d}\tau\ e^{i\omega_0\tau}\ \left[\overline{h_x(t)\,h_x(t-\tau)} + \overline{h_y(t)\,h_y(t-\tau)}\right]\ \times \langle +|\,\widetilde{\rho}_N(t)\,|-\rangle \tag{72}$$

Expanding $e^{i\omega_0\tau}$ into $\cos(\omega_0\tau) + i\sin(\omega_0\tau)$, and taking (51) into account, we get:

$$-\left[\frac{1}{2T_1} + i\Delta\right] \times \langle +|\,\widetilde{\rho}_N(t)\,|-\rangle \tag{73}$$

where the coefficient Δ is defined by:

$$\Delta = \frac{1}{2\hbar^2}\int_0^\infty \mathrm{d}\tau\ \sin(\omega_0\tau)\ \left[\overline{h_x(t)\,h_x(t-\tau)} + \overline{h_y(t)\,h_y(t-\tau)}\right] \tag{74}$$

The physical significance of this coefficient is discussed below.

β. Effect of the longitudinal component of the perturbation

The Oz component of the perturbation does not change the spin state. As a result, the first line in (33) contains $n = p = r = +$ and $k = -$; the exponentials and the matrix elements of σ_z are all equal to unity. In the fourth line, $n = +$ and $p = r = k = -$, the exponentials are again equal to unity, as is the product of two matrix elements of σ_z (each equal to -1), and the final result is the same. As for the second and third lines, we have $n = p = +$ and $k = r = -$, so that the exponentials are equal to 1, whereas the product of the matrix elements of σ_z is now equal to -1; these two terms double the two preceding terms. Taking into account the factor $1/2$ of (44), we finally obtain the contribution:

$$-\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau \ \left[\overline{h_z(t)\, h_z(t-\tau)}\right] \times \langle + |\, \widetilde{\rho}_N(t)\, | - \rangle \tag{75}$$

that leads to the coefficient of transverse relaxation:

$$\frac{1}{T_2^{zz}} = \frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau \ \left[\overline{h_z(t)\, h_z(t-\tau)}\right] \tag{76}$$

γ. Discussion, role of the spectral density

Grouping together both contributions (73) and (75), we can write the complete evolution of the non-diagonal element as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle + |\, \widetilde{\rho}_N(t)\, | - \rangle = -\left[\frac{1}{2T_1} + \frac{1}{T_2^{zz}} + i\Delta\right]\langle + |\, \widetilde{\rho}_N(t)\, | - \rangle \tag{77}$$

Leaving the interaction picture to go back to the density operator $\rho_N(t)$ in the usual laboratory picture, we must add to this evolution the first term appearing on the right-hand side of (36); this yields:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle + |\, \rho_N(t)\, | - \rangle = -\left[\frac{1}{2T_1} + \frac{1}{T_2^{zz}} + i(\omega_0 + \Delta)\right] \times \langle + |\, \rho_N(t)\, | - \rangle \tag{78}$$

(i) Damping

In either picture, the non-diagonal element of the density matrix is damped with a time constant T_2 given by:

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_2^{zz}} \tag{79}$$

which is the sum of two contributions.

– The first is directly related to the longitudinal relaxation process. This process changes the distribution of the populations between the two levels + and −, hence destroying the coherence between these two levels. This rate of destruction of the coherence is the same as the one affecting the populations in (50), but half the one affecting $M_z(t)$ in (54). Note that it is only the transverse components of the perturbation that play a role in this contribution, since they are the ones that can induce transitions between the two levels.

– The second contribution comes only from the longitudinal component of the perturbation. This fluctuating component directly modifies the energy difference $\hbar\omega_0$ between the two levels + and −, and hence the precession velocity of the spins' transverse component. When the different spins have different precession velocities around the Oz axis, their transverse components spread out and their vector sum diminishes. This leads to a decrease of the transverse component of the global spin of the system.

(ii) Frequency shift

Relation (78) shows that the term in Δ is equivalent to a change in the precession frequency ω_0 of the spins. In addition to the damping associated with the relaxation, the perturbation introduces a shift in the evolution frequency of the non-diagonal elements.

The same calculation as the one leading to (58) yields for the frequency shift Δ:

$$\Delta = \frac{1}{4i\sqrt{2\pi}\hbar^2}\int_0^\infty \mathrm{d}\tau \int_{-\infty}^{+\infty} \mathrm{d}\omega\ e^{i\omega\tau}\left[I_{xx}(\omega)+I_{yy}(\omega)\right]\ \left[e^{i\omega_0\tau}-e^{-i\omega_0\tau}\right] \tag{80}$$

This expression contains an integral over $\mathrm{d}\tau$:

$$\frac{1}{i}\int_0^\infty \mathrm{d}\tau \left[e^{i(\omega+\omega_0)\tau}-e^{i(\omega-\omega_0)\tau}\right] = \left[\mathcal{P}\frac{1}{\omega+\omega_0}-\mathcal{P}\frac{1}{\omega-\omega_0}\right]-i\pi\left[\delta(\omega+\omega_0)-\delta(\omega-\omega_0)\right] \tag{81}$$

As the functions $I_{xx}(\omega)$ and $I_{yy}(\omega)$ are even, the terms containing the delta functions cancel out, whereas the terms containing the principal parts double each other. This leads to:

$$\Delta = -\sqrt{\frac{\pi}{2}}\frac{1}{\hbar^2}\int_{-\infty}^{+\infty}\mathrm{d}\omega\ \mathcal{P}\frac{1}{\omega-\omega_0}\left[I_{xx}(\omega)+I_{yy}(\omega)\right] \tag{82}$$

Note that, contrary to the longitudinal relaxation characterized by the time T_1, it is not the power spectrum of the perturbation at the resonant frequency $\omega=\omega_0$ that plays a role. Only non-resonant frequencies contribute to the shift.

δ. Exponential correlation functions

When the correlation function is modeled by an exponential as in relation (47), equalities (74) and (76) become:

$$\frac{1}{T_2^{zz}} = \frac{1}{\hbar^2}\overline{\left[h_z(0)\right]^2}\tau_c \tag{83}$$

and:

$$\Delta = \frac{1}{2\hbar^2}\overline{\left[h_x(0)\right]^2+\left[h_x(0)\right]^2}\ \frac{\omega_0\tau_c}{1+(\omega_0\tau_c)^2} \tag{84}$$

It is interesting to discuss the effect of the fluctuations of the perturbation on the relaxation time T_2^{zz}. Imagine first that the ensemble of spins is placed in a magnetic field along Oz, which does not change in time but has a different value for each spin. We

note H the root mean square of the corresponding fluctuation of the Hamiltonian. If the spins are initially oriented in the same direction, their transverse orientations will start spreading around Oz since each spin has a different precession velocity. The average value of the global transverse orientation of the sample will go to zero over a time of the order $\hbar/H$. This means that the transverse orientation diminishes at a rate of the order of $H/\hbar$, which depends linearly on the perturbation amplitude H. Now in the presence of time dependent fluctuations of the perturbation, result (83) predicts a totally different behavior. The relaxation rate $1/T_2^{zz}$ is of the order of $H^2\tau_c/\hbar^2$ and hence varies as the square of the perturbation amplitude. This evolution rate in the presence of fluctuations is thus $H\tau_c/\hbar$ times the evolution rate in the static case. As the factor $H\tau_c/\hbar \ll 1$ – see the motional narrowing condition (38) – the quadratic relaxation is much slower than the relaxation in the absence of fluctuation. This effect is even stronger when τ_c is shorter, which shows that it is the rapidly changing fluctuations that are responsible for the decrease of the relaxation rate.

Now the width of the magnetic resonance lines is an increasing function of the transverse relaxation rate[3]. Consequently, the shorter the correlation time, the narrower the lines. It often happens that the perturbation fluctuations come from the spins' motion in the sample, in which case the more rapid the motion, the narrower the magnetic resonance lines. This explains the origin of the expression "motional narrowing".

On the other hand, comparing relations (63) and (84) shows that the relaxation probability and the frequency shift have a very different dependence on $\omega_0\tau_c$. The relaxation probability follows a Lorentzian function, with a maximum for $\omega_0\tau_c = 0$, whereas the frequency shift is maximum for $\omega_0\tau_c = 1$. This difference comes from the fact that, as we discussed at the end of § 2-c-γ, it is not the resonant but rather the non-resonant frequencies of the spectral density that determine the frequency shift.

3. Conclusion

As mentioned in the introduction, there are many situations where an ensemble of individual quantum systems is subjected to a random perturbation with a correlation time τ_c very short compared to the other characteristic times of the problem. In a more general way than in §§ D and E of Chapter XIII, we examined in this complement how, in the limit where τ_c is too short for the perturbation to have an effect during that time, the perturbation no longer induces a Rabi type oscillation. This led us to introduce a transition probability between the levels, leading to an exponential (and not oscillating) evolution of the populations. Note that in Complement D_{XIII}, we also obtained, with the Fermi golden rule, a transition probability. In that case, it was the summation over the energies of all the final states that transformed the oscillation into a real and damped exponential. In the present complement, it is the random character of the perturbation that has a similar effect, even though the final state is unique and has a perfectly well defined energy. Another result we obtained concerns the existence of a frequency shift induced by the random perturbation. In the case of an optical excitation such as the one considered in § E-3-b of Chapter XIII, they are called "light shifts", and have numerous applications in atomic physics (Complement C_{XX}).

[3]Figure 7 of Complement F_{IV} shows the variation of these lines, assuming there exists only one longitudinal and transverse relaxation rate $1/T_R$.

Complement F_{XIII}

Exercises

1. Consider a one-dimensional harmonic oscillator of mass m, angular frequency ω_0 and charge q. Let $|\varphi_n\rangle$ and $E_n = (n + 1/2)\hbar\omega_0$ be the eigenstates and eigenvalues of its Hamiltonian H_0.

For $t < 0$, the oscillator is in the ground state $|\varphi_0\rangle$. At $t = 0$, it is subjected to an electric field "pulse" of duration τ. The corresponding perturbation can be written:

$$W(t) = \begin{cases} -q\mathcal{E}X & \text{for } 0 \leq t \leq \tau \\ 0 & \text{for } t < 0 \text{ and } t > \tau \end{cases}$$

$\mathcal{E}$ is the field amplitude and X is the position observable. Let $\mathscr{P}_{0n}$ be the probability of finding the oscillator in the state $|\varphi_n\rangle$ after the pulse.

a. Calculate $\mathscr{P}_{01}$ by using first-order time-dependent perturbation theory. How does $\mathscr{P}_{01}$ vary with τ, for fixed ω_0?

b. Show that, to obtain $\mathscr{P}_{02}$, the time-dependent perturbation theory calculation must be pursued at least to second order. Calculate $\mathscr{P}_{02}$ to this perturbation order.

c. Give the exact expressions for $\mathscr{P}_{01}$ and $\mathscr{P}_{02}$ in which the translation operator used in Complement F_V appears explicitly. By making a limited power series expansion in $\mathcal{E}$ of these expressions, find the results of the preceding questions.

2. Consider two spin 1/2's, $\mathbf{S}_1$ and $\mathbf{S}_2$, coupled by an interaction of the form $a(t)\mathbf{S}_1 \cdot \mathbf{S}_2$; $a(t)$ is a function of time which approaches zero when $|t|$ approaches infinity, and takes on non-negligible values (on the order of a_0) only inside an interval, whose width is of the order of τ, about $t = 0$.

a. At $t = -\infty$, the system is in the state $|+,-\rangle$ (an eigenstate of S_{1z} and S_{2z} with the eigenvalues $+\hbar/2$ and $-\hbar/2$). Calculate, without approximations, the state of the system at $t = +\infty$. Show that the probability $\mathscr{P}(+- \rightarrow -+)$ of finding, at $t = +\infty$, the system in the state $|-,+\rangle$ depends only on the integral $\int_{-\infty}^{+\infty} a(t)\,\mathrm{d}t$.

b. Calculate $\mathscr{P}(+- \rightarrow -+)$ by using first-order time-dependent perturbation theory. Discuss the validity conditions for such an approximation by comparing the results obtained with those of the preceding question.

c. Now assume that the two spins are also interacting with a static magnetic field $\mathbf{B}_0$ parallel to Oz. The corresponding Zeeman Hamiltonian can be written:

$$H_0 = -B_0(\gamma_1 S_{1z} + \gamma_2 S_{2z})$$

where γ_1 and γ_2 are the gyromagnetic ratios of the two spins, assumed to be different.

Assume that $a(t) = a_0\,\mathrm{e}^{-t^2/\tau^2}$. Calculate $\mathscr{P}(+- \rightarrow -+)$ by first-order time-dependent perturbation theory. Considering a_0 and τ as fixed, discuss the variation of $\mathscr{P}(+- \rightarrow -+)$ with respect to B_0.

3. Two-photon transitions between non-equidistant levels

Consider an atomic level of angular momentum $J = 1$, subject to static electric and magnetic fields, both parallel to Oz. It can be shown that three non-equidistant

energy levels are then obtained. The eigenstates $|\varphi_M\rangle$ of $J_z (M = -1, 0, +1)$, of energies E_M, correspond to them. We set $E_1 - E_0 = \hbar\omega_0$, $E_0 - E_{-1} = \hbar\omega'_0$ ($\omega_0 \neq \omega'_0$).

The atom is also subjected to a radiofrequency field rotating at the angular frequency ω in the xOy plane. The corresponding perturbation $W(t)$ can be written:

$$W(t) = \frac{\omega_1}{2}(J_+ e^{-i\omega t} + J_- e^{i\omega t})$$

where ω_1 is a constant proportional to the amplitude of the rotating field.

a. We set (notation identical to that of Chapter XIII):

$$|\psi(t)\rangle = \sum_{M=-1}^{+1} b_M(t) e^{-iE_M t/\hbar} |\varphi_M\rangle$$

Write the system of differential equations satisfied by the $b_M(t)$.

b. Assume that, at time $t = 0$, the system is in the state $|\varphi_{-1}\rangle$. Show that if we want to calculate $b_1(t)$ by time-dependent perturbation theory, the calculation must be pursued to second order. Calculate $b_1(t)$ to this perturbation order.

c. For fixed t, how does the probability $\mathscr{P}_{-1,+1}(t) = |b_1(t)|^2$ of finding the system in the state $|\varphi_1\rangle$ at time t vary with respect to ω? Show that a resonance appears, not only for $\omega = \omega_0$ and $\omega = \omega'_0$, but also for $\omega = (\omega_0 + \omega'_0)/2$. Give a particle interpretation of this resonance.

4. Returning to exercise 5 of Complement H_{XI} and using its notation, assume that the field $\mathbf{B}_0$ is oscillating at angular frequency ω, and can be written $\mathbf{B}_0(t) = \mathbf{B}_0 \cos\omega t$. Assume that $b = 2a$ and that ω is not equal to any Bohr angular frequency of the system (non-resonant excitation).

Introduce the susceptibility tensor χ, of components $\chi_{ij}(\omega)$, defined by:

$$\langle M_i \rangle(t) = \sum_j \mathrm{Re}\left[\chi_{ij}(\omega)\, B_{0j}\, e^{i\omega t}\right]$$

with $i, j = x, y, z$. Using a method analogous to the one in § 2 of Complement A_{XIII}, calculate $\chi_{ij}(\omega)$. Setting $\omega = 0$, find the results of exercise 5 of Complement H_{XI}.

5. The Autler-Townes effect

Consider a three-level system: $|\varphi_1\rangle$, $|\varphi_2\rangle$, and $|\varphi_3\rangle$, of energies E_1, E_2 and E_3. Assume $E_3 > E_2 > E_1$ and $E_3 - E_2 \ll E_2 - E_1$.

This system interacts with a magnetic field oscillating at the angular frequency ω. The states $|\varphi_2\rangle$ and $|\varphi_3\rangle$ are assumed to have the same parity, which is the opposite of that of $|\varphi_1\rangle$, so that the interaction Hamiltonian $W(t)$ with the oscillating magnetic field can connect $|\varphi_2\rangle$ and $|\varphi_3\rangle$ to $|\varphi_1\rangle$. Assume that, in the basis of the three states $|\varphi_1\rangle$, $|\varphi_2\rangle$, $|\varphi_3\rangle$, arranged in that order, $W(t)$ is represented by the matrix:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \omega_1 \sin\omega t \\ 0 & \omega_1 \sin\omega t & 0 \end{bmatrix}$$

where ω_1 is a constant proportional to the amplitude of the oscillating field.

a. Set (notation identical to that of Chapter XIII):

$$|\psi(t)\rangle = \sum_{i=1}^{3} b_i(t) \mathrm{e}^{-iE_i t/\hbar} |\varphi_i\rangle$$

Write the system of differential equations satisfied by the $b_i(t)$.

b. Assume that ω is very close to $\omega_{32} = (E_3 - E_2)/\hbar$. Making approximations analogous to those used in Complement $\mathrm{C_{XIII}}$, integrate the preceding system, with the initial conditions:

$$b_1(0) = b_2(0) = \frac{1}{\sqrt{2}} \qquad\qquad b_3(0) = 0$$

(neglect, on the right-hand side of the differential equations, the terms whose coefficients, $\mathrm{e}^{\pm i(\omega+\omega_{32})t}$, vary very rapidly, and keep only those whose coefficients are constant or vary very slowly, as $\mathrm{e}^{\pm i(\omega-\omega_{32})t}$).

c. The component D_z along Oz of the electric dipole moment of the system is represented, in the basis of the three states $|\varphi_1\rangle$, $|\varphi_2\rangle$, $|\varphi_3\rangle$, arranged in that order, by the matrix:

$$\begin{bmatrix} 0 & d & 0 \\ d & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where d is a real constant (D_z is an odd operator and can connect only states of different parities).

Calculate $\langle D_z \rangle(t) = \langle \psi(t)|D_z|\psi(t)\rangle$, using the vector $|\psi(t)\rangle$ calculated in *b*.

Show that the time evolution of $\langle D_z \rangle(t)$ is given by a superposition of sinusoidal terms. Determine the frequencies v_k and relative intensities π_k of these terms.

These are the frequencies that can be absorbed by the atom when it is placed in an oscillating electric field parallel to Oz. Describe the modifications of this absorption spectrum when, for ω fixed and equal to ω_{32}, ω_1 is increased from zero. Show that the presence of the magnetic field oscillating at the frequency $\omega_{32}/2\pi$ splits the electric dipole absorption line at the frequency $\omega_{21}/2\pi$, and that the separation of the two components of the doublet is proportional to the oscillating magnetic field amplitude (the Autler-Townes doublet).

What happens when, for ω_1 fixed, $\omega - \omega_{32}$ is varied?

6. Elastic scattering by a particle in a bound state. Form factor

Consider a particle (a) in a bound state $|\varphi_0\rangle$ described by the wave function $|\varphi_0(\mathbf{r}_a)$ localized about a point O. Towards this particle (a) is directed a beam of particles (b), of mass m, momentum $\hbar\mathbf{k}_i$, energy $E_i = \hbar^2\mathbf{k}_i^2/2m$ and wave function $\dfrac{1}{(2\pi)^{3/2}}\mathrm{e}^{i\mathbf{k}_i\cdot\mathbf{r}_b}$. Each particle (b) of the beam interacts with particle (a). The corresponding potential energy, W, depends only on the relative position $\mathbf{r}_b - \mathbf{r}_a$ of the two particles.

a. Calculate the matrix element:

$$\langle a : \varphi_0; b : \mathbf{k}_f | W(\mathbf{R}_b - \mathbf{R}_a) | a : \varphi_0; b : \mathbf{k}_i \rangle$$

of $W(\mathbf{R}_b - \mathbf{R}_a)$ between two states in which particle (a) is in the same state $|\varphi_0\rangle$ and particle (b) goes from the state $|\mathbf{k}_i\rangle$ to the state $|\mathbf{k}_f\rangle$. The expression for this matrix element should include the Fourier transform $\overline{W}(\mathbf{k})$ of the potential $W(\mathbf{r}_b - \mathbf{r}_a)$:

$$W(\mathbf{r}_b - \mathbf{r}_a) = \frac{1}{(2\pi)^{3/2}} \int \overline{W}(\mathbf{k})\, \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{r}_b - \mathbf{r}_a)} \mathrm{d}^3k$$

b. Consider the scattering processes in which, under the effect of the interaction W, particle (b) is scattered in a certain direction, with particle (a) remaining in the same quantum state $|\varphi_0\rangle$ after the scattering process (elastic scattering).

Using a method analogous to the one in Chapter XIII [*cf.* comment (ii) of § C-3-b], calculate, in the Born approximation, the elastic scattering cross section of particle (b) by particle (a) in the state $|\varphi_0\rangle$.

Show that this cross section can be obtained by multiplying the cross section for scattering by the potential $W(\mathbf{r})$ (in the Born approximation) by a factor which characterizes the state $|\varphi_0\rangle$, called the "form factor".

7. A simple model of the photoelectric effect

Consider, in a one-dimensional problem, a particle of mass m, placed in a potential of the form $V(x) = -\alpha\delta(x)$, where α is a real positive constant.

Recall (*cf.* exercises 2 and 3 of Complement $\mathrm{K_I}$) that, in such a potential, there is a single bound state, of negative energy $E_0 = -m\alpha^2/2\hbar^2$, associated with a normalized wave function $\varphi_0(x) = \sqrt{m\alpha/\hbar^2}\, \mathrm{e}^{-\frac{m\alpha}{\hbar^2}|x|}$. For each positive value of the energy $E = \hbar^2k^2/2m$, on the other hand, there are two stationary wave functions, corresponding, respectively, to an incident particle coming from the left or from the right. The expression for the first eigenfunction, for example, is:

$$\chi_k(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \left[\mathrm{e}^{ikx} - \dfrac{1}{1 + i\hbar^2k/m\alpha} \mathrm{e}^{-ikx} \right] & \text{for} \quad x < 0 \\ \dfrac{1}{\sqrt{2\pi}} \dfrac{i\hbar^2k/m\alpha}{1 + i\hbar^2k/m\alpha} \mathrm{e}^{ikx} & \text{for} \quad x > 0 \end{cases}$$

a. Show that the $\chi_k(x)$ satisfy the orthonormalization relation (in the extended sense):

$$\langle \chi_k | \chi_{k'} \rangle = \delta(k - k')$$

The following relation [*cf.* formula (47) of Appendix II] can be used:

$$\int_{-\infty}^{0} \mathrm{e}^{iqx}\, \mathrm{d}x = \int_{0}^{\infty} \mathrm{e}^{-iqx} \mathrm{d}x = \operatorname*{Lim}_{\varepsilon \to 0} \frac{1}{\varepsilon + iq}$$

$$= \pi\delta(q) - i\mathcal{P}\left(\frac{1}{q}\right)$$

Calculate the density of states $\rho(E)$ for a positive energy E.

b. Calculate the matrix element $\langle \chi_k | X | \varphi_0 \rangle$ of the position observable X between the bound state $|\varphi_0\rangle$ and the positive energy state $|\chi_k\rangle$ whose wave function was given above.

c. The particle, assumed to be charged (charge q) interacts with an electric field oscillating at the angular frequency ω. The corresponding perturbation is:

$$W(t) = -q\mathcal{E}X \sin \omega t$$

where $\mathcal{E}$ is a constant.

The particle is initially in the bound state $|\varphi_0\rangle$. Assume that $\hbar\omega > -E_0$. Calculate, using the results of § C of Chapter XIII [see, in particular, formula (C-37)], the transition probability w per unit time to an arbitrary positive energy state (the photoelectric or photoionization effect). How does w vary with ω and $\mathcal{E}$?

8. Disorientation of an atomic level due to collisions with rare gas atoms

Consider a motionless atom A at the origin of a coordinate frame $Oxyz$ (Fig. 1). This atom A is in a level of angular momentum $J = 1$, to which correspond the three orthonormal kets $|M\rangle$ ($M = -1, 0, +1$), eigenstates of J_z of eigenvalues $M\hbar$.

A second atom B, in a level of zero angular momentum, is in uniform rectilinear motion in the xOz plane: it is travelling at the velocity v along a straight line parallel to Oz and situated at a distance b from this axis (b is the "impact parameter"). The time origin is chosen at the time when B arrives at point H of the Ox axis ($OH = b$). At time t, atom B is therefore at point M, where $HM = vt$. Call θ the angle between Oz and OM.

The preceding model, which treats the external degrees of freedom of the two atoms classically, permits the simple calculation of the effect on the internal degrees of freedom of atom A (which are treated quantum mechanically) of a collision with atom B (which is, for example, a rare gas atom in the ground state). It can be shown that, because of the Van der Waals forces (*cf.* Complement C_{XI}) between the two atoms, atom A is

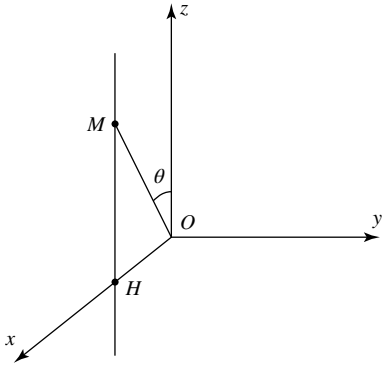


Figure 1

subject to a perturbation W acting on its internal degrees of freedom, and given by:

$$W = \frac{C}{r^6} J_u^2$$

where C is a constant, r is the distance between the two atoms, and J_u is the component of the angular momentum $\mathbf{J}$ of atom A on the OM axis joining the two atoms.

a. Express W in terms of $C, b, v, t, J_z, J_\pm = J_x \pm iJ_y$. Introduce the dimensionless parameter $\tau = vt/b$.

b. Assume that there is no external magnetic field, so that the three states $|+1\rangle$, $|0\rangle$, $|-1\rangle$ of atom A have the same energy.

Before the collision, that is, at $t = -\infty$, atom A is in the state $|-1\rangle$. Using first-order time-dependent perturbation theory, calculate the probability $\mathscr{P}_{-1,+1}$ of finding, after the collision (that is, at $t = +\infty$), atom A in the state $|+1\rangle$. Discuss the variation of $\mathscr{P}_{-1,+1}$ with respect to b and v. Similarly, calculate $\mathscr{P}_{-1,0}$.

c. Now assume that there is a static field $\mathbf{B}_0$ parallel to Oz, so that the three states $|M\rangle$ have an additional energy $M\hbar\omega_0$ (the Zeeman effect), where ω_0 is the Larmor angular frequency in the field $\mathbf{B}_0$.

α. With ordinary magnetic fields ($\mathbf{B}_0 \sim 10^2$ gauss), $\omega_0 \simeq 10^9$ rad.sec^{-1}; b is of the order of 5 Å, and v, of the order of 5×10^2 m.sec^{-1}. Show that, under these conditions, the results of question *b* remain valid.

β. Without going into detailed calculations, explain what happens for much higher values of B_0. Starting with what value of ω_0 (where b and v have the values indicated in α) will the results of *b* no longer be valid?

d. Without going into detailed calculations, explain how to calculate the disorientation probabilities $\mathscr{P}_{-1,+1}$ and $\mathscr{P}_{-1,0}$ for an atom A placed in a gas of atoms B in thermodynamic equilibrium at the temperature T, containing a number n of atoms per unit volume sufficiently small that only binary collisions need be considered.

N.B. We give: $\int_{-\infty}^{+\infty} \mathrm{d}\tau (1+\tau^2)^{-4} = 5\pi/16$

9. Transition probability per unit time under the effect of a random perturbation. Simple relaxation model

This exercice uses the results of Complement E_{XIII}. We consider a system of $\mathcal{N}$ spin 1/2 particles, with gyromagnetic ratio γ, placed in a static field $\mathbf{B}_0$ (set $\omega_0 = -\gamma B_0$). These particles are enclosed in a spherical cell of radius R. Each of them bounces constantly back and forth between the walls. The mean time between two collisions of the same particle with the wall is called the "flight time" τ_v. During this time, the particle "sees" only the field $\mathbf{B}_0$. In a collision with the wall, each particle remains adsorbed on the surface during a mean time τ_a ($\tau_a \ll \tau_v$), during which it "sees", in addition to $\mathbf{B}_0$, a constant microscopic magnetic field $\mathbf{b}$, due to the paramagnetic impurities contained in the wall. The direction of $\mathbf{b}$ varies randomly from one collision to another; the mean amplitude of $\mathbf{b}$ is denoted by b_0.

a. What is the correlation time of the perturbation seen by the spins? Give the physical justification for the following form, to be chosen for the correlation function of the components of the microscopic field $\mathbf{b}$:

$$\overline{b_x(t) b_x(t-\tau)} = \frac{1}{3} b_0^2 \frac{\tau_a}{\tau_v} \mathrm{e}^{-\tau/\tau_a}$$

with analogous expressions for the components along Oy and Oz, all the cross terms $\overline{b_x(t)b_y(t-\tau)}$... being zero.

b. Let $\mathcal{M}_z$ be the component along the Oz axis defined by the field $\mathbf{B}_0$ of the macroscopic magnetization of the $\mathcal{N}$ particles. Show that, under the effect of the collisions with the wall, $\mathcal{M}_z$ "relaxes" with a time constant T_1:

$$\frac{\mathrm{d}\mathcal{M}_z}{\mathrm{d}t} = -\frac{\mathcal{M}_z}{T_1}$$

(T_1 is called the longitudinal relaxation time). Calculate T_1 in terms of γ, B_0, τ_a, τ_v, b_0.

c. Show that studying the variation of T_1 with B_0 permits the experimental determination of the mean adsorption time τ_a.

d. We have at our disposition several cells, of different radii R, constructed from the same material. By measuring T_1, how can we determine experimentally the mean amplitude b_0 of the microscopic field at the wall?

10. Absorption of radiation by a many-particle system forming a bound state. The Doppler effect. Recoil energy. The Mössbauer effect

In Complement $\mathrm{A_{XIII}}$, we consider the absorption of radiation by a charged particle attracted by a fixed center O (the hydrogen atom model for which the nucleus is infinitely heavy). In this exercise, we treat a more realistic situation, in which the incident radiation is absorbed by a system of several particles of finite masses interacting with each other and forming a bound state. Thus, we are studying the effect on the absorption phenomenon of the degrees of freedom of the center of mass of the system.

I-Absorption of radiation by a free hydrogen atom. The doppler effect. Recoil energy

Let $\mathbf{R}_1$ and $\mathbf{P}_1$, $\mathbf{R}_2$ and $\mathbf{P}_2$ be the position and momentum observables of two particles, (1) and (2), of masses m_1 and m_2 and opposite charges q_1 and q_2 (a hydrogen atom). Let $\mathbf{R}$ and $\mathbf{P}$, $\mathbf{R}_G$ and $\mathbf{P}_G$ be the position and momentum observables of the relative particle and of the center of mass (*cf.* Chap. VII, § B). $M = m_1 + m_2$ is the total mass, and $m = m_1 m_2/(m_1 + m_2)$ is the reduced mass. The Hamiltonian H_0 of the system can be written:

$$H_0 = H_e + H_i \tag{1}$$

where:

$$H_e = \frac{1}{2M}\mathbf{P}_G^2 \tag{2}$$

is the translational kinetic energy of the atom, assumed to be free ("external" degrees of freedom), and where H_i (which depends only on $\mathbf{R}$ and $\mathbf{P}$) describes the internal energy of the atom ("internal" degrees of freedom). We denote by $|\mathbf{K}\rangle$ the eigenstates of H_e, with eigenvalues $\hbar^2\mathbf{K}^2/2M$. We concern ourselves with only two eigenstates of H_i, $|\chi_a\rangle$ and $|\chi_b\rangle$, of energies E_a and E_b ($E_b > E_a$). We set:

$$E_b - E_a = \hbar\omega_0 \tag{3}$$

a. What energy must be furnished to the atom to move it from the state $|\mathbf{K}; \chi_a\rangle$ (the atom in the state $|\chi_a\rangle$ with a total momentum $\hbar\mathbf{K}$) to the state $|\mathbf{K}'; \chi_b\rangle$?

b. This atom interacts with a plane electromagnetic wave of wave vector $\mathbf{k}$ and angular frequency $\omega = ck$, polarized along the unit vector $\mathbf{e}$ perpendicular to $\mathbf{k}$. The corresponding vector potential $\mathbf{A}(\mathbf{r}, t)$ is:

$$\mathbf{A}(\mathbf{r}, t) = \mathcal{A}_0 \, \mathbf{e} \, \mathrm{e}^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} + \text{c.c.} \tag{4}$$

where $\mathcal{A}_0$ is a constant. The principal term of the interaction Hamiltonian between this plane wave and the two-particle system can be written (*cf.* Complement A_{XIII}, § 1-b):

$$W(t) = -\sum_{i=1}^{2} \frac{q_i}{m_i} \mathbf{P}_i \cdot \mathbf{A}(\mathbf{R}_i, t) \tag{5}$$

Express $W(t)$ in terms of $\mathbf{R}$, $\mathbf{P}$, $\mathbf{R}_G$, $\mathbf{P}_G$, m, M and q (setting $q_1 = -q_2 = q$), and show that, in the electric dipole approximation which consists of neglecting $\mathbf{k} \cdot \mathbf{R}$ (but not $\mathbf{k} \cdot \mathbf{R}_G$) compared to 1, we have:

$$W(t) = W \, \mathrm{e}^{-i\omega t} + W^{\dagger} \, \mathrm{e}^{i\omega t} \tag{6}$$

where:

$$W = -\frac{q\mathcal{A}_0}{m} \, \mathbf{e} \cdot \mathbf{P} \, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{R}_G} \tag{7}$$

c. Show that the matrix element of W between the state $|\mathbf{K}; \chi_a\rangle$ and the state $|\mathbf{K}'; \chi_b\rangle$ is different from zero only if there exists a certain relation between $\mathbf{K}$, $\mathbf{k}$, $\mathbf{K}'$ (to be specified). Interpret this relation in terms of the total momentum conservation during the absorption of an incident photon by the atom.

d. Show from this that if the atom in the state $|\mathbf{K}; \chi_a\rangle$ is placed in the plane wave (4), resonance occurs when the energy $\hbar\omega$ of the photons associated with the incident wave differs from the energy $\hbar\omega_0$ of the atomic transition $|\chi_a\rangle \rightarrow |\chi_b\rangle$ by a quantity δ which is to be expressed in terms of $\hbar$, ω_0, $\mathbf{K}$, $\mathbf{k}$, M, c (since δ is a corrective term, we can replace ω by ω_0 in the expression for δ). Show that δ is the sum of two terms: one of which, δ_1, depends on $\mathbf{K}$ and on the angle between $\mathbf{K}$ and $\mathbf{k}$ (the Doppler effect); and the other, δ_2, is independent of $\mathbf{K}$. Give a physical interpretation of δ_1 and δ_2 (showing that δ_2 is the recoil kinetic energy of the atom when, having been initially motionless, it absorbs a resonant photon).

Show that δ_2 is negligible compared to δ_1 when $\hbar\omega_0$ is of the order of 10 eV (the domain of atomic physics). Choose, for M, a mass of the order of that of the proton $Mc^2 \simeq 10^9$ eV), and, for $|\mathbf{K}|$, a value corresponding to a thermal velocity at $T = 300°$ K. Would this still be true if $\hbar\omega_0$ were of the order of 10^5 eV (the domain of nuclear physics)?

II-Recoilless absorption of radiation by a nucleus vibrating about its equilibrium position in a crystal. The Mössbauer effect

The system under consideration is now a nucleus of mass M vibrating at the angular frequency Ω about its equilibrium position in a crystalline lattice (the Einstein model; *cf.* Complement A_V, § 2). We again denote by $\mathbf{R}_G$ and $\mathbf{P}_G$ the position and momentum of the center of mass of this nucleus. The vibrational energy of the nucleus is described by the Hamiltonian:

$$H_e = \frac{1}{2M}\mathbf{P}_G^2 + \frac{1}{2}M\Omega^2(X_G^2 + Y_G^2 + Z_G^2) \tag{8}$$

which is that of a three-dimensional isotropic harmonic oscillator. Denote by $|\psi_{n_x,n_y,n_z}\rangle$ the eigenstate of H_e of eigenvalue $(n_x + n_y + n_z + 3/2)\hbar\Omega$. In addition to these external degrees of freedom, the nucleus possesses internal degrees of freedom with which are associated observables which all commute with $\mathbf{R}_G$ and $\mathbf{P}_G$. Let H_i be the Hamiltonian that describes the internal energy of the nucleus. As above, we concern ourselves with two eigenstates of H_i, $|\chi_a\rangle$ and $|\chi_b\rangle$, of energies E_a and E_b, and we set $\hbar\omega_0 = E_b - E_a$. Since $\hbar\omega_0$ falls into the γ-ray domain, we have, of course:

$$\omega_0 \gg \Omega \tag{9}$$

e. What energy must be furnished to the nucleus to allow it to go from the state $|\psi_{0,0,0}; \chi_a\rangle$ (the nucleus in the vibrational state defined by the quantum numbers $n_x = 0$, $n_y = 0$, $n_z = 0$ and the internal state $|\chi_a\rangle$) to the state $|\psi_{n,0,0}; \chi_b\rangle$?

f. This nucleus is placed in an electromagnetic wave of the type defined by (4), whose wave vector $\mathbf{k}$ is parallel to Ox. It can be shown that, in the electric dipole approximation, the interaction Hamiltonian of the nucleus with this plane wave (responsible for the absorption of the γ-rays) can be written as in (6), with:

$$W = \mathcal{A}_0\, S_i(k)\, \mathrm{e}^{ikX_G} \tag{10}$$

where $S_i(k)$ is an operator which acts on the internal degrees of freedom and consequently commutes with $\mathbf{R}_G$ and $\mathbf{P}_G$. Set $s(k) = \langle\chi_b|S_i(k)|\chi_a\rangle$.

The nucleus is initially in the state $|\psi_{0,0,0}; \chi_a\rangle$. Show that, under the influence of the incident plane wave, a resonance appears whenever $\hbar\omega$ coincides with one of the energies calculated in *e*, with the intensity of the corresponding resonance proportional to $|s(k)|^2|\langle\psi_{n,0,0}|\mathrm{e}^{ikX_G}|\psi_{0,0,0}\rangle|^2$, where the value of k is to be specified. Show, furthermore, that condition (9) allows us to replace k by $k_0 = \omega_0/c$ in the expression for the intensity of the resonance.

g. We set:

$$\pi_n(k_0) = |\langle\varphi_n|\mathrm{e}^{ik_0X_G}|\varphi_0\rangle|^2 \tag{11}$$

where the states $|\varphi_n\rangle$ are the eigenstates of a one-dimensional harmonic oscillator of position X_G, mass M and angular frequency Ω.

α. Calculate $\pi_n(k_0)$ in terms of $\hbar$, M, Ω, k_0, n (see also exercise 7 of Complement $\mathrm{M_V}$). Set $\xi = \frac{\hbar^2k_0^2}{2M}/\hbar\Omega$.

Hint: establish a recurrence relation between $\langle\varphi_n|\mathrm{e}^{ik_0X_G}|\varphi_0\rangle$ and $\langle\varphi_{n-1}|\mathrm{e}^{ik_0X_G}|\varphi_0\rangle$, and express all the $\pi_n(k_0)$ in terms of $\pi_0(k_0)$, which is to be calculated directly from the wave function of the harmonic oscillator ground state. Show that the $\pi_n(k_0)$ are given by a Poisson distribution.

β. Verify that $\sum_{n=0}^{\infty} \pi_n(k_0) = 1$

γ. Show that $\sum_{n=0}^{\infty} n\hbar\Omega\, \pi_n(k_0) = \hbar^2\omega_0^2/2Mc^2$.

h. Assume that $\hbar\Omega \gg \hbar^2\omega_0^2/2Mc^2$, i.e. that the vibrational energy of the nucleus is much greater than the recoil energy (very rigid crystalline bonds). Show that the absorption spectrum of the nucleus is essentially composed of a single line of angular frequency ω_0. This line is called the recoilless absorption line. Justify this name. Why does the Doppler effect disappear?

i. Now assume that $\hbar\Omega \ll \hbar^2\omega_0^2/2Mc^2$ (very weak crystalline bonds). Show that the absorption spectrum of the nucleus is composed of a very large number of equidistant lines whose barycenter (obtained by weighting the abscissa of each line by its relative intensity) coincides with the position of the absorption line of the free and initially motionless nucleus. What is the order of magnitude of the width of this spectrum (the dispersion of the lines about their barycenter)? Show that one obtains the results of the first part in the limit $\Omega \to 0$.

REFERENCES

Exercise 3:
References: see Brossel's lectures in (15.2).
Exercise 5:
References: see Townes and Schawlow (12.10), Chap. 10, § 9.
Exercise 6:
References: see Wilson (16.34).
Exercise 9:
References: see Abragam (14.1), Chap. VIII; Slichter (14.2), Chap. 5.
Exercise 10:
References: see De Benedetti (16.23); Valentin (16.1), annex XV.

Chapter XIV

Systems of identical particles

A Statement of the problem 1420
A-1 Identical particles: definition 1420
A-2 Identical particles in classical mechanics 1420
A-3 Identical particles in quantum mechanics: the difficulties of applying the general postulates 1421
B Permutation operators 1425
B-1 Two-particle systems 1426
B-2 Systems containing an arbitrary number of particles 1430
C The symmetrization postulate 1434
C-1 Statement of the postulate 1434
C-2 Removal of exchange degeneracy 1435
C-3 Construction of physical kets 1436
C-4 Application of the other postulates 1440
D Discussion 1443
D-1 Differences between bosons and fermions. Pauli's exclusion principle 1444
D-2 The consequences of particle indistinguishability on the calculation of physical predictions 1446

In Chapter III, we stated the postulates of non-relativistic quantum mechanics, and in Chapter IX, we concentrated on those which concern spin degrees of freedom. Here, we shall see (§ A) that, in reality, these postulates are not sufficient when we are dealing with systems containing several identical particles since, in this case, their application leads to ambiguities in the physical predictions. To eliminate these ambiguities, it is necessary to introduce a new postulate, concerning the quantum mechanical description of systems of identical particles. We shall state this postulate in § C and discuss its physical implications in § D. Before we do so, however, we shall (in § B) define and study permutation operators, which considerably facilitate the reasoning and the calculations.

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

A. Statement of the problem

A-1. Identical particles: definition

Two particles are said to be identical if all their intrinsic properties (mass, spin, charge, etc.) are exactly the same: no experiment can distinguish one from the other. Thus, all the electrons in the universe are identical, as are all the protons and all the hydrogen atoms. On the other hand, an electron and a positron are not identical, since, although they have the same mass and the same spin, they have different electrical charges.

An important consequence can be deduced from this definition: when a physical system contains two identical particles, there is no change in its properties or its evolution if the roles of these two particles are exchanged.

Comment:

Note that this definition is independent of the experimental conditions. Even if, in a given experiment, the charges of the particles are not measured, an electron and a positron can never be treated like identical particles.

A-2. Identical particles in classical mechanics

In classical mechanics, the presence of identical particles in a system poses no particular problems. This special case is treated just like the general case. Each particle moves along a well-defined trajectory, which enables us to distinguish it from the others and "follow" it throughout the evolution of the system.

To treat this point in greater detail, we shall consider a system of two identical particles. At the initial time t_0, the physical state of the system is defined by specifying the position and velocity of each of the two particles; we denote these initial data by $\{\mathbf{r}_0, \mathbf{v}_0\}$ and $\{\mathbf{r}'_0, \mathbf{v}'_0\}$. To describe this physical state and calculate its evolution, we number the two particles: $\mathbf{r}_1(t)$ and $\mathbf{v}_1(t)$ denote the position and velocity of particle (1) at time t, and $\mathbf{r}_2(t)$ and $\mathbf{v}_2(t)$, those of particle (2). This numbering has no physical foundation, as it would if we were dealing with two particles having different natures. It follows that the initial physical state which we have just defined may, in theory, be described by two different "mathematical states" as we can set, either:

$$\begin{aligned} &\mathbf{r}_1(t_0) = \mathbf{r}_0 \quad \mathbf{r}_2(t_0) = \mathbf{r}'_0 \\ &\mathbf{v}_1(t_0) = \mathbf{v}_0 \quad \mathbf{v}_2(t_0) = \mathbf{v}'_0 \end{aligned} \tag{A-1}$$

or:

$$\begin{aligned} &\mathbf{r}_1(t_0) = \mathbf{r}'_0 \quad \mathbf{r}_2(t_0) = \mathbf{r}_0 \\ &\mathbf{v}_1(t_0) = \mathbf{v}'_0 \quad \mathbf{v}_2(t_0) = \mathbf{v}_0 \end{aligned} \tag{A-2}$$

Now, let us consider the evolution of the system. Suppose that the solution of the equations of motion defined by initial conditions (A-1) can be written:

$$\mathbf{r}_1(t) = \mathbf{r}(t) \quad \mathbf{r}_2(t) = \mathbf{r}'(t) \tag{A-3}$$

where $\mathbf{r}(t)$ and $\mathbf{r}'(t)$ are two vector functions. The fact that the two particles are identical implies that the system is not changed if they exchange roles. Consequently, the Lagrangian $\mathcal{L}(\mathbf{r}_1, \mathbf{v}_1; \mathbf{r}_2, \mathbf{v}_2)$ and the classical Hamiltonian $\mathcal{H}(\mathbf{r}_1, \mathbf{p}_1; \mathbf{r}_2, \mathbf{p}_2)$ are invariant under exchange of indices 1 and 2. It follows that the solution of the equations of motion corresponding to the initial state (A-2) is:

$$\mathbf{r}_1(t) = \mathbf{r}'(t) \quad \mathbf{r}_2(t) = \mathbf{r}(t) \tag{A-4}$$

where the functions $\mathbf{r}(t)$ and $\mathbf{r}'(t)$ are the same as in (A-3).

The two possible mathematical descriptions of the physical state under consideration are therefore perfectly equivalent, since they lead to the same physical predictions. The particle which started from $\{\mathbf{r}_0, \mathbf{v}_0\}$ at t_0 is at $\mathbf{r}(t)$ with the velocity $\mathbf{v}(t) = \mathrm{d}\mathbf{r}/\mathrm{d}t$ at time t, and the one which started from $\{\mathbf{r}'_0, \mathbf{v}'_0\}$ is at $\mathbf{r}'(t)$ with the velocity $\mathbf{v}'(t) = \mathrm{d}\mathbf{r}'/\mathrm{d}t$ (Fig. 1). Under these conditions, all we need to do is choose, at the initial time, either one of the two possible "mathematical states" and ignore the existence of the other one. Thus, we treat the system as if the two particles were actually of different natures. The numbers (1) and (2), with which we label them arbitrarily at t_0, then act like intrinsic properties to distinguish the two particles. Since we can follow each particle step-by-step along its trajectory (arrows in Figure 1), we can determine the locations of the particle numbered (1) and the one numbered (2) at any time.

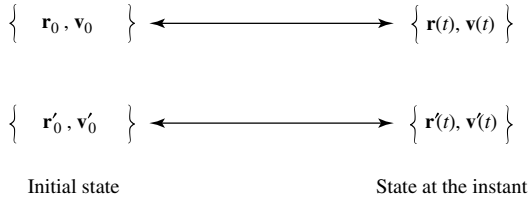


Figure 1: Position and velocity of each of the two particles at the initial time t_0 and at time t.

A-3. Identical particles in quantum mechanics: the difficulties of applying the general postulates

A-3-a. Qualitative discussion of a first simple example

It is immediately apparent that the situation is radically different in quantum mechanics, since the particles no longer have definite trajectories. Even if, at t_0, the wave packets associated with two identical particles are completely separated in space, their subsequent evolution may mix them. We then "lose track" of the particles; when we detect one particle in a region of space in which both of them have a non-zero position probability, we have no way of knowing if the particle detected is the one numbered (1) or the one numbered (2). Except in special cases – for example, when the two wave packets never overlap – the numbering of the two particles becomes ambiguous when their positions are measured, since, as we shall see, there exist several distinct "paths" taking the system from its initial state to the state found in the measurement.

To investigate this point in greater detail, consider a concrete example; a collision between two identical particles in their center of mass frame (Fig. 2). Before the collision, we have two completely separate wave packets, directed towards each other (Fig. 2a). We can agree, for example, to denote by (1) the particle on the left and by (2), the one on the right. During the collision (Fig. 2b), the two wave packets overlap. After the collision, the region of space in which the probability density of the two particles is non-zero[1] looks like a spherical shell whose radius increases over time (Fig. 2c). Suppose that a detector placed in the direction which makes an angle θ with the initial velocity of wave packet (l) detects a particle. It is then certain (because momentum is conserved in the collision) that the other particle is moving away in the opposite direction. However, it is impossible to know if the particle detected at D is the one initially numbered (1) or the one numbered (2). Thus, there are two different "paths" that could have led the system from the initial state shown in Figure 2a to the final state found in the measurement. These two paths are represented schematically in Figures 3a and 3b. Nothing enables us to determine which one was actually followed.

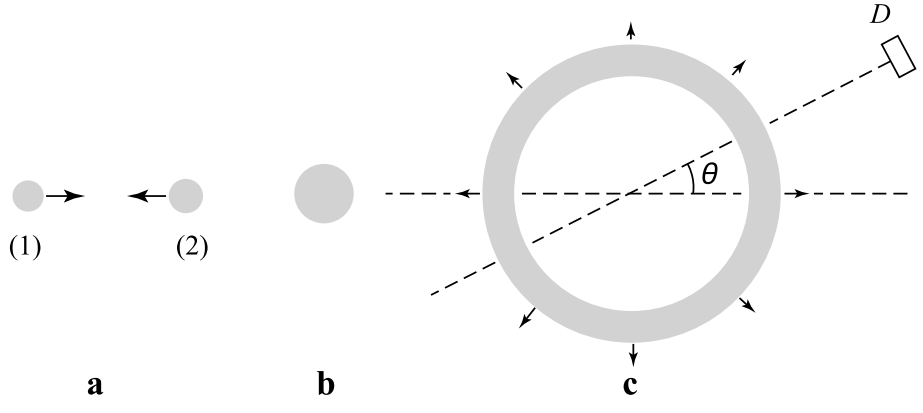


Figure 2: Collision between two identical particles in the center of mass frame: schematic representation of the probability density of the two particles. Before the collision (fig. a), the two wave packets are clearly separated and can be labeled. During the collision (fig. b), the two wave packets overlap. After the collision (fig. c), the probability density is non-zero in a region shaped like a spherical shell whose radius increases over time. Because the two particles are identical, it is impossible, when a particle is detected at D, to know with which wave packet, (1) or (2), it was associated before the collision.

A fundamental difficulty then arises in quantum mechanics when using the postulates of Chapter III. In order to calculate the probability of a given measurement result it is necessary to know the final state vectors associated with this result. Here, there are two, which correspond respectively to Figures 3a and 3b. These two kets are distinct (and, furthermore, orthogonal). Nevertheless, they are associated with a single physical

[1]The two-particle wave function depends on six variables (the components of the two particles coordinates $\mathbf{r}$ and $\mathbf{r}'$) and is not easily represented in 3 dimensions. Figure 2 is therefore very schematic: the grey regions are those to which both $\mathbf{r}$ and $\mathbf{r}'$ must belong for the wave function to take on significant values.

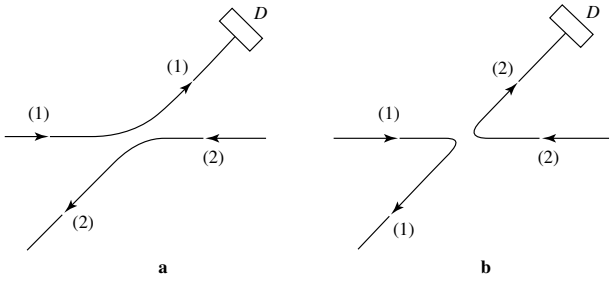


Figure 3: Schematic representation of two types of "paths" which the system could have followed in going from the initial state to the state found in the measurement. Because the two particles are identical, we cannot determine the path that was actually followed.

state since it is impossible to imagine a more complete measurement that would permit distinguishing between them. Under these conditions, should one calculate the probability using path 3a, path 3b or both? In the latter case, should one take the sum of the probabilities associated with each path, or the sum of their probability amplitudes (and in this case, with what sign)? These different possibilities lead, as we shall verify later, to different predictions.

The answer to the preceding questions will be given in § D after we have stated the symmetrization postulate. Before going on, we shall study another example that will aid us in understanding the difficulties related to the indistinguishability of two particles.

A-3-b. Origin of the difficulties: Exchange degeneracy

In the preceding example, we considered two wave packets which, initially, did not overlap; this enabled us to label each of them arbitrarily with a number, (1) or (2). Ambiguities appeared, however, when we tried to determine the mathematical state (or ket) associated with a given result of a position measurement. Actually, the same difficulty arises in the choice of the mathematical ket used to describe the initial physical state. This type of difficulty is related to the concept of "exchange degeneracy" which we shall introduce in this section. To simplify the reasoning, we shall first consider a different example, so as to confine ourselves to a finite-dimensional space. Then, we shall generalize the concept of exchange degeneracy, showing that it can be generalized to all quantum mechanical systems containing identical particles.

α. *Exchange degeneracy for a system of two spin 1/2 particles*

Let us consider a system composed of two identical spin 1/2 particles, confining ourselves to the study of its spin degrees of freedom. As in § A-2, we shall distinguish between the physical state of the system and its mathematical description (a ket in state space).

It would seem natural to suppose that, if we made a complete measurement of each of the two spins, we would then know the physical state of the total system perfectly. Here, we shall assume that the component along Oz of one of them is equal to $+\hbar/2$ and that of the other one, $-\hbar/2$ (this is the equivalent for the two spins of the specification

of $\{\mathbf{r}_0, \mathbf{v}_0\}$ and $\{\mathbf{r}'_0, \mathbf{v}'_0\}$ in § A-2).

To describe the system mathematically, we number the particles: $\mathbf{S}_1$ and $\mathbf{S}_2$ denote the two spin observables, and $\{|\varepsilon_1, \varepsilon_2\rangle\}$ (where ε_1 and ε_2 can be equal to + or −) is the orthonormal basis of the state space formed by the common eigenkets of S_{1z} (eigenvalue $\varepsilon_1\hbar/2$) and S_{2z} (eigenvalue $\varepsilon_2\hbar/2$).

Just as in classical mechanics, two different "mathematical states" could be associated with the same physical state. Either one of the two orthogonal kets:

$$|\varepsilon_1 = + \ , \ \varepsilon_2 = -\rangle \tag{A-5a}$$

$$|\varepsilon_1 = - \ , \ \varepsilon_2 = +\rangle \tag{A-5b}$$

can, *a priori*, describe the physical state considered here.

These two kets span a two-dimensional subspace whose normalized vectors are of the form:

$$\alpha|+,-\rangle \ + \ \beta|-,+\rangle \tag{A-6}$$

with:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{A-7}$$

By the superposition principle, all mathematical kets (A-6) can represent the same physical state as (A-5a) or (A-5b) (one spin pointing up and the other one pointing down). This is called "*exchange degeneracy*".

Exchange degeneracy creates fundamental difficulties, since application of the postulates of Chapter III to the various kets (A-6) can lead to physical predictions that depend on the ket chosen. Let us determine, for example, the probability of finding the components of the two spins along Ox both equal to $+\hbar/2$. With this measurement result is associated a single ket of the state space. According to formula (A-20) of Chapter IV, this ket can be written:

$$\begin{aligned}\frac{1}{\sqrt{2}}[|\varepsilon_1 = +\rangle + |\varepsilon_1 = -\rangle] \otimes \frac{1}{\sqrt{2}}[|\varepsilon_2 = +\rangle + |\varepsilon_2 = -\rangle] \\ = \frac{1}{2}\Big[|+,+\rangle \ + \ |-,+\rangle \ + \ |+,-\rangle \ + \ |-,-\rangle\Big]\end{aligned} \tag{A-8}$$

Consequently, the desired probability, for the vector (A-6), is equal to:

$$\left|\frac{1}{2}(\alpha + \beta)\right|^2 \tag{A-9}$$

This probability does depend on the coefficients α and β. It is not possible, therefore, to describe the physical state under consideration by the set of kets (A-6) or by any one of them chosen arbitrarily. The exchange degeneracy must be removed. That is, we must indicate unambiguously which of the kets (A-6) is to be used.

Comment:

In this example, exchange degeneracy appears only in the initial state, since we chose the same value for the components of the two spins in the final state. In the general case (for example, if the measurement result corresponds to two different eigenvalues of S_x), exchange degeneracy appears in both the initial and the final state.

β. Generalization

The difficulties related to exchange degeneracy arise in the study of all systems containing an arbitrary number N of identical particles ($N > 1$).

Consider, for example, a three-particle system. With each of the three particles, taken separately, are associated a state space and observables acting in this space. Thus, we are led to number the particles: $\mathcal{E}(1)$, $\mathcal{E}(2)$ and $\mathcal{E}(3)$ will denote the three one-particle state spaces, and the corresponding observables will be labeled by the same indices. The state space of the three-particle system is the tensor product:

$$\mathcal{E} = \mathcal{E}(1) \otimes \mathcal{E}(2) \otimes \mathcal{E}(3) \tag{A-10}$$

Now, consider an observable $B(1)$, initially defined in $\mathcal{E}(1)$. We shall assume that $B(1)$ alone constitutes a C.S.C.O. in $\mathcal{E}(1)$ [or that $B(1)$ actually denotes several observables which form a C.S.C.O.]. The fact that the three particles are identical implies that the observables $B(2)$ and $B(3)$ exist and that they constitute C.S.C.O.'s in $\mathcal{E}(2)$ and $\mathcal{E}(3)$ respectively. $B(1)$, $B(2)$ and $B(3)$ have the same spectrum, $\{b_n; n = 1, 2, \ldots\}$. Using the bases that define these three observables in $\mathcal{E}(1)$, $\mathcal{E}(2)$ and $\mathcal{E}(3)$, we can construct, by taking the tensor product, an orthonormal basis of $\mathcal{E}$, which we shall denote by:

$$|1 : b_i \ ; \ 2 : b_j \ ; \ 3 : b_k\rangle \quad ; \quad i, j, k = 1, 2, \tag{A-11}$$

The kets $|1 : b_i \ ; \ 2 : b_j \ ; \ 3 : b_k\rangle$ are common eigenvectors of the extensions of $B(1)$, $B(2)$ and $B(3)$ in $\mathcal{E}$, with respective eigenvalues b_i, b_j et b_k.

Since the three particles are identical, we cannot measure $B(1)$ or $B(2)$ or $B(3)$, since the numbering has no physical significance. However, we can measure the physical quantity B for each of the three particles. Suppose that such a measurement has resulted in three different eigenvalues, b_n, b_p and b_q. Exchange degeneracy then appears, since the state of the system after this measurement can, *a priori*, be represented by any one of the kets of the subspace of $\mathcal{E}$ spanned by the six basis vectors:

$$\begin{aligned} &|1 : b_n \ ; \ 2 : b_p \ ; \ 3 : b_q\rangle, \quad |1 : b_q \ ; \ 2 : b_n \ ; \ 3 : b_p\rangle, \quad |1 : b_p \ ; \ 2 : b_q \ ; \ 3 : b_n\rangle \\ &|1 : b_n \ ; \ 2 : b_q \ ; \ 3 : b_p\rangle, \quad |1 : b_p \ ; \ 2 : b_n \ ; \ 3 : b_q\rangle, \quad |1 : b_q \ ; \ 2 : b_p \ ; \ 3 : b_n\rangle \end{aligned} \tag{A-12}$$

Therefore, *a complete measurement on each of the particles does not permit the determination of a unique ket* of the state space of the system.

Comment:

The indeterminacy due to exchange degeneracy is, of course, less important if two of the eigenvalues found in the measurement are equal. This indeterminacy disappears in the special case in which the three results are identical.

B. Permutation operators

Before stating the additional postulate that enables us to remove the indeterminacy related to exchange degeneracy, we shall study certain operators, defined in the total state space of the system under consideration, which actually permute the various particles of the system. The use of these permutation operators will simplify the calculations and reasoning in §§ C and D.

B-1. Two-particle systems

B-1-a. Definition of the permutation operator P_{21}

Consider a system composed of two particles with the same spin s. Here it is not necessary for these two particles to be identical; it is sufficient that their individual state spaces be isomorphic. Therefore, to avoid the problems that arise when the two particles are identical, we shall assume that they are not: the numbers (1) and (2) with which they are labeled indicate their natures. For example, (1) will denote a proton and (2), an electron.

We choose a basis, $\{|u_i\rangle\}$, in the state space $\mathcal{E}(1)$ of particle (1). Since the two particles have the same spin, $\mathcal{E}(2)$ is isomorphic to $\mathcal{E}(1)$, and it can be spanned by the same basis. By taking the tensor product, we construct, in the state space $\mathcal{E}$ of the system, the basis:

$$\{|1 : u_i \ ; \ 2 : u_j\rangle\} \tag{B-1}$$

Since the order of the vectors is of no importance in a tensor product, we have:

$$|2 : u_j \ ; \ 1 : u_i\rangle \equiv |1 : u_i \ ; \ 2 : u_j\rangle \tag{B-2}$$

However, note that:

$$|1 : u_j \ ; \ 2 : u_i\rangle \neq |1 : u_i \ ; \ 2 : u_j\rangle \quad \text{if} \quad i \neq j \tag{B-3}$$

The permutation operator P_{21} is then defined as the linear operator whose action on the basis vectors is given by:

$$P_{21}|1 : u_i \ ; \ 2 : u_j\rangle = |2 : u_i \ ; \ 1 : u_j\rangle = |1 : u_j \ ; \ 2 : u_i\rangle \tag{B-4}$$

Its action on any ket of $\mathcal{E}$ can easily be obtained by expanding this ket[2] on the basis (B-1).

Comment:

If we choose the basis formed by the common eigenstates of the position observable $\mathbf{R}$ and the spin component S_z, (B-4) can be written:

$$P_{21}|1 : \mathbf{r}, \ \varepsilon; 2 : \mathbf{r}', \varepsilon'\rangle = |1 : \mathbf{r}', \varepsilon' \ ; \ 2 : \mathbf{r}, \varepsilon\rangle \tag{B-5}$$

Any ket $|\psi\rangle$ of the state space $\mathcal{E}$ can be represented by a set of $(2s+1)^2$ functions of six variables:

$$|\psi\rangle = \sum_{\varepsilon,\varepsilon'} \int \mathrm{d}^3r \ \mathrm{d}^3r' \ \psi_{\varepsilon,\varepsilon'}(\mathbf{r}, \mathbf{r}')|1 : \mathbf{r}, \varepsilon \ ; \ 2 : \mathbf{r}', \varepsilon'\rangle \tag{B-6}$$

with:

$$\psi_{\varepsilon,\varepsilon'}(\mathbf{r}, \mathbf{r}') = \langle 1 : \mathbf{r}, \varepsilon \ ; \ 2 : \mathbf{r}', \varepsilon'|\psi\rangle \tag{B-7}$$

[2] It can easily be shown that the operator P_{21} so defined does not depend on the $\{|u_i\rangle\}$ basis chosen.

We then have:

$$P_{21}|\psi\rangle = \sum_{\varepsilon,\varepsilon'} \int d^3r\ d^3r'\ \psi_{\varepsilon,\varepsilon'}(\mathbf{r},\mathbf{r}')|1:\mathbf{r}',\varepsilon'\ ;\ 2:\mathbf{r},\varepsilon\rangle \tag{B-8}$$

By changing the names of the dummy variables:

$$\begin{aligned}&\varepsilon \leftrightarrow \varepsilon'\\ &\mathbf{r} \leftrightarrow \mathbf{r}'\end{aligned} \tag{B-9}$$

we transform formula (B-8) into:

$$P_{21}|\psi\rangle = \sum_{\varepsilon,\varepsilon'} \int d^3r\ d^3r'\ \psi_{\varepsilon,\varepsilon'}(\mathbf{r}',\mathbf{r})|1:\mathbf{r},\varepsilon\ ;\ 2:\mathbf{r}',\varepsilon'\rangle \tag{B-10}$$

Consequently, the functions:

$$\psi'_{\varepsilon,\varepsilon'}(\mathbf{r},\mathbf{r}') = \langle 1:\mathbf{r},\varepsilon\ ;\ 2:\mathbf{r}',\varepsilon'|P_{21}|\psi\rangle \tag{B-11}$$

which represent the ket $|\psi'\rangle = P_{21}|\psi\rangle$ can be obtained from the functions (B-7) which represent the ket $|\psi\rangle$ by inverting $(\mathbf{r},\varepsilon)$ and $(\mathbf{r}',\varepsilon')$:

$$\psi'_{\varepsilon,\varepsilon'}(\mathbf{r},\mathbf{r}') = \psi_{\varepsilon',\varepsilon}(\mathbf{r}',\mathbf{r}) \tag{B-12}$$

B-1-b. Properties of P_{21}

We see directly from definition (B-4) that:

$$(P_{21})^2 = 1 \tag{B-13}$$

The operator P_{21} is its own inverse.

It can easily be shown that P_{21} *is Hermitian*:

$$P_{21}^\dagger = P_{21} \tag{B-14}$$

The matrix elements of P_{21} in the $\{|1:u_i\ ;\ 2:u_j\rangle\}$ basis are:

$$\begin{aligned}\langle 1:u_{i'}\ ;\ 2:u_{j'}|P_{21}|1:u_i\ ;\ 2:u_j\rangle &= \langle 1:u_{i'}\ ;\ 2:u_{j'}|1:u_j\ ;\ 2:u_i\rangle\\ &= \delta_{i'j}\delta_{j'i}\end{aligned} \tag{B-15}$$

Those of $P_{21}^\dagger$ are, by definition:

$$\begin{aligned}\langle 1:u_{i'}\ ;\ 2:u_{j'}|P_{21}^\dagger|1:u_i\ ;\ 2:u_j\rangle &= (\langle 1:u_i\ ;\ 2:u_j|P_{21}|1:u_{i'}\ ;\ 2:u_{j'}\rangle)^*\\ &= (\langle 1:u_i\ ;\ 2:u_j|1:u_{j'}\ ;\ 2:u_{i'}\rangle)^*\\ &= \delta_{ij'}\,\delta_{ji'}\end{aligned} \tag{B-16}$$

Each of the matrix elements of $P_{21}^\dagger$ is therefore equal to the corresponding matrix element of P_{21}. This leads to relation (B-14).

It follows from (B-13) and (B-14) that P_{21} is also *unitary*:

$$P_{21}^\dagger P_{21} = P_{21}P_{21}^\dagger = 1 \tag{B-17}$$

B-1-c. Symmetric and antisymmetric kets. Symmetrizer and antisymmetrizer

According to relation (B-14), the eigenvalues of P_{21} must be real. Since, according to (B-13), their squares are equal to 1, these eigenvalues are simply $+1$ and -1. The eigenvectors of P_{21} associated with the eigenvalue $+1$ are called *symmetric*, those corresponding to the eigenvalue -1, *antisymmetric*:

$$\begin{aligned} &P_{21}|\psi_S\rangle = |\psi_S\rangle \qquad \Longrightarrow |\psi_S\rangle \quad \text{symmetric} \\ &P_{21}|\psi_A\rangle = -|\psi_A\rangle \qquad \Longrightarrow |\psi_A\rangle \quad \text{antisymmetric} \end{aligned} \tag{B-18}$$

Now consider the two operators:

$$S = \frac{1}{2}(1 + P_{21}) \tag{B-19a}$$

$$A = \frac{1}{2}(1 - P_{21}) \tag{B-19b}$$

These operators are *projectors*, since (B-13) implies that:

$$S^2 = S \tag{B-20a}$$

$$A^2 = A \tag{B-20b}$$

and, in addition, (B-14) enables us to show that:

$$S^\dagger = S \tag{B-21a}$$

$$A^\dagger = A \tag{B-21b}$$

S and A are projectors onto orthogonal subspaces, since, according to (B-13):

$$SA = AS = 0 \tag{B-22}$$

These subspaces are supplementary, since definitions (B-19) yield:

$$S + A = 1 \tag{B-23}$$

If $|\psi\rangle$ is an arbitrary ket of the state space $\mathcal{E}$, $S|\psi\rangle$ is a symmetric ket and $A|\psi\rangle$, an antisymmetric ket, as it is easy to see, using (B-13) again, that:

$$\begin{aligned} &P_{21}S|\psi\rangle = S|\psi\rangle \\ &P_{21}A|\psi\rangle = -A|\psi\rangle \end{aligned} \tag{B-24}$$

For this reason, S and A are called, respectively, a *symmetrizer* and an *antisymmetrizer*.

Comment:

The same symmetric ket is obtained by applying S to $P_{21}|\psi\rangle$ or to $|\psi\rangle$ itself:

$$SP_{21}|\psi\rangle = S|\psi\rangle \tag{B-25}$$

For the antisymmetrizer, we have, similarly:

$$AP_{21}|\psi\rangle = -A|\psi\rangle \tag{B-26}$$

B-1-d. Transformation of observables by permutation

Consider an observable $B(1)$, initially defined in $\mathcal{E}(1)$ and then extended into $\mathcal{E}$. It is always possible to construct the $\{|u_i\rangle\}$ basis in $\mathcal{E}(1)$ from eigenvectors of $B(1)$ (the corresponding eigenvalues will be written b_i). Let us now calculate the action of the operator $P_{21}B(1)P_{21}^\dagger$ on an arbitrary basis ket of $\mathcal{E}$:

$$\begin{aligned} P_{21}B(1)P_{21}^\dagger|1:u_i\ ;\ 2:u_j\rangle &= P_{21}B(1)|1:u_j\ ;\ 2:u_i\rangle \\ &= b_j P_{21}|1:u_j\ ;\ 2:u_i\rangle \\ &= b_j|1:u_i\ ;\ 2:u_j\rangle \end{aligned} \tag{B-27}$$

We would obtain the same result by applying the observable $B(2)$ directly to the basis ket chosen. Consequently:

$$P_{21}B(1)P_{21}^\dagger = B(2) \tag{B-28}$$

The same reasoning shows that:

$$P_{21}B(2)P_{21}^\dagger = B(1) \tag{B-29}$$

In $\mathcal{E}$, there are also observables, such as $B(1)+C(2)$ or $B(1)C(2)$, which involve both indices simultaneously. We obviously have:

$$P_{21}[B(1)+C(2)]P_{21}^\dagger = B(2)+C(1) \tag{B-30}$$

Similarly, using (B-17), we find:

$$\begin{aligned} P_{21}B(1)C(2)P_{21}^\dagger &= P_{21}B(1)P_{21}^\dagger P_{21}C(2)P_{21}^\dagger \\ &= B(2)C(1) \end{aligned} \tag{B-31}$$

These results can be generalized to all observables in $\mathcal{E}$ which can be expressed in terms of observables of the type of $B(1)$ and $C(2)$, to be denoted by $\mathcal{O}(1,2)$:

$$P_{21}\mathcal{O}(1,2)P_{21}^\dagger = \mathcal{O}(2,1) \tag{B-32}$$

$\mathcal{O}(2,1)$ is the observable obtained from $\mathcal{O}(1,2)$ by exchanging indices 1 and 2 throughout.

An observable $\mathcal{O}_S(1,2)$ is said to be *symmetric* if:

$$\mathcal{O}_S(2,1) = \mathcal{O}_S(1,2) \tag{B-33}$$

According to (B-32), all symmetric observables satisfy:

$$P_{21}\mathcal{O}_S(1,2) = \mathcal{O}_S(1,2)P_{21} \tag{B-34}$$

that is:

$$[\mathcal{O}_S(1,2), P_{21}] = 0 \tag{B-35}$$

Symmetric observables commute with the permutation operator.

B-2. Systems containing an arbitrary number of particles

In the state space of a system composed of N particles with the same spin (temporarily assumed to be of different natures), $N!$ permutation operators can be defined (one of which is the identity operator). If N is greater than 2, the properties of these operators are more complex than those of P_{21}. To have an idea of the changes involved when N is greater than 2, we shall briefly study the case in which $N = 3$.

B-2-a. Definition of the permutation operators

Consider, therefore, a system of three particles that are not necessarily identical, but have the same spin. As in § B-1-a, we construct a basis of the state space of the system by taking a tensor product:

$$\{|1 : u_i \ ; \ 2 : u_j \ ; \ 3 : u_k\rangle\} \tag{B-36}$$

In this case, there exist six permutation operators, which we shall denote by:

$$P_{123} \ , \ P_{312} \ , \ P_{231} \ , \ P_{132} \ , \ P_{213} \ , \ P_{321} \tag{B-37}$$

By definition, the operator P_{npq} (where n, p, q is an arbitrary permutation of the numbers 1, 2, 3) is the linear operator whose action on the basis vectors obeys:

$$P_{npq}|1 : u_i \ ; \ 2 : u_j \ ; \ 3 : u_k\rangle = |n : u_i \ ; \ p : u_j \ ; \ q : u_k\rangle \tag{B-38}$$

For example:

$$\begin{aligned} P_{231}|1 : u_i \ ; \ 2 : u_j \ ; \ 3 : u_k\rangle &= |2 : u_i \ ; \ 3 : u_j \ ; \ 1 : u_k\rangle \\ &= |1 : u_k : 2 : u_i \ ; \ 3 : u_j\rangle \end{aligned} \tag{B-39}$$

P_{123} therefore coincides with the identity operator. The action of P_{npq} on any ket of the state space can easily be obtained by expanding this ket on the basis (B-36).

The $N!$ permutation operators associated with a system of N particles with the same spin could be defined analogously.

B-2-b. Properties

α. *The set of permutation operators constitutes a group*

This can easily be shown for the operators (B-37):

(*i*) P_{123} is the identity operator.

(*ii*) The product of two permutation operators is also a permutation operator. We can show, for example, that:

$$P_{312}P_{132} = P_{321} \tag{B-40}$$

To do so, we apply the left-hand side to an arbitrary basis ket:

$$\begin{aligned} P_{312}P_{132}|1 : u_i \ ; \ 2 : u_j \ ; \ 3 : u_k\rangle & \\ &= P_{312}|1 : u_i \ ; \ 3 : u_j \ ; \ 2 : u_k\rangle \\ &= P_{312}|1 : u_i \ ; \ 2 : u_k \ ; \ 3 : u_j\rangle \\ &= |3 : u_i \ ; \ 1 : u_k \ ; \ 2 : u_j\rangle \\ &= |1 : u_k \ ; \ 2 : u_j \ ; \ 3 : u_i\rangle \end{aligned} \tag{B-41}$$

The action of P_{321} effectively leads to the same result:

$$\begin{aligned} P_{321}|1:u_i\ ;\ 2:u_j\ ;\ 3:u_k\rangle &= |3:u_i\ ;\ 2:u_j\ ;\ 1:u_k\rangle \\ &= |1:u_k\ ;\ 2:u_j\ ;\ 3:u_i\rangle \end{aligned} \tag{B-42}$$

(*iii*) Each permutation operator has an inverse, which is also a permutation operator. Reasoning as in (*ii*), we can easily show that:

$$\begin{aligned} &P_{123}^{-1} = P_{123}\ ; \quad P_{312}^{-1} = P_{231}\ ; \quad P_{231}^{-1} = P_{312} \\ &P_{132}^{-1} = P_{132}\ ; \quad P_{213}^{-1} = P_{213}\ ; \quad P_{321}^{-1} = P_{321} \end{aligned} \tag{B-43}$$

Note that *the permutation operators do not commute with each other.* For example:

$$P_{132}P_{312} = P_{213} \tag{B-44}$$

which, compared to (B-40), shows that the commutator of P_{132} and P_{312} is not zero.

β. Transpositions. Parity of a permutation operator

A *transposition* is a permutation which simply exchanges the roles of two of the particles, without touching the others. Of the operators (B-37), the last three are transposition operators[3]. Transposition operators are Hermitian, and each of them is the same as its inverse, so that they are also unitary [the proofs of these properties are identical to those for (B-14), (B-13) and (B-17)].

Any permutation operator can be broken down into a product of transposition operators. For example, the second operator (B-37) can be written:

$$P_{312} = P_{132}P_{213} = P_{321}P_{132} = P_{213}P_{321} = P_{132}P_{213}(P_{132})^2 = \ldots \tag{B-45}$$

This decomposition is not unique. However, for a given permutation, it can be shown that the parity of the number of transpositions into which it can be broken down is always the same: it is called the *parity of the permutation.* Thus, the first three operators (B-37) are even, and the last three, odd. For any N, there are always as many even permutations as odd ones.

γ. Permutation operators are unitary

Permutation operators, which are products of transposition operators, all of which are unitary, are therefore also unitary. However, they are not necessarily Hermitian, since transposition operators do not generally commute with each other.

Finally, note that the adjoint of a given permutation operator has the same parity as that of the operator, since it is equal to the product of the same transposition operators, taken in the opposite order.

B-2-c. Completely symmetric or antisymmetric kets. Symmetrizer and antisymmetrizer

Since the permutation operators do not commute for $N > 2$, it is not possible to construct a basis formed by common eigenvectors of these operators. Nevertheless, we shall see that there exist certain kets which are simultaneously eigenvectors of all the permutation operators.

[3]Of course, for $N = 2$, the only possible permutation is a transposition.

We shall denote by P_α an arbitrary permutation operator associated with a system of N particles with the same spin; α represents an arbitrary permutation of the first N integers. A ket $|\psi_s\rangle$ such that:

$$P_\alpha|\psi_s\rangle = |\psi_s\rangle \tag{B-46}$$

for any permutation P_α, is said to be *completely symmetric*. Similarly, a *completely antisymmetric* ket $|\psi_A\rangle$ satisfies, by definition[4]:

$$P_\alpha|\psi_A\rangle = \varepsilon_\alpha|\psi_A\rangle \tag{B-47}$$

where:

$$\begin{aligned}&\varepsilon_\alpha = +1 \text{ if } P_\alpha \text{ is an even permutation}\\ &\varepsilon_\alpha = -1 \text{ if } P_\alpha \text{ is an odd permutation}\end{aligned} \tag{B-48}$$

The set of completely symmetric kets constitutes a vector subspace $\mathcal{E}_S$ of the state space $\mathcal{E}$; the set of completely antisymmetric kets, a subspace $\mathcal{E}_A$.

Now consider the two operators:

$$S = \frac{1}{N!}\sum_\alpha P_\alpha \tag{B-49}$$

$$A = \frac{1}{N!}\sum_\alpha \varepsilon_\alpha P_\alpha \tag{B-50}$$

where the summations are performed over the $N!$ permutations of the first N integers, and ε_α is defined by (B-48). We shall show that S and A are the projectors onto $\mathcal{E}_S$ and $\mathcal{E}_A$ respectively. For this reason, they are called a *symmetrizer* and an *antisymmetrizer*.

S and A are Hermitian operators:

$$S^\dagger = S \tag{B-51}$$

$$A^\dagger = A \tag{B-52}$$

The adjoint $P_\alpha^\dagger$ of a given permutation operator is, as we saw above (*cf.* § B-2-b-γ), another permutation operator, of the same parity (which coincides, furthermore, with P_α^{-1}). Taking the adjoints of the right-hand sides of the definitions of S and A therefore amounts simply to changing the order of the terms in the summations (since the set of the P_α^{-1} is again the permutation group).

Also, if P_{α_0} is an arbitrary permutation operator, we have:

$$P_{\alpha_0}S = SP_{\alpha_0} = S \tag{B-53a}$$

$$P_{\alpha_0}A = AP_{\alpha_0} = \varepsilon_{\alpha_0}A \tag{B-53b}$$

This is due to the fact that $P_{\alpha_0}P_\alpha$ is also a permutation operator:

$$P_{\alpha_0}P_\alpha = P_\beta \tag{B-54}$$

[4]According to the property stated in § B-2-b-β, this definition can also be based solely on the transposition operators: any transposition operator leaves a completely symmetric ket invariant and transforms a completely antisymmetric ket into its opposite.

such that:

$$\varepsilon_\beta = \varepsilon_{\alpha_0}\varepsilon_\alpha \tag{B-55}$$

If, for P_{α_0} fixed, we choose successively for P_α all the permutations of the group, we see that the P_β are each identical to one and only one of these permutations (in, of course, a different order). Consequently:

$$P_{\alpha_0}S = \frac{1}{N!}\sum_\alpha P_{\alpha_0}P_\alpha = \frac{1}{N!}\sum_\beta P_\beta = S \tag{B-56a}$$

$$P_{\alpha_0}A = \frac{1}{N!}\sum_\alpha \varepsilon_\alpha P_{\alpha_0}P_\alpha = \frac{1}{N!}\varepsilon_{\alpha_0}\sum_\beta \varepsilon_\beta P_\beta = \varepsilon_{\alpha_0}A \tag{B-56b}$$

Similarly, we could prove analogous relations in which S and A are multiplied by P_{α_0} from the right.

From (B-53), we see that:

$$\begin{aligned} S^2 &= S \\ A^2 &= A \end{aligned} \tag{B-57}$$

and, moreover:

$$AS = SA = 0 \tag{B-58}$$

This is because:

$$\begin{aligned} S^2 &= \frac{1}{N!}\sum_\alpha P_\alpha S = \frac{1}{N!}\sum_\alpha S = S \\ A^2 &= \frac{1}{N!}\sum_\alpha \varepsilon_\alpha P_\alpha A = \frac{1}{N!}\sum_\alpha \varepsilon_\alpha^2 A = A \end{aligned} \tag{B-59}$$

as each summation includes $N!$ terms; furthermore:

$$AS = \frac{1}{N!}\sum_\alpha \varepsilon_\alpha P_\alpha S = \frac{1}{N!}S\sum_\alpha \varepsilon_\alpha = 0 \tag{B-60}$$

since half the ε_α are equal to $+1$ and half equal to -1 (*cf.* § B-2-b-β).

S and A are therefore *projectors*. They project respectively onto $\mathcal{E}_S$ and $\mathcal{E}_A$ since, according to (B-53), their action on any ket $|\psi\rangle$ of the state space yields a completely symmetric or completely antisymmetric ket:

$$P_{\alpha_0}S|\psi\rangle = S|\psi\rangle \tag{B-61a}$$

$$P_{\alpha_0}A|\psi\rangle = \varepsilon_{\alpha_0}A|\psi\rangle \tag{B-61b}$$

Comments:

(*i*) The completely symmetric ket constructed by the action of S on $P_\alpha|\psi\rangle$, where P_α is an arbitrary permutation, is the same as that obtained from $|\psi\rangle$, since expressions (B-53) indicate that:

$$SP_\alpha|\psi\rangle = S|\psi\rangle \tag{B-62}$$

As for the corresponding completely antisymmetric kets, they differ at most by their signs:

$$AP_\alpha|\psi\rangle = \varepsilon_\alpha A|\psi\rangle \tag{B-63}$$

(*ii*) For $N > 2$, the symmetrizer and antisymmetrizer are not projectors onto supplementary subspaces. For example, when $N = 3$, it is easy to obtain [by using the fact that the first three permutations (B-37) are even and the others odd] the relation:

$$S + A = \frac{1}{3}(P_{123} + P_{231} + P_{312}) \neq 1 \tag{B-64}$$

In other words, the state space is not the direct sum of the subspace $\mathcal{E}_S$ of completely symmetric kets and the subspace $\mathcal{E}_A$ of completely antisymmetric kets.

B-2-d. Transformation of observables by permutation

We have indicated (§ B-2-b-β) that any permutation operator of an N-particle system can be broken down into a product of transposition operators analogous to the operator P_{21} studied in § B-1. For these transposition operators, we can use the arguments of § B-1-d to determine the behavior of the various observables of the system when they are multiplied from the left by an arbitrary permutation operator P_α and from the right by $P_\alpha^\dagger$.

In particular, the observables $\mathcal{O}_S(1, 2, \ldots, N)$ which are completely symmetric under exchange of the indices 1, 2, ..., N, commute with all the transposition operators, and, therefore, with all the permutation operators:

$$[\mathcal{O}_S(1, 2, \ldots, N), P_\alpha] = 0 \tag{B-65}$$

C. The symmetrization postulate

C-1. Statement of the postulate

> When a system includes several identical particles, only certain kets of its state space can describe its physical states. Physical kets are, depending on the nature of the identical particles, either completely symmetric or completely antisymmetric with respect to permutation of these particles. Those particles for which the physical kets are symmetric are called *bosons*, and those for which they are antisymmetric, *fermions*.

The symmetrization postulate thus limits the state space for a system of identical particles. This space is no longer, as it was in the case of particles of different natures, the tensor product $\mathcal{E}$ of the individual state spaces of the particles constituting the system. It is only a subspace of $\mathcal{E}$, namely $\mathcal{E}_S$ or $\mathcal{E}_A$, depending on whether the particles are bosons or fermions.

From the point of view of this postulate, particles existing in nature are divided into two categories. All currently known particles obey the following *empirical rule*[5]: particles of half-integral spin (electrons, positrons, protons, neutrons, muons, etc.) are fermions, and particles of integral spin (photons, mesons, etc.) are bosons.

[5]The "spin-statistics theorem", proven in quantum field theory, makes it possible to consider this

Comment:

Once this rule has been verified for the particles which are called "elementary", it holds for all other particles as well, inasmuch as they are composed of these elementary particles. Consider a system of many identical composite particles. Permuting two of them is equivalent to simultaneously permuting all the particles composing the first one with the corresponding particles (necessarily identical to the aforementioned ones) of the second one. This permutation must leave the ket describing the state of the system unchanged if the composite particles being studied are formed only of elementary bosons or if each of them contains an even number of fermions (no sign change, or an even number of sign changes); in this case, the particles are bosons. On the other hand, composite particles containing an odd number of fermions are themselves fermions (an odd number of sign changes in the permutation). Now, the spin of these composite particles is necessarily integral in the first case and half-integral in the second one (Chap. X, § C-3-c). They therefore obey the rule just stated. For example, atomic nuclei are known to be composed of neutrons and protons, which are fermions (spin 1/2). Consequently, nuclei whose mass number A (the total number of nucleons) is even are bosons, and those whose mass number is odd are fermions. Thus, the nucleus of the ^{3}He isotope of helium is a fermion, and that of the ^{4}He isotope, a boson.

C-2. Removal of exchange degeneracy

We shall begin by examining how this new postulate removes the exchange degeneracy and the corresponding difficulties.

The discussion of § A can be summarized in the following way. Let $|u\rangle$ be a ket which can mathematically describe a well-defined physical state of a system containing N identical particles. For any permutation operator P_α, $P_\alpha|u\rangle$ can describe this physical state as well as $|u\rangle$. The same is true for any ket belonging to the subspace $\mathcal{E}_u$ spanned by $|u\rangle$ and all its permutations $P_\alpha|u\rangle$. Depending on the ket $|u\rangle$ chosen, the dimension of $\mathcal{E}_u$ can vary between 1 and $N!$. If this dimension is greater than 1, several mathematical kets correspond to the same physical state: there is then an exchange degeneracy.

The new postulate which we have introduced considerably restricts the class of mathematical kets able to describe a physical state: these kets must belong to $\mathcal{E}_S$ for bosons, or to $\mathcal{E}_A$ for fermions. We shall be able to say that the difficulties related to exchange degeneracy are eliminated if we can show that $\mathcal{E}_u$ contains *a single* ket of $\mathcal{E}_S$ or *a single* ket of $\mathcal{E}_A$.

To do so, we shall use the relations $S = SP_\alpha$ or $A = \varepsilon_\alpha AP_\alpha$, proven in (B-53). We obtain:

$$S|u\rangle = SP_\alpha|u\rangle \tag{C-1a}$$

$$A|u\rangle = \varepsilon_\alpha AP_\alpha|u\rangle \tag{C-1b}$$

These relations express the fact that the projections onto $\mathcal{E}_S$ and $\mathcal{E}_A$ of the various kets which span $\mathcal{E}_u$ and, consequently, of all the kets of $\mathcal{E}_u$, are collinear. The symmetrization postulate thus unambiguously indicates (to within a constant factor) *the* ket of $\mathcal{E}_u$ which

rule to be a consequence of very general hypotheses. However, these hypotheses may not all be correct, and discovery of a boson of half-integral spin or a fermion of integral spin remains possible. It is not inconceivable that, for certain particles, the physical kets might have more complex symmetry properties than those envisaged here.

must be associated with the physical state considered: $S|u\rangle$ for bosons and $A|u\rangle$ for fermions. This ket will be called the *physical ket*.

Comment:

It is possible for all the kets of $\mathcal{E}_u$ to have a zero projection onto $\mathcal{E}_A$ (or $\mathcal{E}_S$). In this case, the symmetrization postulate excludes the corresponding physical state. Later (§§ C-3-b and C-3-c), we shall see examples of such a situation when dealing with fermions.

C-3. Construction of physical kets

C-3-a. The construction rule

The discussion of the preceding section leads directly to the following rule for the construction of the *unique* ket (the physical ket) corresponding to a given physical state of a system of N identical particles:

(*i*) Number the particles arbitrarily, and construct the ket $|u\rangle$ corresponding to the physical state considered and to the numbers given to the particles.

(*ii*) Apply S or A to $|u\rangle$, depending on whether the identical particles are bosons or fermions.

(*iii*) Normalize the ket so obtained.

We shall describe some simple examples to illustrate this rule.

C-3-b. Application to systems of two identical particles

Consider a system composed of two identical particles. Suppose that one of them is known to be in the individual state characterized by the normalized ket $|\varphi\rangle$, and the other one, in the individual state characterized by the normalized ket $|\chi\rangle$.

First of all, we shall envisage the case in which the two kets, $|\varphi\rangle$ and $|\chi\rangle$, are distinct. The preceding rule is applied in the following way:

(*i*) We label with the number 1, for example, the particle in the state $|\varphi\rangle$, and with the number 2, the one in the state $|\chi\rangle$. This yields:

$$|u\rangle = |1 : \varphi \ ; \ 2 : \chi\rangle \tag{C-2}$$

(*ii*) We symmetrize $|u\rangle$ if the particles are bosons:

$$S|u\rangle = \frac{1}{2}[|1 : \varphi \ ; \ 2 : \chi\rangle + |1 : \chi \ ; \ 2 : \varphi\rangle] \tag{C-3a}$$

We antisymmetrize $|u\rangle$ if the particles are fermions:

$$A|u\rangle = \frac{1}{2}[|1 : \varphi \ ; \ 2 : \chi\rangle - |1 : \chi \ ; \ 2 : \varphi\rangle] \tag{C-3b}$$

(*iii*) The kets (C-3a) and (C-3b), in general, are not normalized. If we assume $|\varphi\rangle$ and $|\chi\rangle$ to be orthogonal, the normalization constant is very simple to calculate. All we have to do to normalize $S|u\rangle$ or $A|u\rangle$ is replace the factor $1/2$ appearing in formulas (C-3) by $1/\sqrt{2}$. The normalized physical ket, in this case, can therefore be written:

$$|\varphi\ ;\ \chi\rangle = \frac{1}{\sqrt{2}}\left[|1:\varphi\ ;\ 2:\chi\rangle + \varepsilon|1:\chi\ ;\ 2:\varphi\rangle\right] \tag{C-4}$$

with $\varepsilon = +1$ for bosons and -1 for fermions.

We shall now assume that the two individual states, $|\varphi\rangle$ and $|\chi\rangle$, are identical:

$$|\varphi\rangle = |\chi\rangle \tag{C-5}$$

(C-2) then becomes:

$$|u\rangle = |1:\varphi\ ;\ 2:\varphi\rangle \tag{C-6}$$

$|u\rangle$ is already symmetric. If the two particles are bosons, (C-6) is then the physical ket associated with the state in which the two bosons are in the same individual state $|\varphi\rangle$. If, on the other hand, the two particles are fermions, we see that:

$$A|u\rangle = \frac{1}{2}\Big[|1:\varphi\ ;\ 2:\varphi\rangle - |1:\varphi\ ;\ 2:\varphi\rangle\Big] = 0 \tag{C-7}$$

Consequently, there exists no ket of $\mathcal{E}_A$ able to describe the physical state in which two fermions are in the same individual state $|\varphi\rangle$. Such a physical state is therefore excluded by the symmetrization postulate. We have thus established, for a special case, a fundamental result known as "*Pauli's exclusion principle*": *two identical fermions cannot be in the same individual state*. This result has some very important physical consequences which we shall discuss in § D-1.

C-3-c. Generalization to an arbitrary number of particles

These ideas can be generalized to an arbitrary number N of particles. To see how this can be done, we shall first treat the case $N = 3$.

Consider a physical state of the system defined by specifying the three individual normalized states $|\varphi\rangle$, $|\chi\rangle$ and $|\omega\rangle$. The state $|u\rangle$ which enters into the rule of § a can be chosen in the form:

$$|u\rangle = |1:\varphi\ ;\ 2:\chi\ ;\ 3:\omega\rangle \tag{C-8}$$

We shall discuss the cases of bosons and fermions separately.

α. The case of bosons

The application of S to $|u\rangle$ gives:

$$\begin{aligned} S|u\rangle &= \frac{1}{3!}\sum_{\alpha} P_\alpha|u\rangle \\ &= \frac{1}{6}\Big[|1:\varphi\ ;\ 2:\chi\ ;\ 3:\omega\rangle + |1:\omega\ ;\ 2:\varphi\ ;\ 3:\chi\rangle + |1:\chi\ ;\ 2:\omega\ ;\ 3:\varphi\rangle \\ &\quad + |1:\varphi\ ;\ 2:\omega\ ;\ 3:\chi\rangle + |1:\chi\ ;\ 2:\varphi\ ;\ 3:\omega\rangle + |1:\omega\ ;\ 2:\chi\ ;\ 3:\varphi\rangle\Big] \end{aligned} \tag{C-9}$$

It then suffices to normalize the ket (C-9).

First of all, let us assume that the three kets $|\varphi\rangle$, $|\chi\rangle$ and $|\omega\rangle$ are orthogonal. The six kets appearing on the right-hand side of (C-9) are then also orthogonal. To normalize (C-9), all we must do is replace the factor $1/6$ by $1/\sqrt{6}$.

If the two states $|\varphi\rangle$ and $|\chi\rangle$ coincide, while remaining orthogonal to $|\omega\rangle$, only three distinct kets now appear on the right-hand side of (C-9). It can easily be shown that the normalized physical ket can then be written:

$$|\varphi\ ;\ \varphi\ ;\ \omega\rangle = \frac{1}{\sqrt{3}}\Big[|1:\varphi\ ;\ 2:\varphi\ ;\ 3:\omega\rangle + |1:\varphi\ ;\ 2:\omega\ ;\ 3:\varphi\rangle + |1:\omega\ ;\ 2:\varphi\ ;\ 3:\varphi\rangle\Big] \tag{C-10}$$

Finally, if the three states $|\varphi\rangle$, $|\chi\rangle$, $|\omega\rangle$ are the same, the ket:

$$|u\rangle = |1:\varphi\ ;\ 2:\varphi\ ;\ 3:\varphi\rangle \tag{C-11}$$

is already symmetric and normalized.

β. The case of fermions

The application of A to $|u\rangle$ leads to:

$$A|u\rangle = \frac{1}{3!}\sum_{\alpha}\varepsilon_\alpha P_\alpha|1:\varphi\ ;\ 2:\chi\ ;\ 3:\omega\rangle \tag{C-12}$$

The signs of the various terms of the sum (C-12) are determined by the same rule as those of a 3×3 determinant. This is why it is convenient to write $A|u\rangle$ in the form of a *Slater determinant*:

$$A|u\rangle = \frac{1}{3!}\begin{vmatrix} |1:\varphi\rangle & |1:\chi\rangle & |1:\omega\rangle \\ |2:\varphi\rangle & |2:\chi\rangle & |2:\omega\rangle \\ |3:\varphi\rangle & |3:\chi\rangle & |3:\omega\rangle \end{vmatrix} \tag{C-13}$$

$A|u\rangle$ is zero if two of the individual states $|\varphi\rangle$, $|\chi\rangle$ or $|\omega\rangle$ coincide, since the determinant (C-13) then has two identical columns. We obtain Pauli's exclusion principle, already mentioned in § C-3-b: the same quantum mechanical state cannot be simultaneously occupied by several identical fermions.

Finally, note that if the three states $|\varphi\rangle$, $|\chi\rangle$, $|\omega\rangle$ are orthogonal, the six kets appearing on the right-hand side of (C-12) are orthogonal. All we must then do to normalize $A|u\rangle$ is replace the factor $1/3!$ appearing in (C-12) or (C-13) by $1/\sqrt{3!}$.

If, now, the system being considered contains more than three identical particles, the situation actually remains similar to the one just described. It can be shown that, for N identical bosons, it is always possible to construct the physical state $S|u\rangle$ from arbitrary individual states $|\varphi\rangle$, $|\chi\rangle$, ... On the other hand, for fermions, the physical ket $A|u\rangle$ can be written in the form of an $N\times N$ Slater determinant; this excludes the case in which two individual states coincide (the ket $A|u\rangle$ is then zero). This shows, and we shall return to this in detail in § D, how different the consequences of the new postulate can be for fermion and boson systems.

C-3-d. Construction of a basis in the physical state space

Consider a system of N identical particles. Starting with a basis, $\{|u_i\rangle\}$, in the state space of a single particle, we can construct the basis:

$$\{|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle\}$$

in the tensor product space $\mathcal{E}$. However, since the physical state space of the system is not $\mathcal{E}$, but rather one of the subspaces $\mathcal{E}_S$ or $\mathcal{E}_A$, the problem arises of how to determine a basis in this physical state space.

By application of S (or A) to the various kets of the basis:

$$\{|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle\}$$

we can obtain a set of vectors spanning $\mathcal{E}_S$ (or $\mathcal{E}_A$). Let $|\varphi\rangle$ be an arbitrary ket of $\mathcal{E}_S$, for example (the case in which $|\varphi\rangle$ belongs to $\mathcal{E}_A$ can be treated in the same way). $|\varphi\rangle$, which belongs to $\mathcal{E}$, can be expanded in the form:

$$|\varphi\rangle = \sum_{i,j,\ldots,p} a_{i,j,\ldots,p} |1 : u_i \ ; \ 2 : u_j \ ; \ \ldots N : u_p\rangle \tag{C-14}$$

Since $|\varphi\rangle$, by hypothesis, belongs to $\mathcal{E}_S$, we have $S|\varphi\rangle = |\varphi\rangle$, and we simply apply the operator S to both sides of (C-14) to show that $|\varphi\rangle$ can be expressed in the form of a linear combination of the various kets $S|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$.

However, it must be noted that the various kets $S|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$ are not independent. Let us permute the roles of the various particles in one of the kets $|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$ of the initial basis (before symmetrization). On this new ket, application of S or A leads, according to (B-62) and (B-63), to the same ket of $\mathcal{E}_S$ or $\mathcal{E}_A$ (possibly with a change of sign).

Thus, we are led to introduce the concept of an *occupation number*: by definition, for the ket $|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$, the occupation number n_k of the individual state $|u_k\rangle$ is equal to the number of times the state $|u_k\rangle$ appears in the sequence $\{|u_i\rangle, |u_j\rangle \ldots |u_p\rangle\}$, that is, the number of particles in the state $|u_k\rangle$ (we have, obviously, $\sum_k n_k = N$). Two different kets $|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$ for which the occupation numbers are equal can be obtained from each other by the action of a permutation operator. Consequently, after the action of the symmetrizer S (or the antisymmetrizer A), they give the same physical state, which we shall denote by $|n_1, n_2, \ldots, n_k, \ldots\rangle$:

$$|n_1, n_2, \ldots, n_k, \ldots\rangle = cS|\underbrace{1 : u_1 \ ; \ 2 : u_1 \ ; \ \ldots n_1 \ ; \ u_1}_{\substack{n_1 \text{ particles} \\ \text{in the state } |\ u_1\rangle}} \ ; \ \underbrace{n_1 + 1 : u_2 \ ; \ \ldots \ ; \ n_1 + n_2 : u_2}_{\substack{n_2 \text{ particles} \\ \text{in the state } |u_2\rangle}} \ ; \ \ldots\rangle \tag{C-15}$$

For fermions, S would be replaced by A in (C-15) (c is a factor which permits the normalization of the state obtained in this way[6]). We shall not study the states $|n_1, n_2, \ldots n_k, \ldots\rangle$ in detail here; we shall confine ourselves to giving some of their important properties:

[6] A simple calculation yields: $c = \sqrt{N!/n_1!n_2!\ldots}$ for bosons and $\sqrt{N!}$ for fermions.

(*i*) The scalar product of two kets $|n_1, n_2, \ldots n_k, \ldots\rangle$ and $|n'_1, n'_2, \ldots n'_k, \ldots\rangle$ is different from zero only if all the occupation numbers are equal ($n_k = n'_k$ for all k).

By using (C-15) and definitions (B-49) and (B-50) of S and A, we can obtain the expansion of the two kets under consideration on the orthonormal basis, $\{|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle\}$. It is then easy to see that, if the occupation numbers are not all equal, these two kets cannot simultaneously have non-zero components on the same basis vector.

(*ii*) If the particles under study are bosons, the kets $|n_1, n_2, \ldots n_k, \ldots\rangle$, in which the various occupation numbers n_k are arbitrary (with, of course $\sum_k n_k = N$) form an orthonormal basis of the physical state space.

Let us show that, for bosons, the kets $|n_1, n_2, \ldots n_k, \ldots\rangle$ defined by (C-15) are never zero. To do so, we replace S by its definition (B-49). There then appear, on the right-hand side of (C-15), various orthogonal kets $|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$, all with positive coefficients. The ket $|n_1, n_2, \ldots n_k, \ldots\rangle$ cannot, therefore, be zero.

The $|n_1, n_2, \ldots n_k, \ldots\rangle$ form a basis in $\mathcal{E}_S$ since these kets span $\mathcal{E}_S$, are all non-zero, and are orthogonal to each other.

(*iii*) If the particles under study are fermions, a basis of the physical state space $\mathcal{E}_A$ is obtained by choosing the set of kets $|n_1, n_2, \ldots n_k, \ldots\rangle$ in which all the occupation numbers are equal either to 1 or to 0 (again with $\sum_k n_k = N$).

The preceding proof is not applicable to fermions because of the minus signs which appear before the odd permutations in definition (B-50) of A. Furthermore, we saw in § c that two identical fermions cannot occupy the same individual quantum state: if any one of the occupation numbers is greater than 1, the vector defined by (C-15) is zero. On the other hand, it is never zero if all the occupation numbers are equal to one or zero; this is because two particles are then never in the same individual quantum state, so that the kets $|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$ and $P_\varepsilon|1 : u_i \ ; \ 2 : u_j \ ; \ \ldots \ ; \ N : u_p\rangle$ are always distinct and orthogonal. Relation (C-15) therefore defines a non-zero physical ket in this case. The rest of the proof is the same as for bosons.

C-4. Application of the other postulates

It remains for us to show how the general postulates of Chapter III can be applied in light of the symmetrization postulate introduced in § C-1, and to verify that no contradictions arise. More precisely, we shall see how measurement processes can be described with kets belonging only to either $\mathcal{E}_S$ or $\mathcal{E}_A$, and we shall show that the time evolution process does not take the ket $|\psi(t)\rangle$ associated with the state of the system out of this subspace. Thus, all the quantum mechanical formalism can be applied inside either $\mathcal{E}_S$ or $\mathcal{E}_A$.

C-4-a. Measurement postulates

α. Probability of finding the system in a given physical state

Consider a measurement performed on a system of identical particles. The ket $|\psi(t)\rangle$ describing the quantum state of the system before the measurement must, according to the symmetrization postulate, belong to $\mathcal{E}_S$ or to $\mathcal{E}_A$, depending on whether the system is formed of bosons or fermions. To apply the postulates of Chapter III

concerning measurements, we must take the scalar product of $|\psi(t)\rangle$ with the ket $|u\rangle$ corresponding to the physical state of the system after the measurement. This ket $|u\rangle$ is to be constructed by applying the rule given in § C-3-a. The probability amplitude $\langle u|\psi(t)\rangle$ can therefore be expressed in terms of two vectors, both belonging either to $\mathcal{E}_S$ or to $\mathcal{E}_A$. In § D-2, we shall discuss a certain number of examples of such calculations.

If the measurement envisaged is a "complete" measurement (yielding, for example, the positions and spin components S_z for all the particles), the physical ket $|u\rangle$ is unique (to within a constant factor). On the other hand, if the measurement is "incomplete" (for example, a measurement of the spins only, or a measurement bearing on a single particle), several orthogonal physical kets are obtained, and the corresponding probabilities must then be summed.

β. *Physical observables: invariance of $\mathcal{E}_S$ and $\mathcal{E}_A$*

In certain cases, it is possible to specify the measurement performed on the system of identical particles by giving the explicit expression of the corresponding observable in terms of $\mathbf{R}_1$, $\mathbf{P}_1$, $\mathbf{S}_1$, $\mathbf{R}_2$, $\mathbf{P}_2$, $\mathbf{S}_2$, etc.

We shall give some concrete examples of observables which can be measured in a three-particle system:

- Position of the center of mass $\mathbf{R}_G$, total momentum $\mathbf{P}$ and total angular momentum $\mathbf{L}$:

$$\mathbf{R}_G = \frac{1}{3}(\mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3) \tag{C-16}$$

$$\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2 + \mathbf{P}_3 \tag{C-17}$$

$$\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2 + \mathbf{L}_3 \tag{C-18}$$

- Electrostatic repulsion energy:

$$W = \frac{q^2}{4\pi\varepsilon_0}\left(\frac{1}{|\mathbf{R}_1 - \mathbf{R}_2|} + \frac{1}{|\mathbf{R}_2 - \mathbf{R}_3|} + \frac{1}{|\mathbf{R}_3 - \mathbf{R}_1|}\right) \tag{C-19}$$

- Total spin:

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3 \tag{C-20}$$

etc.

It is clear from these expressions that the observables associated with the physical quantities considered involve the various particles symmetrically. This important property follows directly from the fact that the particles are identical. In (C-16), for example, $\mathbf{R}_1$, $\mathbf{R}_2$ and $\mathbf{R}_3$ have the same coefficient, since the three particles have the same mass. It is the equality of the charges which is at the basis of the symmetric form of (C-19). In general, since no physical properties are modified when the roles of the N identical particles are permuted, these N particles must play a symmetric role[7] in any actually measurable observable. Mathematically, the corresponding observable G, which we shall call a *physical* observable, must be invariant under all permutations of the N identical

[7] Note that this reasoning is valid for fermions as well as for bosons.

particles. It must therefore commute with all the permutation operators P_α of the N particles (*cf.* § B-2-d):

$$[G, P_\alpha] = 0 \qquad \text{for all } P_\alpha \tag{C-21}$$

For a system of two identical particles, for example, the observable $\mathbf{R}_1 - \mathbf{R}_2$ (the vector difference of the positions of the two particles), which is not invariant under the effect of the permutation P_{21} ($\mathbf{R}_1 - \mathbf{R}_2$ changes sign) is not a physical observable; indeed, a measurement of $\mathbf{R}_1 - \mathbf{R}_2$ assumes that particle (1) can be distinguished from particle (2). On the other hand, we can measure the distance between the two particles, that is, $\sqrt{(\mathbf{R}_1 - \mathbf{R}_2)^2}$, which is symmetric.

Relation (C-21) implies that $\mathcal{E}_S$ and $\mathcal{E}_A$ are both invariant under the action of a physical observable G. Let us show that, if $|\psi\rangle$ belongs to $\mathcal{E}_A$, $G|\psi\rangle$ also belongs to $\mathcal{E}_A$ (the same proof also applies, of course, to $\mathcal{E}_S$). The fact that $|\psi\rangle$ belongs to $\mathcal{E}_A$ means that:

$$P_\alpha|\psi\rangle = \varepsilon_\alpha|\psi\rangle \tag{C-22}$$

Now let us calculate $P_\alpha G|\psi\rangle$. According to (C-21) and (C-22), we have:

$$P_\alpha G|\psi\rangle = GP_\alpha|\psi\rangle = \varepsilon_\alpha G|\psi\rangle \tag{C-23}$$

Since the permutation P_α is arbitrary, (C-23) expresses the fact that $G|\psi\rangle$ is completely antisymmetric and therefore belongs to $\mathcal{E}_A$.

All operations normally performed on an observable – in particular, the determination of eigenvalues and eigenvectors – can therefore be applied to G entirely within one of the subspaces, $\mathcal{E}_S$ or $\mathcal{E}_A$. Only the eigenkets of G belonging to the physical subspace, and the corresponding eigenvalues, are retained.

Comments:

(*i*) All the eigenvalues of G which exist in the total space $\mathcal{E}$ are not necessarily found if we restrict ourselves to the subspace $\mathcal{E}_S$ (or $\mathcal{E}_A$). The effect of the symmetrization postulate on the spectrum of a symmetric observable G may therefore be to abolish certain eigenvalues. On the other hand, it adds no new eigenvalues to this spectrum, since, because of the global invariance of $\mathcal{E}_S$ (or $\mathcal{E}_A$) under the action of G, any eigenvector of G in $\mathcal{E}_S$ (or $\mathcal{E}_A$) is also an eigenvector of G in $\mathcal{E}$ with the same eigenvalue.

(*ii*) Consider the problem of writing mathematically, in terms of the observables $\mathbf{R}_1$, $\mathbf{P}_1$, $\mathbf{S}_1$, etc., the observables corresponding to the different types of measurement envisaged in § α. This problem is not always simple. For example, for a system of three identical particles, we shall try to write the observables corresponding to the simultaneous measurement of the three positions in terms of $\mathbf{R}_1$, $\mathbf{R}_2$ and $\mathbf{R}_3$. We can resolve this problem by considering several physical observables chosen such that we can, using the results obtained by measuring them, unambiguously deduce the position of each particle (without, of course, being able to associate a numbered particle with each position). For example, we can choose the set:

$$X_1 + X_2 + X_3 \ , \ X_1X_2 + X_2X_3 + X_3X_1 \ , \ X_1X_2X_3$$

(and the corresponding observables for the Y and Z coordinates). However, this point of view is essentially formal. Rather than trying to write the expressions for the observables in all cases, it is simpler to follow the method used in § α, in which we confined ourselves to using the physical eigenkets of the measurement.

C-4-b. Time-evolution postulates

The Hamiltonian of a system of identical particles must be a physical observable. We shall write, for example, the Hamiltonian describing the motion of the two electrons of the helium atom about the nucleus, assumed to be motionless[8]:

$$H(1,2) = \frac{\mathbf{P}_1^2}{2m_e} + \frac{\mathbf{P}_2^2}{2m_e} - \frac{2e^2}{R_1} - \frac{2e^2}{R_2} + \frac{e^2}{|\mathbf{R}_1 - \mathbf{R}_2|} \tag{C-24}$$

The first two terms represent the kinetic energy of the system; they are symmetric because the two masses are equal. The next two terms are due to the attraction of the nucleus (whose charge is twice that of the proton). The electrons are obviously equally affected by this attraction. Finally, the last term describes the mutual interaction of the electrons. It is also symmetric, since neither of the two electrons is in a privileged position. It is clear that this argument can be generalized to any system of identical particles. Consequently, all the permutation operators commute with the Hamiltonian of the system:

$$[H, P_\alpha] = 0 \tag{C-25}$$

Under these conditions, if the ket $|\psi(t_0)\rangle$ describing the state of the system at a given time t_0 is a physical ket, the same must be true of the ket $|\psi(t)\rangle$ obtained from $|\psi(t_0)\rangle$ by solving the Schrödinger equation. According to this equation:

$$|\psi(t + \mathrm{d}t)\rangle = \left(1 + \frac{\mathrm{d}t}{i\hbar} H\right) |\psi(t)\rangle \tag{C-26}$$

Now, applying P_α and using relation (C-25):

$$P_\alpha |\psi(t + \mathrm{d}t)\rangle = (1 + \frac{\mathrm{d}t}{i\hbar} H) P_\alpha |\psi(t)\rangle \tag{C-27}$$

If $|\psi(t)\rangle$ is an eigenvector of P_α, $|\psi(t + \mathrm{d}t)\rangle$ is also an eigenvector of P_α, with the same eigenvalue. Since $|\psi(t_0)\rangle$, by hypothesis, is a completely symmetric or completely antisymmetric ket, this property is conserved over time.

The symmetrization postulate is therefore also compatible with the postulate that gives the time evolution of physical systems: the Schrödinger equation does not remove the ket $|\psi(t)\rangle$ from $\mathcal{E}_S$ or $\mathcal{E}_A$.

D. Discussion

In this final section, we shall examine the consequences of the symmetrization postulate on the physical properties of systems of identical particles. First of all, we shall indicate

[8] Here, we shall consider only the most important terms of this Hamiltonian. See Complement $\mathrm{B_{XIV}}$ for a more detailed study of the helium atom.

the fundamental differences introduced by Pauli's exclusion principle between systems of identical fermions and systems of identical bosons. Then, we shall discuss the implications of the symmetrization postulate concerning the calculation of the probabilities associated with the various physical processes.

D-1. Differences between bosons and fermions. Pauli's exclusion principle

In the statement of the symmetrization postulate, the difference between bosons and fermions may appear insignificant. Actually, this simple sign difference in the symmetry of the physical ket has extremely important consequences. As we saw in § C-3, the symmetrization postulate does not restrict the individual states accessible to a system of identical bosons. On the other hand, it requires fermions to obey Pauli's exclusion principle: two identical fermions cannot occupy the same individual quantum state.

The exclusion principle was formulated initially in order to explain the properties of many-electron atoms (§ D-1-a below and Complement A_{XIV}). It can now be seen to be more than a principle applicable only to electrons: it is a consequence of the symmetrization postulate, valid for all systems of identical fermions. Predictions based on this principle, which are often spectacular, have always been confirmed experimentally. We shall give some examples of them.

D-1-a. Ground state of a system of independent identical particles

The Hamiltonian of a system of identical particles (bosons or fermions) is always symmetric with respect to permutations of these particles (§ C-4). Consider such a system in which the various particles are independent, that is, do not interact with each other (at least in a first approximation). The corresponding Hamiltonian is then a sum of one-particle operators of the form:

$$H(1,2,\ldots,N) = h(1) + h(2) + \ldots + h(N) \tag{D-1}$$

$h(1)$ is a function only of the observables associated with the particle numbered (1); the fact that the particles are identical [which implies a symmetric Hamiltonian $H(1,2,\ldots,N)$] requires this function h to be the same in the N terms of expression (D-1). In order to determine the eigenstates and eigenvalues of the total Hamiltonian $H(1,2,\ldots,N)$, we simply calculate those of the individual Hamiltonian $h(j)$ in the state space $\mathcal{E}(j)$ of one of the particles:

$$h(j)|\varphi_n\rangle = e_n|\varphi_n\rangle \ ; \quad |\varphi_n\rangle \in \mathcal{E}(j) \tag{D-2}$$

For the sake of simplicity, we shall assume that the spectrum of $h(j)$ is discrete and non-degenerate.

If we are considering a system of identical bosons, the physical eigenvectors of the Hamiltonian $H(1,2,\ldots,N)$ can be obtained by symmetrizing the tensor products of N arbitrary individual states $|\varphi_n\rangle$:

$$|\phi^{(S)}_{n_1,n_2,\ldots,n_N}\rangle = c\sum_\alpha P_\alpha |1:\varphi_{n_1} \ ; \ 2:\varphi_{n_2} \ ; \ \ldots \ ; \ N:\varphi_{n_N}\rangle \tag{D-3}$$

where the corresponding energy is the sum of the N individual energies:

$$E_{n_1,n_2,\ldots,n_N} = e_{n_1} + e_{n_2} + \ldots + e_{n_N} \tag{D-4}$$

[it can easily be shown that each of the kets appearing on the right-hand side of (D-3) is an eigenket of H with the eigenvalue (D-4); this is also true of their sum]. In particular, if e_1 is the smallest eigenvalue of $h(j)$, and $|\varphi_1\rangle$ is the associated eigenstate, the ground state of the system is obtained when the N identical bosons are all in the state $|\varphi_1\rangle$. The energy of this ground state is therefore:

$$E_{1,1,\ldots,1} = Ne_1 \tag{D-5}$$

and its state vector is:

$$|\phi^{(S)}_{n_1,n_2,\ldots,n_N}\rangle = |1:\varphi_1\ ;\ 2:\varphi_1\ ;\ \ldots\ ;\ N:\varphi_1\rangle \tag{D-6}$$

Now, suppose that the N identical particles considered are fermions. It is no longer possible for these N particles all to be in the individual state $|\varphi_1\rangle$. To obtain the ground state of the system, Pauli's exclusion principle must be taken into account. If the individual energies e_n are arranged in increasing order:

$$e_1 < e_2 < \ldots < e_{n-1} < e_n < e_{n+1} < \ldots, \tag{D-7}$$

the ground state of the system of N identical fermions has an energy of:

$$E_{1,2,\ldots,N} = e_1 + e_2 + \ldots + e_N \tag{D-8}$$

and it is described by the normalized physical ket:

$$|\phi^{(A)}_{1,2,\ldots,N}\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} |1:\phi_1\rangle & |1:\phi_2\rangle & |1:\phi_3\rangle \\ |2:\phi_1\rangle & |2:\phi_2\rangle & |2:\phi_3\rangle \\ \vdots & & \\ |N:\phi_1\rangle & |N:\phi_2\rangle & |N:\phi_3\rangle \end{vmatrix} \tag{D-9}$$

The highest individual energy e_N found in the ground state is called the *Fermi energy* of the system.

Pauli's exclusion principle thus plays a role of primary importance in all domains of physics in which many-electron systems are involved, such as atomic and molecular physics (*cf.* Complements A_{XIV} and B_{XIV}) and solid state physics (*cf.* Complement C_{XIV}), and in all those in which many-proton and many-neutron systems are involved, such as nuclear physics[9].

Comment:

In most cases, the individual energies e_n are actually degenerate. Each of them can then enter into a sum such as (D-8) a number of times equal to its degree of degeneracy.

[9]The ket representing the state of a nucleus must be antisymmetric both with respect to the set of protons and with respect to the set of neutrons.

D-1-b. Quantum statistics

The object of statistical mechanics is to study systems composed of a very large number of particles (in numerous cases, the mutual interactions between these particles are weak enough to be neglected in a first approximation). Since we do not know the microscopic state of the system exactly, we content ourselves with describing it globally by its macroscopic properties (pressure, temperature, density, etc.). A particular macroscopic state corresponds to a whole set of microscopic states. We then use probabilities: the statistical weight of a macroscopic state is proportional to the number of distinct microscopic states that correspond to it, and the system, at thermodynamic equilibrium, is in its most probable macroscopic state (with any constraints that may be imposed taken into account). To study the macroscopic properties of the system, it is therefore essential to determine how many different microscopic states possess certain characteristics and, in particular, a given energy.

In classical statistical mechanics (Maxwell-Boltzmann statistics), the N particles of the system are treated as if they were of different natures, even if they are actually identical. Such a microscopic state is defined by specifying the individual state of each of the N particles. Two microscopic states are considered to be distinct when these N individual states are the same but the permutation of the particles is different.

In quantum statistical mechanics, the symmetrization postulate must be taken into account. A microscopic state of a system of identical particles is characterized by the enumeration of the N individual states which form it, the order of these states being of no importance since their tensor product must be symmetrized or anti-symmetrized. The numbering of the microscopic states therefore does not lead to the same result as in classical statistical mechanics. In addition, Pauli's principle radically differentiates systems of identical bosons and systems of identical fermions: the number of particles occupying a given individual state cannot exceed one for fermions, while it can take on any value for bosons (*cf.* § C-3). Different statistical properties result: bosons obey *Bose-Einstein statistics* and fermions, *Fermi-Dirac statistics*. This is the origin of the terms "bosons" and "fermions".

The physical properties of systems of identical fermions and systems of identical bosons are very different. This subject will be discussed in more detail in the first three chapters of Volume III. The differences can be observed, for example, at low temperatures, when the particles tend to accumulate in the individual states of lowest energy. Identical bosons may then exhibit a phenomenon called *Bose-Einstein condensation* of particles (Complements B_{XV} and C_{XV}); by contrast identical fermions, subject to the restrictions of Pauli's principle, build a Fermi sphere (Complement C_{XIV}) and can undergo only a pair condensation (Chapter XVII). Bose-Einstein condensation is at the origin of the remarkable properties (superfluidity) of the ^{4}He isotope of helium, in particular the superfluid properties (Complement D_{XV}) of its liquid at low temperatures (a few K) . The ^{3}He isotope, which is a fermion (*cf.* Comment of § C-1), has very different properties and is superfluid only at much lower temperatures because of pair condensation.

D-2. The consequences of particle indistinguishability on the calculation of physical predictions

In quantum mechanics, all the predictions concerning the properties of a system are expressed in terms of probability amplitudes (scalar products of two state vectors)

or matrix elements of an operator. It is then not surprising that the symmetrization or antisymmetrization of state vectors causes special *interference effects* to appear in systems of identical particles. First, we shall specify these effects, and then we shall see how they disappear under certain conditions (the particles of the system, although identical, then behave as if they were of different natures). To simplify the discussion, we shall confine ourselves to systems containing only two identical particles.

D-2-a. Interferences between direct and exchange processes

α. Predictions concerning a measurement on a system of identical particles: the direct term and the exchange term

Consider a system of two identical particles, one of which is known to be in the individual state $|\varphi\rangle$ and the other, in the individual state $|\chi\rangle$. We shall assume $|\varphi\rangle$ and $|\chi\rangle$ to be orthogonal, so that the state of the system is described by the normalized physical ket [*cf.* formula (C-4)]:

$$|\varphi\ ;\ \chi\rangle = \frac{1}{\sqrt{2}}[1+\varepsilon P_{21}]|1:\varphi\ ;\ 2:\chi\rangle \tag{D-10}$$

where:

$$\begin{aligned} &\varepsilon = +1 \quad \text{if the particles are bosons} \\ &\varepsilon = -1 \quad \text{if the particles are fermions} \end{aligned} \tag{D-11}$$

With the system in this state, suppose that we want to measure on each of the two particles the same physical quantity B with which the observables $B(1)$ and $B(2)$ are associated. For the sake of simplicity, we shall assume that the spectrum of B is entirely discrete and non-degenerate:

$$B|u_i\rangle = b_i|u_i\rangle \tag{D-12}$$

What is the probability of finding certain given values in this measurement (b_n for one of the particles and $b_{n'}$ for the other one)? We shall begin by assuming b_n and $b_{n'}$ to be different, so that the corresponding eigenvectors $|u_n\rangle$ and $|u_{n'}\rangle$ are orthogonal. Under these conditions, the normalized physical ket defined by the result of this measurement can be written:

$$|u_n\ ;\ u_{n'}\rangle = \frac{1}{\sqrt{2}}[1+\varepsilon P_{21}]|1:u_n\ ;\ 2:u_{n'}\rangle \tag{D-13}$$

which gives the probability amplitude associated with this result:

$$\langle u_n\ ;\ u_{n'}|\varphi\ ;\ \chi\rangle = \frac{1}{2}\langle 1:u_n\ ;\ 2:u_{n'}|(1+\varepsilon P_{21}^{\dagger})(1+\varepsilon P_{21})|1:\varphi\ ;\ 2:\chi\rangle \tag{D-14}$$

Using properties (B-13) and (B-14) of the operator P_{21}, we can write:

$$\frac{1}{2}(1+\varepsilon P_{21}^{\dagger})(1+\varepsilon P_{21}) = 1+\varepsilon P_{21} \tag{D-15}$$

(D-14) then becomes:

$$\langle u_n\ ;\ u_{n'}|\varphi\ ;\ \chi\rangle = \langle 1:u_n\ ;\ 2:u_{n'}|(1+\varepsilon P_{21})|1:\varphi\ ;\ 2:\chi\rangle \tag{D-16}$$

Letting $1 + \varepsilon P_{21}$ act on the bra, we obtain:

$$\begin{aligned}\langle u_n\ ;\ u_{n'}|\varphi\ ;\ \chi\rangle &= \langle 1 : u_n\ ;\ 2 : u_{n'}|(1 : \varphi\ ;\ 2 : \chi\rangle \\ &\quad + \varepsilon\langle 1 : u'_n\ ;\ 2 : u_n|1 : \varphi\ ;\ 2 : \chi\rangle \\ &= \langle 1 : u_n|1 : \varphi\rangle\langle 2 : u_{n'}|2 : \chi\rangle \\ &\quad + \varepsilon\langle 1 : u'_n|1 : \varphi\rangle\langle 2 : u_n|2 : \chi\rangle \\ &= \langle u_n|\varphi\rangle\langle u'_n|\chi\rangle + \varepsilon\langle u'_n|\varphi\rangle\langle u_n|\chi\rangle \end{aligned} \tag{D-17}$$

The numbering has disappeared from the probability amplitude, which is now expressed directly in terms of the scalar products $\langle u_n|\varphi\rangle \ldots \langle u_n|\chi\rangle$. Also, the probability amplitude appears either as a sum (for bosons) or a difference (for fermions) of two terms, with which we can associate the diagrams of Figures 4a and 4b.

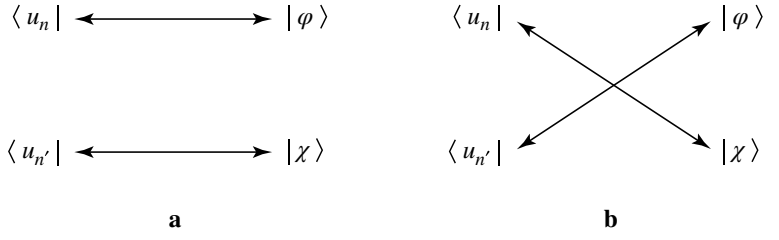


Figure 4: Schematic representation of the direct term and the exchange term associated with a measurement performed on a system of two identical particles. Before the measurement, one of the particles is known to be in the state $|\varphi\rangle$ and the other one, in the state $|\chi\rangle$. The measurement result obtained corresponds to a situation in which one particle is in the state $|u_n\rangle$ and the other one, in the state $|u_{n'}\rangle$. Two probability amplitudes are associated with such a measurement; they are represented schematically by figures a and b. These amplitudes interfere with a + sign for bosons and with a – sign for fermions.

We can interpret result (D-17) in the following way. The two kets $|\varphi\rangle$ and $|\chi\rangle$ associated with the initial state can be connected to the two bras $\langle u_n|$ and $\langle u_{n'}|$ associated with the final state by two different "paths", represented schematically by Figures 4a and 4b. With each of these paths is associated a probability amplitude, $\langle u_n|\varphi\rangle\langle u_{n'}|\chi\rangle$ or $\langle u_{n'}|\varphi\rangle\langle u_n|\chi\rangle$, and *these two amplitudes interfere with a + sign for bosons and a – sign for fermions.* Thus, we obtain the answer to the question posed in § A-3-a above: the desired probability $\mathcal{P}(b_n; b_{n'})$ is equal to the square of the modulus of (D-17):

$$\mathcal{P}(b_n; b_{n'}) = |\langle u_n|\varphi\rangle\langle u_{n'}|\chi\rangle + \varepsilon\langle u_{n'}|\varphi\rangle\langle u_n|\chi\rangle|^2 \tag{D-18}$$

One of the two terms on the right-hand side of (D-17), the one which corresponds, for example, to path 4-a, is often called the *direct term.* The other term is called the *exchange term.*

Comment:

Let us examine what happens if the two particles, instead of being identical, are of different natures. We shall then choose as the initial state of the system the tensor product ket:

$$|\psi\rangle = |1 : \varphi \ ; \ 2 : \chi\rangle \tag{D-19}$$

Now, consider a measurement instrument which, although the two particles, (1) and (2), are not identical, is not able to distinguish between them. If it yields the results b_n and $b_{n'}$, we do not know if b_n is associated with particle (l) or particle (2) (for example, for a system composed of a muon μ^- and an electron e^-, the measurement device may be sensitive only to the charge of the particles, giving no information about their masses). The two eigenstates $|1 : u_n \ ; \ 2 : u_{n'}\rangle$ and $|1 : u_{n'} \ ; \ 2 : u_n\rangle$ (which, in this case, represent different physical states) then correspond to the same measurement result. Since they are orthogonal, we must add the corresponding probabilities, which gives:

$$\begin{aligned}\mathcal{P}'(b_n; b_{n'}) &= |\langle 1 : u_n \ ; \ 2 : u_{n'}|1 : \varphi \ ; \ 2 : \chi\rangle|^2 \\ &\qquad + |\langle 1 : u_{n'} \ ; \ 2 : u_n|1 : \varphi \ ; \ 2 : \chi\rangle|^2 \\ &= |\langle u_n|\varphi\rangle|^2|\langle u_{n'}|\chi\rangle|^2 + |\langle u_{n'}|\varphi\rangle|^2|\langle u_n|\chi\rangle|^2\end{aligned} \tag{D-20}$$

Comparison of (D-18) with (D-20) clearly reveals the significant difference in the physical predictions of quantum mechanics depending on whether the particles under consideration are identical or not.

Now consider the case in which the two states $|u_n\rangle$ and $|u_{n'}\rangle$ are the same. When the two particles are fermions, the corresponding physical state is excluded by Pauli's principle, and the probability $\mathcal{P}(b_n; b_n)$ is zero. On the other hand, if the two particles are bosons, we have:

$$|u_n \ ; \ u_n\rangle = |1 : u_n \ ; \ 2 : u_n\rangle \tag{D-21}$$

and, consequently:

$$\begin{aligned}\langle u_n \ ; \ u_n|\varphi \ ; \ \chi\rangle &= \frac{1}{\sqrt{2}}\langle 1 : u_n \ ; \ 2 : u_n|(1 + P_{21})|1 : \varphi \ ; \ 2 : \chi\rangle \\ &= \sqrt{2}\langle u_n|\varphi\rangle\langle u_n|\chi\rangle\end{aligned} \tag{D-22}$$

which gives:

$$\mathcal{P}(b_n; b_n) = 2|\langle u_n|\varphi\rangle\langle u_n|\chi\rangle|^2 \tag{D-23}$$

Comments:

(i) Let us compare this result with the one which would be obtained in the case, already considered above, in which the two particles are different. We must

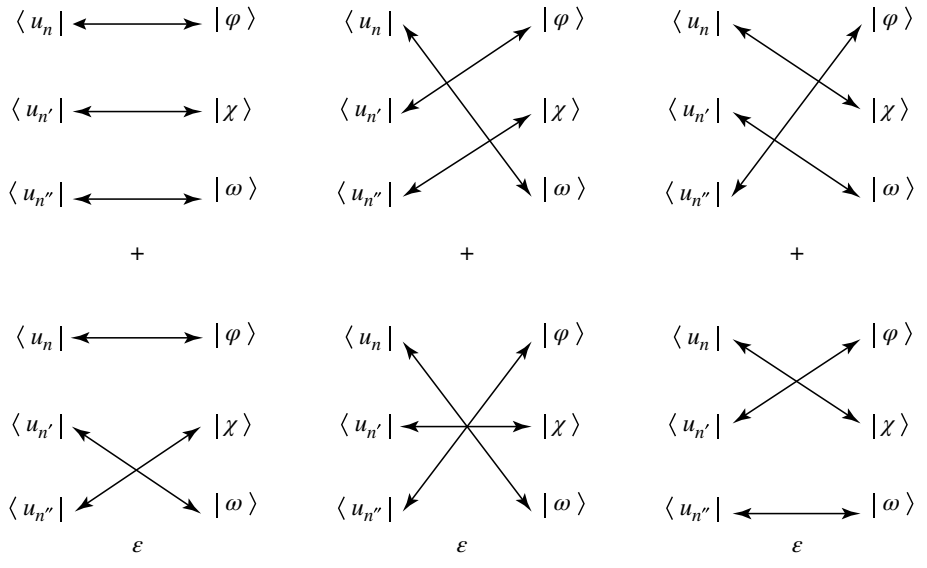


Figure 5: Schematic representation of the six probability amplitudes associated with a system of three identical particles. Before the measurement, one particle is known to be in the state $|\varphi\rangle$, another, in the state $|\chi\rangle$, and the last one, in the state $|\omega\rangle$. The result obtained corresponds to a situation in which one particle is in the state $|u_n\rangle$, another, in the state $|u_{n'}\rangle$, and the last one, in the state $|u_{n''}\rangle$. The six amplitudes interfere with a sign which is shown beneath each one ($\varepsilon = +1$ for bosons, -1 for fermions).

---

then replace $|\varphi\ ;\ \chi\rangle$ by $|1:\varphi\ ;\ 2:\chi\rangle$ and $|u_n\ ;\ u_n\rangle$ by $|1:u_n\ ;\ 2:u_n\rangle$, which gives the value for the probability amplitude:

$$\langle u_n|\varphi\rangle\langle u_n|\chi\rangle \tag{D-24}$$

and, consequently:

$$\mathcal{P}'(b_n;b_n) = |\langle u_n|\varphi\rangle\langle u_n|\chi\rangle|^2 \tag{D-25}$$

(*ii*) For a system containing N identical particles, there are, in general, $N!$ distinct exchange terms which add (or subtract) in the probability amplitude. For example, consider a system of three identical particles in the individual states $|\varphi\rangle$, $|\phi\rangle$ and $|\omega\rangle$, and the probability of finding, in a measurement, the results b_n, $b_{n'}$ and $b_{n''}$. The possible "paths" are then shown in Figure 5. There are six such paths (all different if the three eigenvalues, b_n, $b_{n'}$ and $b_{n''}$ are different). Some always contribute to the probability amplitude with a + sign, others with an ε sign (+ for bosons and − for fermions).

β. Example: elastic collision of two identical particles

To understand the physical meaning of the exchange term, let us examine a concrete example (already alluded to in § A-3-a): that of the elastic collision of two identical

particles in their center of mass frame[10]. Unlike the situation in § α above, here we must take into account the evolution of the system between the initial time when it is in the state $|\psi_i\rangle$ and the time t when the measurement is performed. However, as we shall see, this evolution does not change the problem radically, and the exchange term enters the problem as before.

In the initial state of the system (Fig. 6a), the two particles are moving towards each other with opposite momenta. We choose the Oz axis along the direction of these momenta, and we denote their modulus by p. One of the particles thus possesses the momentum $p\mathbf{e_z}$, and the other one, the momentum $-p\mathbf{e}_z$ (where $\mathbf{e}_z$ is the unit vector of the Oz axis). We shall write the physical ket $|\psi_i\rangle$ representing this initial state in the form:

$$|\psi_i\rangle = \frac{1}{\sqrt{2}}(1 + \varepsilon P_{21})|1 : p\mathbf{e}_z \ ; \ 2 : -p\mathbf{e}_z\rangle \tag{D-26}$$

$|\psi_i\rangle$ describes the state of the system at t_0, before the collision.

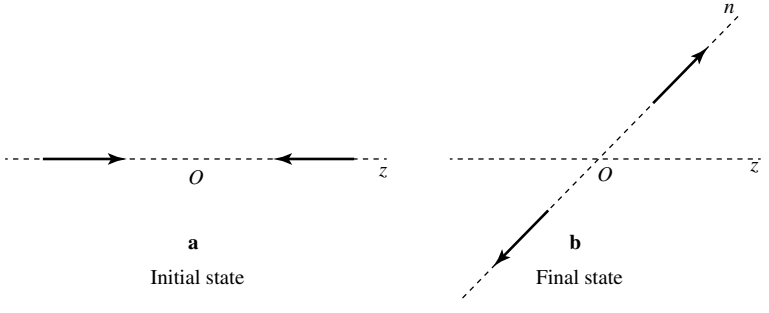


Figure 6: Collision between two identical particles in the center of mass frame: the momenta of the two particles in the initial state (fig. a) and in the final state found in the measurement (fig. b) are represented. For the sake of simplicity, we ignore the spin of the particles.

The Schrödinger equation which governs the time evolution of the system is linear. Consequently, there exists a linear operator $U(t, t')$, which is a function of the Hamiltonian H, such that the state vector at time t is given by:

$$|\psi(t)\rangle = U(t, t_0)|\psi_i\rangle \tag{D-27}$$

(Complement $\mathrm{F_{III}}$). In particular, after the collision, the state of the system at time t_1 is represented by the physical ket:

$$|\psi(t_1)\rangle = U(t_1, t_0)|\psi_i\rangle \tag{D-28}$$

[10]We shall give a simplified treatment of this problem, intended only to illustrate the relation between the direct term and the exchange term. In particular, we ignore the spin of the two particles. However, the calculations of this section remain valid in the case in which the interactions are not spin-dependent and the two particles are initially in the same spin state.

Note that, since the Hamiltonian H is symmetric, the evolution operator U commutes with the permutation operator:

$$[U(t,t'),P_{21}]=0 \tag{D-29}$$

Now, let us calculate the probability amplitude of the result envisaged in § A-3-a, in which the particles are detected in the two opposite directions of the *On* axis, of unit vector $\mathbf{n}$ (Fig. 6b). We denote the physical ket associated with this final state by:

$$|\psi_f\rangle=\frac{1}{\sqrt{2}}(1+\varepsilon P_{21})|1:p\mathbf{n}\ ;\ 2:-p\mathbf{n}\rangle \tag{D-30}$$

The desired probability amplitude can therefore be written:

$$\begin{aligned}\langle\psi_f|\psi(t_1)\rangle &= \langle\psi_f|U(t_1,t_0)|\psi_i\rangle\\ &= \frac{1}{2}\langle 1:p\mathbf{n}\ ;\ 2:-p\mathbf{n}|(1+\varepsilon P_{21}^{\dagger})\\ &\quad\times U(t_1,t_0)(1+\varepsilon P_{21})|1:p\mathbf{e}_z\ ;\ 2:-p\mathbf{e}_z\rangle\end{aligned} \tag{D-31}$$

According to relation (D-29) and the properties of the operator P_{21}, we finally obtain:

$$\begin{aligned}|\psi_f\rangle &= U(t,t_0)|\psi_i\rangle\\ &= \langle 1:p\mathbf{n}\ ;\ 2:-p\mathbf{n}|(1+\varepsilon P_{21}^{\dagger})U(t_1,t_0)|1:p\mathbf{e}_z\ ;\ 2:-p\mathbf{e}_z\rangle\\ &= \langle 1:p\mathbf{n}\ ;\ 2:-p\mathbf{n}|U(t_1,t_0)|1:p\mathbf{e}_z\ ;\ 2:-p\mathbf{e}_z\rangle\\ &\quad+\varepsilon\langle 1:-p\mathbf{n}\ ;\ 2:p\mathbf{n}|U(t_1,t_0)|1:p\mathbf{e}_z\ ;\ 2:-p\mathbf{e}_z\rangle\end{aligned} \tag{D-32}$$

The direct term corresponds, for example, to the process shown in Figure 7a, and the exchange term is then represented by Figure 7b. Again, the probability amplitudes associated with these two processes must be added or subtracted. This causes an interference term to appear when the square of the modulus of expression (D-32) is taken. Note also that this expression is simply multiplied by ε if $\mathbf{n}$ is changed to –$\mathbf{n}$, so that the corresponding probability is invariant under this change.

D-2-b. Situations in which the symmetrization postulate can be ignored

If the application of the symmetrization postulate were always indispensable, it would be impossible to study the properties of a system containing a restricted number of particles, because it would be necessary to take into account all the particles in the universe which are identical to those in the system. We shall see in this section that this is not the case. In fact, under certain special conditions, identical particles behave as if they were actually different, and it is not necessary to take the symmetrization postulate into account in order to obtain correct physical predictions. It seems natural to expect, considering the results of § D-2-a, that such a situation would arise whenever the exchange terms introduced by the symmetrization postulate are zero. We shall give two examples.

α. *Identical particles situated in two distinct regions of space*

Consider two identical particles, one of which is in the individual state $|\varphi\rangle$ and the other, in the state $|\chi\rangle$. To simplify the notation, we shall ignore their spin. Suppose that

the domain of the wave functions representing the kets $|\varphi\rangle$ and $|\chi\rangle$ are well separated in space:

$$\begin{cases} \varphi(\mathbf{r}) = \langle \mathbf{r}|\varphi\rangle = 0 \quad \text{if} \quad \mathbf{r} \notin D \\ \chi(\mathbf{r}) = \langle \mathbf{r}|\chi\rangle = 0 \quad \text{if} \quad \mathbf{r} \notin \Delta \end{cases} \tag{D-33}$$

where the domains D and Δ do not overlap. The situation is analogous to the classical mechanical one (§ A-2): as long as the domains D and Δ do not overlap, each of the particles can be "tracked"; we therefore expect application of the symmetrization postulate to be unnecessary.

In this case, we can envisage measuring an observable related to one of the two particles. All we need is a measurement device placed so that it cannot record what happens in the domain D, or in the domain Δ. If it is D which is excluded in this way, the measurement will only concern the particle in Δ, an vice versa.

Now, imagine a measurement concerning the two particles simultaneously, but performed with two distinct measurement devices, one of which is not sensitive to phenomena occurring in Δ, and the other, to those in D. How can the probability of obtaining a given result be calculated? Let $|u\rangle$ and $|v\rangle$ be the individual states associated respectively with the results of the two measurement devices. Since the two particles are identical, the symmetrization postulate must, in theory, be taken into account. In the probability amplitude associated with the measurement result, the direct term is then $\langle u|\varphi\rangle\langle v|\chi\rangle$, and the exchange term is $\langle u|\chi\rangle\langle v|\varphi\rangle$. Now, the spatial disposition of the measurement devices implies that:

$$\begin{aligned} u(\mathbf{r}) &= \langle \mathbf{r}|u\rangle = 0 \quad \text{if} \quad \mathbf{r} \in \Delta \\ v(\mathbf{r}) &= \langle \mathbf{r}|v\rangle = 0 \quad \text{if} \quad \mathbf{r} \in D \end{aligned} \tag{D-34}$$

According to (D-33) and (D-34), the wave functions $u(\mathbf{r})$ and $\chi(\mathbf{r})$ do not overlap; neither do $v(\mathbf{r})$ and $\varphi(\mathbf{r})$, so that:

$$\langle u|\chi\rangle = \langle v|\varphi\rangle = 0 \tag{D-35}$$

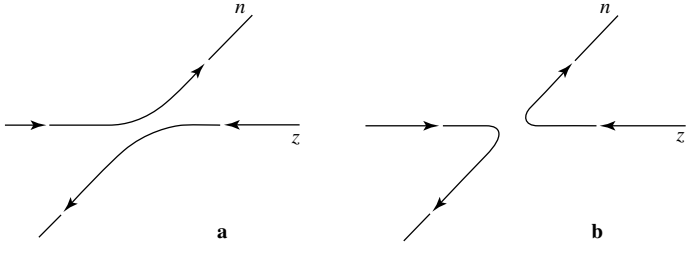


Figure 7: Collision between two identical particles in the center of mass frame: schematic representation of the physical processes corresponding to the direct term and the exchange term. The scattering amplitudes associated with these two processes interfere with a plus sign for bosons and a minus sign for fermions.

The exchange term is therefore zero. Consequently, it is unnecessary, in this situation, to use the symmetrization postulate. We obtain the desired result directly by reasoning as if the particles were of different natures, labeling, for example, the one in the domain D with the number 1, and the one situated in Δ with the number 2. Before the measurement, the state of the system is then described by the ket $|1 : \varphi \ ; \ 2 : \chi\rangle$, and with the measurement result envisaged is associated the ket $|1 : u \ ; \ 2 : v\rangle$. Their scalar product gives the probability amplitude $\langle u|\varphi\rangle\langle v|\chi\rangle$.

This argument shows that the existence of identical particles does not prevent the separate study of restricted systems, composed of a small number of particles.

Comment:

In the initial state chosen, the two particles are situated in two distinct regions of space. In addition, we have defined the state of the system by specifying two individual states. We might wonder if, after the system has evolved, it is still possible to study one of the two particles and ignore the other one. For this to be the case, it is necessary, not only that the two particles remain in two distinct regions of space, but also that they do not interact. Whether the particles are identical or not, an interaction always introduces correlations between them, and it is no longer possible to describe each of them by a state vector.

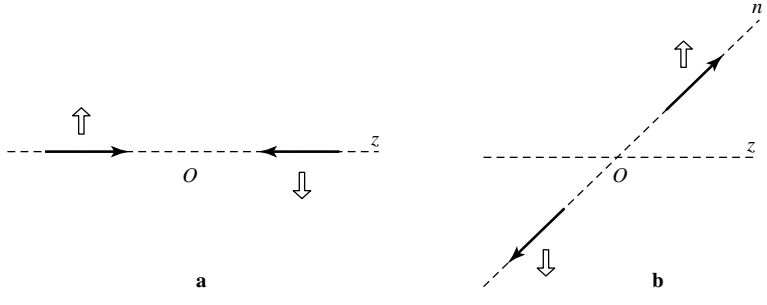


Figure 8: Collision between two identical spin 1/2 particles in the center of mass frame: a schematic representation of the momenta and spins of the two particles in the initial state (fig. a) and in the final state found in the measurement (fig. b). If the interactions between the two particles are spin-independent, the orientation of the spins does not change during the collision. When the two particles are not in the same spin state before the collision (the case of the figure), it is possible to determine the "path" followed by the system in arriving at a given final state. For example, the only scattering process which leads to the final state of figure b and which has a non-zero amplitude is of the type shown in Figure 7a.

β. Particles which can be identified by the direction of their spins

Consider an elastic collision between two identical spin 1/2 particles (electrons, for example), assuming that spin-dependent interactions can be neglected, so that the spin states of the two particles are conserved during the collision. If these spin states are initially orthogonal, they enable us to distinguish between the two particles at all times, as if they were not identical; consequently, the symmetrization postulate should again have no effect here.

We can show this, using the calculation of § D-2-a-β. The initial physical ket will be, for example (Fig. 8a):

$$|\psi_i\rangle = \frac{1}{\sqrt{2}}(1 - P_{21})|1 : p\mathbf{e}_z, + \ ; \ 2 : -p\mathbf{e}_z, -\rangle \tag{D-36}$$

(where the symbol + or − added after each momentum indicates the sign of the spin component along a particular axis). The final state we are considering (Fig. 8b) will be described by:

$$|\psi_f\rangle = \frac{1}{\sqrt{2}}(1 - P_{21})|1 : p\mathbf{n}, + \ ; \ 2 : -p\mathbf{n}, -\rangle \tag{D-37}$$

Under these conditions, only the first term of (D-32) is different from zero, since the second one can be written:

$$\langle 1 : -p\mathbf{n}, - \ ; \ 2 : p\mathbf{n}, +|U(t_1, t_0)|1 : p\mathbf{e}_z, + \ ; \ 2 : -p\mathbf{e}_z, -\rangle \tag{D-38}$$

This is the matrix element of a spin-independent operator (by hypothesis) between two kets whose spin states are orthogonal; it is therefore zero. Consequently, we would obtain the same result if we treated the two particles directly as if they were different, that is, if we did not antisymmetrize the initial and final kets and if we associated index 1 with the spin state $|+\rangle$ and index 2 with the spin state $|-\rangle$. Of course, this is no longer possible if the evolution operator U, that is, the Hamiltonian H of the system, is spin-dependent.

References and suggestions for further reading:

The importance of interference between direct and exchange terms is stressed in Feynman III (1.2), § 3.4 and Chap. 4.

Quantum statistics: Reif (8.4). Kittel (8.2).

Permutation groups: Messiah (1.17), app. D, § IV; Wigner (2.23), Chap. 13; Bacry (10.31), §§ 41 and 42.

The effect of the symmetrization postulate on molecular spectra: Herzberg (12.4), Vol. I, Chap. III, § 2f.

An article giving a popularized version: Gamow (1.27).

A_{XIV} : MANY-ELECTRON ATOMS; ELECTRONIC CONFIGURATIONS	Simple study of many-electron atoms in the central-field approximation. Discusses the consequences of the Pauli exclusion principle and introduces the concept of an electronic configuration. Remains qualitative.
B_{XIV} : ENERGY LEVELS OF THE HELIUM ATOM: CONFIGURATIONS, TERMS, MULTIPLETS	Study, in the case of the helium atom, of the effect of the electronic repulsion between electrons and of the magnetic interactions. Introduces the concepts of terms and multiplets. Can be reserved for later study.
C_{XIV} : PHYSICAL PROPERTIES OF AN ELECTRON GAS. APPLICATION TO SOLIDS	Study of the ground stante of a gas of free electrons enclosed in a "box". Introduces the concept of Fermi energy and periodic boundary conditions. Generalization to electrons in solids and qualitative discussion of the relation between electrical conductivity and the position of the Fermi level. Moderately difficult. The physical discussions are emphasized. Can be considered to be a sequel of F_{XI}.
D_{XIV} : EXERCISES	

Complement A_{XIV}

Many-electron atoms. Electronic configurations

1 The central-field approximation 1459
1-a Difficulties related to electron interactions 1460
1-b Principle of the method 1460
1-c Energy levels of the atom 1461
2 Electron configurations of various elements 1463

The energy levels of the hydrogen atom were studied in detail in Chapter VII. Such a study is considerably simplified by the fact that the hydrogen atom possesses a single electron, so that Pauli's principle is not relevant. In addition, by using the center of mass frame, we can reduce the problem to the calculation of the energy levels of a single particle (the relative particle) subjected to a central potential.

In this complement, we shall consider many-electron atoms, for which these simplifications cannot be made. In the center-of-mass frame, we must solve a problem involving several non-independent particles. This is a complex problem and we shall give only an approximate solution, using the central-field approximation (which will be outlined, without going into details of the calculations). In addition Pauli's principle, as we shall show, plays an important role.

1. The central-field approximation

Consider a Z-electron atom. Since the mass of its nucleus is much larger (several thousand times) than that of the electrons, the center-of-mass of the atom practically coincides with the nucleus, which we shall therefore assume to be motionless at the coordinate origin[1]. The Hamiltonian describing the motion of the electrons, neglecting relativistic corrections and, in particular, spin-dependent terms, can be written:

$$H = \sum_{i=1}^{z} \frac{\mathbf{P}_i^2}{2m_e} - \sum_{i=1}^{z} \frac{Ze^2}{R_i} + \sum_{i<j} \frac{e^2}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{1}$$

We have numbered the electrons arbitrarily from 1 to Z, and we have set:

$$e^2 = \frac{q^2}{4\pi\varepsilon_0} \tag{2}$$

where q is the electron charge. The first term of the Hamiltonian (1) represents the total kinetic energy of the system of Z electrons. The second one arises from the attraction exerted on each of them by the nucleus, which bears a positive charge equal to $-Zq$. The last one describes the mutual repulsion of the electrons [note that the summation is carried out over the $Z(Z-1)/2$ different ways of pairing the Z-electrons].

The Hamiltonian (1) is too complicated for us to solve its eigenvalue equation exactly, even in the simplest case, that of helium ($Z = 2$).

[1] Making this approximation amounts to neglecting the nuclear finite mass effect.

1-a. Difficulties related to electron interactions

In the absence of the mutual interaction term $\sum_{i<j} \frac{e^2}{|\mathbf{R}_i - \mathbf{R}_j|}$ in H, the electrons would be independent. It would then be easy to determine the energies of the atom. We would simply sum the energies of the Z electrons placed individually in the Coulomb potential $-Ze^2/r$, and the theory presented in Chapter VII would yield the result immediately. As for the eigenstates of the atom, they could be obtained by antisymmetrizing the tensor product of the stationary states of the various electrons.

It is then the presence of the mutual interaction term that makes it difficult to solve the problem exactly. We might try to treat this term by perturbation theory. However, a rough evaluation of its relative magnitude shows that this would not yield a good approximation. We expect the distance $|\mathbf{R}_i - \mathbf{R}_j|$ between two electrons to be, on the average, roughly the distance R_i of an electron from the nucleus. The ratio ρ of the third term of formula (1) to the second one is therefore approximately equal to:

$$\rho \simeq \frac{\frac{1}{2}Z(Z-1)}{Z^2} \tag{3}$$

ρ varies between 1/4 for $Z = 2$ and 1/2 for Z much larger than 1. Consequently, the perturbation treatment of the mutual interaction term would yield, at most, more or less satisfactory results for helium ($Z = 2$), but it is out of the question to apply it to other atoms (ρ is already equal to 1/3 for $Z = 3$). A more elaborate approximation method must therefore be found.

1-b. Principle of the method

To understand the concept of a central field, we shall use a semi-classical argument. Consider a particular electron (i). In a first approximation, the existence of the $Z - 1$ other electrons affects it only because their charge distribution partially compensates the electrostatic attraction of the nucleus. In this approximation, the electron (i) can be considered to move in a potential that depends only on its position $\mathbf{r}_i$ and takes into account the average effect of the repulsion of the other electrons. We choose a potential $V_c(r_i)$ that depends only on the modulus of $\mathbf{r}_i$ and call it the "central potential" of the atom under consideration. Of course, this can only be an approximation: since the motion of the electron (i) actually influences that of the $(Z - 1)$ other electrons, it is not possible to ignore the correlations which exist between them. Moreover, when the electron (i) is in the immediate vicinity of another electron (j), the repulsion exerted by the latter becomes preponderant, and the corresponding force is not central. However, the idea of an average potential appears more valid in quantum mechanics, where we consider the delocalization of the electrons as distributing their charges throughout an extended region of space.

These considerations lead us to write the Hamiltonian (1) in the form:

$$H = \sum_{i=1}^{z} \left[\frac{\mathbf{P}_i^2}{2m_e} + V_c(R_i) \right] + W \tag{4}$$

with:

$$W = -\sum_{i=1}^{z} \frac{Ze^2}{R_i} + \sum_{i<j} \frac{e^2}{|\mathbf{R}_i - \mathbf{R}_j|} - \sum_{i=1}^{z} V_c(R_i) \tag{5}$$

If the central potential $V_c(r_i)$ is suitably chosen, W should play the role of a small correction in the Hamiltonian H. The central-field approximation then consists of neglecting this correction, that is of choosing the approximate Hamiltonian:

$$H_0 = \sum_{i=1}^{z} \left[\frac{\mathbf{P}_i^2}{2m_e} + V_c(R_i) \right] \tag{6}$$

W will then be treated like a perturbation of H_0 (*cf.* Complement B_{XIV}, § 2). The diagonalization of H_0 leads to a problem of *independent particles*: to obtain the eigenstates of H_0, we simply determine those of the one-electron Hamiltonian:

$$\frac{\mathbf{P}^2}{2m_e} + V_c(R) \tag{7}$$

Definitions (4) and (5) do not, of course, determine the central potential $V_c(r)$, since we always have $H = H_0 + W$, for all $V_c(r)$. However, in order to treat W like a perturbation, $V_c(r)$ must be wisely chosen. We shall not take up the problem of the existence and determination of such an optimal potential here. This is a complex problem. The potential $V_c(r)$ to which a given electron is subjected depends on the spatial distribution of the $(Z-1)$ other electrons, and this distribution, in turn, depends on the potential $V_c(r)$, since the wave functions of the $(Z-1)$ electrons must also be calculated from $V_c(r)$. We must therefore arrive at a coherent solution (one generally says "self-consistent"), for which the wave functions determined from $V_c(r)$ give a charge distribution which reconstitutes this same potential $V_c(r)$.

1-c. Energy levels of the atom

While the exact determination of the potential $V_c(r)$ requires rather long calculations, the short- and long-distance behavior of this potential is simple to predict. We expect, for small r, the electron (i) under consideration to be inside the charge distribution created by the other electrons, so that it "sees" only the attractive potential of the nucleus. On the other hand, for large r, that is, outside the "cloud" formed by the $(Z-1)$ electrons treated globally, it is as if we had a single point charge situated at the coordinate origin and equal to the sum of the charges of the nucleus and the "cloud" [the $(Z-1)$ electrons screen the field of the nucleus]. Consequently (Fig. 1):

$$\begin{aligned} V_c(r) &\backsimeq -\frac{e^2}{r} \quad \text{for large } r \\ V_c(r) &\backsimeq -\frac{Ze^2}{r} \quad \text{for small } r \end{aligned} \tag{8}$$

For intermediate values of r, the variation of $V_c(r)$ can be more or less complicated, depending on the atom under consideration.

Although these considerations are qualitative, they give an idea of the spectrum of the one-electron Hamiltonian (7). Since $V_c(r)$ is not simply proportional to $1/r$, the accidental degeneracy found for the hydrogen atom (Chap. VII, § C-4-b) is no longer observed. The eigenvalues of the Hamiltonian (7) depend on the two quantum numbers n and l [however, they remain independent of m, since $V_c(r)$ is central]. l, of course, characterizes the eigenvalue of the operator L^2, and n is, by definition (as for the hydrogen

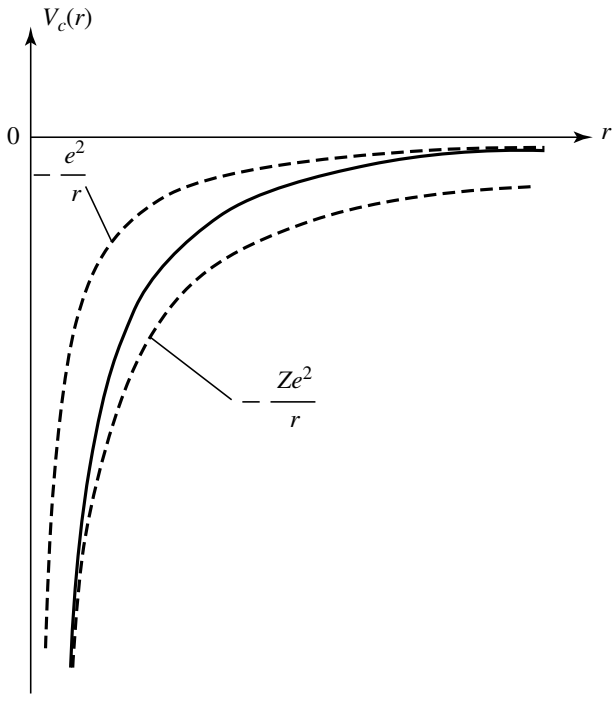


Figure 1: Variation of the central potential $V_c(r)$ with respect to r. The dashed-line curves represent the behavior of this potential at short distances $(-Ze^2/r)$ and at long distances $(-e^2/r)$.

atom), the sum of the azimuthal quantum number l, and the radial quantum number k introduced in solving the radial equation corresponding to l; n and l are therefore integer and satisfy:

$$0 \leqslant l \leqslant n-1 \tag{9}$$

Obviously, for a given value of l, the energies $E_{n,l}$ increase with n:

$$E_{n,l} > E_{n',l} \quad \text{if } n > n' \tag{10}$$

For fixed n, the energy is lower when the corresponding eigenstate is more "penetrating", that is, when the probability density of the electron in the vicinity of the nucleus is larger [according to (8), the screening effect is then smaller]. The energies $E_{n,l}$ associated with the same value of n can therefore be arranged in order of increasing angular momenta:

$$E_{n,0} < E_{n,1} < \ldots < E_{n,n-1} \tag{11}$$

It so happens that the hierarchy of states is approximately the same for all atoms, although the absolute values of the corresponding energies obviously vary with Z. Figure 2

indicates this hierarchy, as well as the $2(2l+1)$-fold degeneracy of each state (the factor 2 comes from the electron spin). The various states are represented in spectroscopic notation (*cf.* Chap. VII, § C-4-b). Those shown inside the same brackets are very close to each other, and may even, in certain atoms, practically coincide (we stress the fact that Figure 2 is simply a schematic representation intended to situate the eigenvalues $E_{n,l}$ with respect to each other; no attempt is made to establish an even moderately realistic energy scale).

Note the great difference between the energy spectrum shown and that of the hydrogen atom (*cf.* Chap. VII, Fig. 4). As we have already pointed out, the energy depends here on the orbital quantum number l, and, in addition, the order of the states is different. For example, Figure 2 indicates that the $4s$ shell has a slightly lower energy than that of the $3d$ shell. This is explained, as mentioned above, by the fact that the $4s$ wave function is more penetrating. Analogous inversions occur for the $n = 4$ and $n = 5$ shells, etc. This demonstrates the importance of inter-electron repulsion.

2. Electron configurations of various elements

In the central-field approximation, the eigenstates of the total Hamiltonian H_0 of the atom are Slater determinants, constructed from the individual electron states associated with the energy states $E_{n,l}$ that we have just described. This is therefore the situation envisaged in § D-1-a of Chapter XIV: the ground state of the atom is obtained when the Z electrons occupy the lowest individual states compatible with Pauli's principle. The maximum number of electrons that can have a given energy $E_{n,l}$ is equal to the $2(2l + 1)$-fold degeneracy of this energy level. The set of individual states associated with the same energy $E_{n,l}$ is called a *shell*. The list of occupied shells with the number of electrons found in each is called the *electronic configuration*. The notation used will be specified below in a certain number of examples. The concept of a configuration also plays an important role in the chemical properties of atoms. Knowledge of the wave functions of the various electrons and of the corresponding energies makes it possible to interpret the number, stability, and geometry of the chemical bonds which can be formed by this atom (*cf.* Complement $\mathrm{E_{XII}}$).

To determine the electronic configuration of a given atom in its ground state, we simply "fill" the various shells successively, in the order indicated in Figure 2 (starting, of course, with the $1s$ level), until the Z electrons are exhausted. This is what we shall do, in a rapid review of Mendeleev's table.

In the ground state of the hydrogen atom, the single electron of this atom occupies the $1s$ level. The electronic configuration of the next element (helium, $Z = 2$) is:

$$\mathrm{He} : 1s^2 \tag{12}$$

which means that the two electrons occupy the two orthogonal states of the $1s$ shell (same spatial wave function, orthogonal spin states). Then comes lithium ($Z = 3$), whose electronic configuration is:

$$\mathrm{Li} : 1s^2, 2s \tag{13}$$

The $1s$ shell can accept only two electrons, so the third one must go into the level directly above it, that is, according to Figure 2, into the $2s$ shell. This shell can accept a second

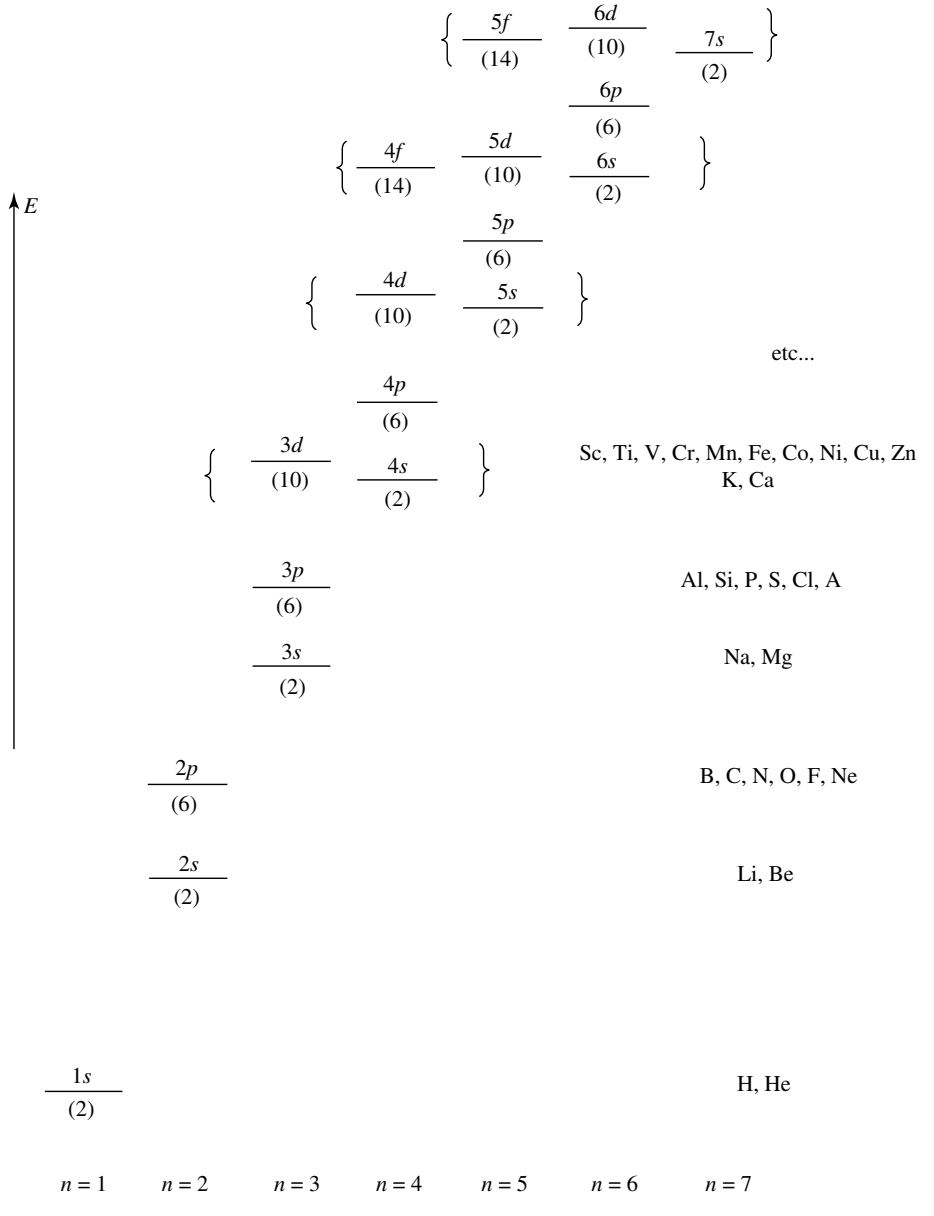


Figure 2: Schematic representation of the hierarchy of energy levels (electronic shells) in a central potential of the type shown in Figure 1. For each value of n, the energy increases with l. The degeneracy of each level is indicated in parentheses. The levels that appear inside the same bracket are very close to each other, and their relative disposition can vary from one atom to another. On the right-hand side of the figure, we have indicated the chemical symbols of the atoms for which the electronic shell appearing on the same line is the outermost shell occupied in the ground state configuration.

electron, which gives beryllium ($Z = 4$) the electronic configuration:

$$\text{Be} : 1s^2, 2s^2 \tag{14}$$

For $Z > 4$, the $2p$ shell (*cf.* Fig. 2) is the first to be gradually filled, and so on. As the number Z of electrons increases, higher and higher electronic shells are brought in (on the right-hand side of Figure 2, we have shown, opposite each of the lowest shells, the symbols of the atoms for which this shell is the outermost). Thus, we obtain the configurations of the ground state for all the atoms. This explains Mendeleev's classification. However, it must be noted that levels that are very close to each other (those grouped in brackets in Figure 2) may be filled in a very irregular fashion. For example, although Figure 2 gives the $4s$ shell a lower energy than that of the $3d$ shell, chromium ($Z = 24$) has five $3d$ electrons although the $4s$ shell is incomplete. Similar irregularities arise for copper ($Z = 29$), niobium ($Z = 41$), etc.

Comments:

(*i*) The electronic configurations which we have analyzed characterize the ground state of various atoms in the central-field approximation. The lowest excited states of the Hamiltonian H_0 are obtained when one of the electrons moves to an individual energy level which is higher than the last shell occupied in the ground state. We shall see, for example, in Complement B_{XIV}, that the first excited configuration of the helium atom is:

$$1s,\ 2s \tag{15}$$

(*ii*) A single non-zero Slater determinant is associated with an electronic configuration ending with a complete shell, since there are then as many orthogonal individual states as there are electrons. Thus, the ground state of the rare gases ($\ldots$, ns^2, np^6) is non-degenerate, as is that of the alkaline-earths ($\ldots$, ns^2). On the other hand, when the number of external electrons is smaller than the degree of degeneracy of the outermost shell, the ground state of the atom is degenerate. For the alkalines ($\ldots, ns$), the degree of degeneracy is equal to 2; for carbon ($1s^2$, $2s^2$, $2p^2$), it is equal to $C_6^2 = 15$, since two individual states can be chosen arbitrarily from the six orthogonal states constituting the $2p$ shell.

(*iii*) It can be shown that, for a complete shell, the total angular momentum is zero, as are the total orbital angular momentum and the total spin (the sums, respectively, of the orbital angular momenta and the spins of the electrons occupying this shell). Consequently, the angular momentum of an atom[2] is due only to its outer electrons. Thus, the total angular momentum of a helium atom in its ground state is zero, and that of an alkali metal is equal to 1/2 (a single external electron of zero orbital angular momentum and spin 1/2).

[2]The angular momentum being discussed here is that of the electronic cloud of the atom. The nucleus also possesses an angular momentum which should be added to this one.

References and suggestions for further reading:

Pauling and Wilson (1.9), Chap. IX; Levine (12.3), Chap. 11, § 1, 2 and 3; Kuhn (11.1), Chap. IV, §§ A and B; Schiff (1.18), § 47; Slater (1.6), Chap. 6; Landau and Lifshitz (1.19), §§ 68, 69 and 70. See also references of Chap. XI (Hartree and Hartree-Fock methods).

The shell model in nuclear physics: Valentin (16.1), Chap. VI; Preston (16.4), Chap. 7; Deshalit and Feshbach (16.6), Chap. IV and V. See also articles by Mayer (16.20); Peierls (16.21) and Baranger (16.22).

Complement B_{XIV}

Energy levels of the helium atom. Configurations, terms, multiplets

1 The central-field approximation. Configurations 1467

1-a The electrostatic Hamiltonian 1467

1-b The ground state configuration and first excited configurations 1468

1-c Degeneracy of the configurations 1468

2 The effect of the inter-electron electrostatic repulsion: exchange energy, spectral terms 1469

2-a Choice of a basis of $\mathcal{E}(n, l; n', l')$ adapted to the symmetries of W 1470

2-b Spectral terms. Spectroscopic notation 1472

2-c Discussion 1473

3 Fine-structure levels; multiplets 1478

In the preceding complement, we studied many-electron atoms in the central-field approximation in which the electrons are independent. This enabled us to introduce the concept of a configuration. We shall evaluate the corrections that must be made to this approximation, taking into account the inter-electron electrostatic repulsion more precisely. In order to simplify the reasoning, we shall confine ourselves to the simplest many-electron atom, the helium atom. We shall show that, under the effect of the inter-electron electrostatic repulsion, the configurations of this atom (§ 1) split into spectral terms (§ 2), which give rise to fine-structure multiplets (§ 3) when smaller terms in the atomic Hamiltonian (magnetic interactions) are taken into account. The concepts we shall bring out in this treatment can be generalized to more complex atoms.

1. The central-field approximation. Configurations

1-a. The electrostatic Hamiltonian

As in the preceding complement, we shall take into account only the electrostatic forces at first, writing the Hamiltonian of the helium atom [formula (C-24) of Chapter XIV] in the form:

$$H = H_0 + W \tag{1}$$

where:

$$H_0 = \frac{\mathbf{P}_1^2}{2m_e} + \frac{\mathbf{P}_2^2}{2m_e} + V_c(R_1) + V_c(R_2) \tag{2}$$

and:

$$W = -\frac{2e^2}{R_1} - \frac{2e^2}{R_2} + \frac{e^2}{|\mathbf{R}_1 - \mathbf{R}_2|} - V_c(R_1) - V_c(R_2) \tag{3}$$

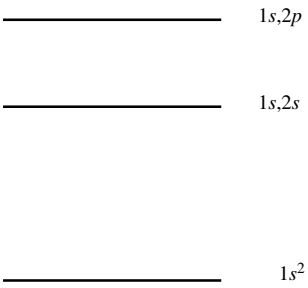


Figure 1: The ground state configuration and first excited configurations of the helium atom (the energies are not shown to scale).

The central potential $V_c(r)$ is chosen so as to make W a small correction of H_0.

When W is neglected, the electrons can be considered to be independent (although their average electrostatic repulsion is partially taken into account by the potential V_c). The energy levels of H_0 then define the electronic configurations we shall study in this section. We shall then examine the effect of W by using stationary perturbation theory in § 2.

1-b. The ground state configuration and first excited configurations

According to the discussion of Complement A_{XIV} (§ 2), the configurations of the helium atom are specified by the quantum numbers n, l and n', l' of the two electrons (placed in the central potential V_c). The corresponding energy E_c can be written:

$$E_c = E_{n,l} + E_{n',l'} \tag{4}$$

Thus (Fig. 1), the ground state configuration, written $1s^2$, is obtained when the two electrons are in the $1s$ shell; the first excited configuration, $1s\ 2s$, when one electron is in the $1s$ shell and the other one is in the $2s$ shell. Similarly, the second excited configuration is the $1s$, $2p$ configuration.

The excited configurations of the helium atom are of the form $1s$, $n'l'$. Actually, there also exist "doubly excited" configurations of the type nl, $n'l'$ (with n, $n' > 1$). But, for helium, their energy is greater than the ionization energy E_I of the atom (the limit of the energy of the configuration $1s$, $n'l'$ when $n' \to \infty$). Most of the corresponding states, therefore, are very unstable: they tend to dissociate rapidly into an ion and an electron and are called "autoionizing states". However, there exist levels belonging to doubly excited configurations which are not autoionizing, but which decay by emitting photons. Some of the corresponding spectral lines have been observed experimentally.

1-c. Degeneracy of the configurations

Since V_c is central and not spin-dependent, the energy of a configuration does not depend on the magnetic quantum numbers m and m' ($-l \leqslant m \leqslant l$, $-l' \leqslant m' \leqslant l'$) or on the spin quantum numbers ε and ε' ($\varepsilon = \pm$, $\varepsilon' = \pm$) associated with the two electrons. Most of the configurations, therefore, are degenerate; it is this degeneracy we shall now calculate.

A state belonging to a configuration is defined by specifying the four quantum numbers (n, l, m, ε) and $(n', l', m', \varepsilon')$ of each electron. Since the electrons are identical

particles, the symmetrization postulate must be taken into account. The physical ket associated with this state can, according to the results of § C-3-b of Chapter XIV, be written in the form:

$$|n,l,m,\varepsilon;n',l',m',\varepsilon'>=\frac{1}{\sqrt{2}}(1-P_{21})|1:n,l,m,\varepsilon;2:n',l',m',\varepsilon'> \tag{5}$$

Pauli's principle excludes the states of the system for which the two electrons would be in the same individual quantum state ($n=n'$, $l=l'$, $m=m'$, $\varepsilon=\varepsilon'$). According to the discussion of § C-3-b of Chapter XIV, the set of physical kets (5) for which n, l, n', l' are fixed and which are not null (that is, not excluded by Pauli's principle) constitute an orthonormal basis in the subspace $\mathcal{E}(n,l;n',l')$ of $\mathcal{E}_A$ associated with the configuration nl, $n'l'$.

To evaluate the degeneracy of a configuration $nl,n'l'$, we shall distinguish between two cases:

(i) The two electrons are not in the same shell (we do not have $n=n'$ and $l=l'$).

The individual states of the two electrons can never coincide, and m, m', ε, ε' can independently take on any value. The degeneracy of the configuration, consequently, is equal to:

$$2(2l+1)\times 2(2l'+1)=4(2l+1)(2l'+1) \tag{6}$$

The $1s$, $2s$ and $1s$, $2p$ configurations enter into this category; their degeneracies are equal to 4 and 12 respectively.

(ii) The two electrons are in the same shell ($n=n'$ and $l=l'$).

In this case, the states for which $m=m'$ and $\varepsilon=\varepsilon'$ must be excluded. Since the number of distinct individual quantum states is equal to $2(2l+1)$, the degree of degeneracy of the nl^2 configuration is equal to the number of pairs that can be formed from these individual states (*cf.* § C-3-b of Chapter XIV), that is:

$$C^2_{2(2l+1)}=(2l+1)(4l+1) \tag{7}$$

Thus, the $1s^2$ configuration, which enters into this category, is not degenerate. It is useful to expand the Slater determinant corresponding to this configuration. If, in (5), we set $n=n'=1$, $l=l'=m=m'=0$, $\varepsilon=+$, $\varepsilon'=-$, we obtain, writing the spatial part as a common factor:

$$|1s^2>=|1:1,0,0;2:1,0,0>\otimes\frac{1}{\sqrt{2}}(|1:+;2:->-|1:-;2:+>) \tag{8}$$

In the spin part of (8), we recognize the expression for the singlet state $|S=0,\ M_s=0>$, where S and M_s are the quantum numbers related to the total spin $\mathbf{S}=\mathbf{S}_1+\mathbf{S}_2$ (*cf.* Chap. X, § B-4). Thus, although the Hamiltonian H_0 does not depend on the spins, the constraints introduced by the symmetrization postulate require the total spin of the ground state to have the value $S=0$.

2. The effect of the inter-electron electrostatic repulsion: exchange energy, spectral terms

We shall now study the effect of W by using stationary perturbation theory. To do so, we must diagonalize the restriction of W inside the subspace $\mathcal{E}(n,l;n',l')$ associated

with the nl, $n'l'$ configuration. The eigenvalues of the corresponding matrix give the corrections of the configuration energy E_c to first order in W; the associated eigenstates are the zero-order eigenstates.

To calculate the matrix which represents W inside $\mathcal{E}(n,l;n',l')$, we can choose any basis, in particular, the basis of kets (5). Actually, it is to our advantage to use a basis well adapted to the symmetries of W. We shall see that we can choose a basis in which the restriction of W is already diagonal.

2-a. Choice of a basis of $\mathcal{E}(n,l;n',l')$ adapted to the symmetries of W

α. Total orbital momentum $\mathbf{L}$ *and total spin* $\mathbf{S}$

W does not commute with the individual orbital angular momenta $\mathbf{L}_1$ and $\mathbf{L}_2$ of each electron. However, we have already shown (*cf.* Chap. X, § A-2) that, if $\mathbf{L}$ denotes the total orbital angular momentum:

$$\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2 \tag{9}$$

we have:

$$[W, \mathbf{L}] = [\frac{e^2}{R_{12}}, \mathbf{L}] = 0 \tag{10}$$

Therefore, $\mathbf{L}$ is a constant of the motion[1]. Moreover, since W does not act in the spin state space, this is also true for the total spin $\mathbf{S}$:

$$[W, \mathbf{S}] = \mathbf{0} \tag{11}$$

Now, consider the set of the four operators, $\mathbf{L}^2$, $\mathbf{S}^2$, L_z, S_z. They commute with each other and with W. We shall show that they constitute a C.S.C.O. in the subspace $\mathcal{E}(n,l;n',l')$ of $\mathcal{E}_A$. This will enable us in § 2-b to find directly the eigenvalues of the restriction of W in this subspace.

To do this, we shall return to the space $\mathcal{E}$, the tensor product of the state spaces $\mathcal{E}(1)$ and $\mathcal{E}(2)$ relative to the two electrons, assumed to be numbered arbitrarily. The subspace $\mathcal{E}(n,l;n',l')$ of $\mathcal{E}_A$ associated with the nl, nl' configuration can be obtained[2] by antisymmetrizing the various kets of the subspace $\mathcal{E}_{n,l}(1) \otimes \mathcal{E}_{n',l'}(2)$ of $\mathcal{E}$. If we choose the basis $|1 : n, l, m, \varepsilon > \otimes |2 : n', l', m', \varepsilon' >$ in this subspace, we obtain the basis of physical kets (5) by antisymmetrization.

However, we know from the results of Chapter X that we can also choose in $\mathcal{E}_{n,l}(1) \otimes \mathcal{E}_{n',l'}(2)$ another basis composed of common eigenvectors of $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, S_z and entirely defined by the specification of the corresponding eigenvalues. We shall write this basis:

$$|1 : n, l; 2 : n', l'; L, M_L > \otimes |S, M_S > \tag{12}$$

with:

$$\begin{cases} L = l + l', l + l' - 1, \ldots, |l - l'| \\ S = 1, 0 \end{cases} \tag{13}$$

[1] This result is related to the fact that, under a rotation involving both electrons, the distance between them, R_{12}, is invariant. However, it changes if only one of the two electrons is rotated. This is why W commutes with neither $\mathbf{L}_1$ nor $\mathbf{L}_2$.

[2] We could also start with the subspace $\mathcal{E}_{n',l'}(1) \otimes \mathcal{E}_{n,l}(2)$ [*cf.* comment (*i*) of § B-2-c of Chapter XIV, p. 1433].

Since $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, S_z are all symmetric operators (they commute with P_{21}), the vectors (12) remain, after antisymmetrization, eigenvectors of $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, S_z with the same eigenvalues (some of them may, of course, have a zero projection onto $\mathcal{E}_A$, in which case the corresponding physical states are excluded by Pauli's principle; see § β below). The non-zero kets obtained by antisymmetrization of (12) are therefore orthogonal, since they correspond to different eigenvalues of at least one of the four observables under consideration. Since they span $\mathcal{E}(n,l;n',l')$, they constitute an orthonormal basis of this subspace, which we shall write:

$$\Big\{|n,l;n',l';L,M_L;S,M_S>\Big\} \tag{14}$$

with:

$$|n,l;n',l';L,M_L;S,M_S> = c(1-P_{21})\Big\{|1:n,l;2:n',l';L,M_L>\otimes|S,M_S>\Big\} \tag{15}$$

where c is a normalization constant. $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, S_z therefore form a C.S.C.O. inside $\mathcal{E}(n,l;n',l')$.

Now, we shall introduce the permutation operator $P_{21}^{(S)}$ in the spin state space:

$$P_{21}^{(S)}|1:\varepsilon;2:\varepsilon'>=|1:\varepsilon';2:\varepsilon> \tag{16}$$

We showed in § B-4 of Chapter X [*cf.* comment (*ii*)] that:

$$P_{21}^{(S)}|S,M_S>=(-1)^{S+1}|S,M_S> \tag{17}$$

Furthermore, if $P_{21}^{(0)}$ is the permutation operator in the state space of the orbital variables, we have:

$$P_{21}=P_{21}^{(0)}\otimes P_{21}^{(S)} \tag{18}$$

Using (17) and (18), we can, finally, put (15) in the form:

$$|n,l;n',l';L,M_L;S,M_S> = c\Big\{[1-(-1)^{S+1}\ P_{21}^{(0)}]|1:n,l;2:n',l';L,M_L>\Big\}\otimes|S,M_S> \tag{19}$$

β. Constraints imposed by the symmetrization postulate

We have seen that the dimension of the space $\mathcal{E}(n,l;n',l')$ is not always equal to $4(2l+1)(2l'+1)$, that is, to the dimension of $\mathcal{E}_{n,l}(1)\otimes\mathcal{E}_{n',l'}(2)$. Certain kets of $\mathcal{E}_{n,l}(1)\otimes\mathcal{E}_{n',l'}(2)$ can therefore have a zero projection onto $\mathcal{E}(n,l;n',l')$. It is interesting to study the consequences for the basis (14) of this constraint imposed by the symmetrization postulate.

First of all, assume that the two electrons do not occupy the same shell. It is then easy to see that the orbital part of (19) is a sum or a difference of two orthogonal kets and, consequently, is never zero[3]. Since the same is true of $|S,M_S>$, we see that all the

[3]The normalization constant c is then equal to $\frac{1}{\sqrt{2}}$

possible values of L and S [*cf.* formula (13)] are allowed. For example, for the $1s$, $2s$ configuration, we can have $S = 0$, $L = 0$ and $S = 1$, $L = 0$; for the $1s$, $2p$ configuration, we can have $S = 0$, $L = 1$ and $S = 1$, $L = 1$, etc.

If we now assume that the two electrons occupy the same shell, we have $n = n'$ and $l = l'$, and certain of the kets (19) can be zero. Let us write $|1 : n, l; 2 : n', l'; L, M_L >$ in the form:

$$|1 : n, l; 2 : n', l'; L, M_L > = \sum_m \sum_{m'} < l, l'; m, m'|L, m_L > |1 : n, l, m; 2 : n', l', m' > \quad (20)$$

According to relation (25) of Complement B_X:

$$< l, l; m, m'|L, M_L >= (-1)^L < l, l; m', m|L, M_L > \quad (21)$$

By using (20), we then get:

$$P_{21}^{(0)}|1 : n, l; 2 : n, l; L, M_L >= (-1)^L|1 : n, l; 2 : n, l; L, M_L > \quad (22)$$

Substituting this result into (19), we obtain[4]:

$$|n, l; n, l; L, M_L; S, M_S >= \begin{cases} 0 & \text{if } L + S \text{ is odd} \\ |1 : n, l; 2 : n, l; L, M_L > \otimes |S, M_S > & \text{if } L + S \text{ is even} \end{cases} \quad (23)$$

Therefore, L and S cannot be arbitrary: $L + S$ must be even. In particular, for the $1s^2$ configuration, we must have $L = 0$, so $S = 1$ is excluded. This is a result found previously.

Finally, note that the symmetrization postulate introduces a close correlation between the symmetry of the orbital part and that of the spin part of the physical ket (19). Since the total ket must be antisymmetric, and the spin part, depending on the value of S, is symmetric ($S = 1$) or antisymmetric ($S = 0$), the orbital part must be antisymmetric when $S = 1$ and symmetric when $S = 0$. We shall see the importance of this point later.

2-b. Spectral terms. Spectroscopic notation

W commutes with the four observables $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, S_z, which form a C.S.C.O. inside $\mathcal{E}(n, l; n', l')$. It follows that the restriction of W inside $\mathcal{E}(n, l; n', l')$ is diagonal in the basis:

$$\left\{ |n, l; n', l'; L, M_L; S, M_S > \right\} \qquad \text{and has eigenvalues of:}$$

$$\delta(L, S) =< n, l; n', l'; L, M_L; S, M_S|W|n, l; n', l'; L, M_L; S, M_S > \quad (24)$$

This energy depends neither on M_L nor on M_S, since relations (10) and (11) imply that W commutes not only with L_z and S_z but also with $L_\pm$ and $S_\pm$: W is therefore a scalar

[4] The normalization constant is then $1/2$.

operator in both the orbital state space and the spin state space (*cf.* Complement B_{VI}, §§ 5-b and 6-c).

Inside each $nl, n'l'$ configuration, we thus obtain energy levels $E_c(n, l; n', l') + \delta(L, S)$, labeled by their values of L and S. Each of them is $(2L + 1)(2S + 1)$-fold degenerate. Such levels are called *spectral terms* and denoted in the following way. With each value of L is associated, in spectroscopic notation (Chap. VII, § C-4-b) a letter of the alphabet; we write the corresponding capital letter and add, at the upper left, a number equal to $2S + 1$. For example, the $1s^2$ configuration leads to a single spectral term, written ^1S (the ^3S, as we have seen, is forbidden by Pauli's principle). The $1s$, $2s$ configuration produces two terms, ^1S (non-degenerate) and ^3S (three-fold degenerate); the $1s$, $2p$ configuration, two terms, ^1P (degeneracy 3) and ^3P (degeneracy 9). For a more complicated configuration such as, for example, $2p^2$, we obtain (*cf.* § 2-a-β) the spectral terms ^1S, ^1D and ^3P ($L + S$ must be even), etc.

Under the effect of the electrostatic repulsion, the degeneracy of each configuration is therefore partially removed (the $1s^2$ configuration, which is non-degenerate, is simply shifted). We shall study this effect in greater detail in the simple example of the $1s, 2s$ configuration. We shall try to understand why the two terms ^1S and ^3S resulting from this configuration, and whose total spin values are different, have different energies although the original Hamiltonian is purely electrostatic.

2-c. Discussion

α. Energies of the spectral terms arising from the 1s, 2s configuration

In the $1s$, $2s$ configuration, $l = l' = L = 0$. It is then easy to obtain from (20):

$$|1 : n = 1, l = 0; 2 : n' = 2, l' = 0; L = M_L = 0 > \\ = |1 : n = 1, l = m = 0; 2 : n' = 2, l' = m' = 0 > \tag{25}$$

a vector that we shall write, more simply, $|1 : 1s; 2 : 2s >$. If $|^3S, M_s >$ and $|^1S, 0 >$ denote the states corresponding to the two spectral terms ^3S and ^1S arising from the $1s$, $2s$ configuration, we obtain, substituting (25) into (19):

$$|^3S, M_S > = \frac{1}{\sqrt{2}}[(1 - P_{21}^{(0)})|1 : 1s; 2 : 2s >] \otimes |S = 1, M_S > \tag{26a}$$

$$|^1S, 0 > \quad = \frac{1}{\sqrt{2}}[(1 + P_{21}^{(0)})|1 : 1s; 2 : 2s >] \otimes |S = 0, M_S = 0 > \tag{26b}$$

Since W does not act on the spin variables, the eigenvalues given by (24) can be written:

$$\delta(^3S) = \frac{1}{2} < 1 : 1s; 2 : 2s|(1 - P_{21}^{(0)})W(1 - P_{21}^{(0)})|1 : 1s; 2 : 2s > \tag{27a}$$

$$\delta(^1S) = \frac{1}{2} < 1 : 1s; 2 : 2s|(1 + P_{21}^{(0)})W(1 + P_{21}^{(0)})|1 : 1s; 2 : 2s > \tag{27b}$$

(we have used the fact that $P_{21}^{(0)}$ is Hermitian). Moreover, $P_{21}^{(0)}$ commutes with W, and the square of $P_{21}^{(0)}$ is the identity operator. Therefore:

$$(1 \pm P_{21}^{(0)})W(1 \pm P_{21}^{(0)}) = (1 \pm P_{21}^{(0)})^2 W = 2(1 \pm P_{21}^{(0)})W \tag{28}$$

Finally, we obtain:

$$\delta(^3S) = K - J \tag{29a}$$

$$\delta(^1S) = K + J \tag{29b}$$

with:

$$K = \ < 1:1s;2:2s|W|1:1s;2:2s > \tag{30}$$

$$J = \ < 1:1s;2:2s|P_{21}^{(0)}W|1:1s;2:2s > = < 1:2s;2:1s|W|1:1s;2:2s > \tag{31}$$

K therefore represents an overall shift of the energy of the two terms and does not contribute to their separation. J is more interesting, as it introduces an energy difference between the ^3S and ^1S terms (*cf.* Fig. 2). We shall therefore study it in a little more detail.

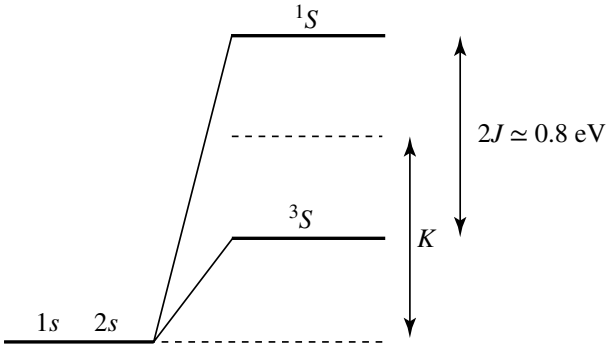


Figure 2: The relative position of the spectral terms ^1S and ^3S arising from the $1s$, $2s$ configuration of the helium atom. K represents an overall shift of the configuration. The removal of the degeneracy is proportional to the exchange integral J.

β. The exchange integral

When we substitute expression (3) for W into (31), there appear terms of the form:

$$< 1:2s;2:1s|V_c(R_1)|1:1s;2:2s > \\ = < 1:2s|V_c(R_1)|1:1s >< 2:1s|2:2s > \tag{32}$$

Now, the scalar product of the two orthogonal states, $|2:1s>$ and $|2:2s>$ is zero. Expression (32) is then equal to zero. The same type of reasoning shows that the terms that arise from the operators $V_c(R_2)$, $-2e^2/R_1$, $-2e^2/R_2$ are also zero, since each of these operators acts only in the single-electron spaces while the state of the two electrons is different in the ket and bra of (31). Finally, there remains:

$$J = < 1:2s;2:1s|\frac{e^2}{|\mathbf{R}_1 - \mathbf{R}_2|}|1:1s;2:2s > \tag{33}$$

J therefore involves only the electrostatic repulsion between the electrons.

Let $\varphi_{n,l,m}(\mathbf{r})$ be the wave functions associated with the states $|n,l,m>$ (the stationary states of an electron in the central potential V_c):

$$\varphi_{n,l,m}(\mathbf{r}) = <\mathbf{r}|n,l,m> \tag{34}$$

In the $\{|\mathbf{r}>\}$ representation, the calculation of J from (33) yields:

$$J = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \; \varphi^*_{2,0,0}(\mathbf{r}_1) \; \varphi^*_{1,0,0}(\mathbf{r}_2) \; \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \; \varphi_{1,0,0}(\mathbf{r}_1) \; \varphi_{2,0,0}(\mathbf{r}_2) \tag{35}$$

This integral is called the "exchange integral". We shall not calculate it explicitly here; we point out, however, that it is positive.

γ. The physical origin of the energy difference between the two spectral terms

We see from expressions (26) and (27) that the origin of the energy separation of the ${}^3S$ and ${}^1S$ terms lies in the symmetry differences of the orbital parts of these terms. As we emphasized at the end of § 2-a, a triplet term ($S = 1$) must have an orbital part which is antisymmetric under exchange of the two electrons ; hence the $-$ sign before $P_{21}^{(0)}$ in (26a) and (27a). On the other hand, a singlet term ($S = 0$) must have a symmetric orbital part [+ sign in (26b) and (27b)].

This explains the relative position of the ${}^3S$ and ${}^1S$ terms shown in Figure 2. For the singlet term, the orbital wave function is symmetric with respect to exchange of the two electrons, which then have a non-zero probability of being at the same point in space. This is why the electrostatic repulsion, which gives an energy of e^2/r_{12} which is large when the electrons are near each other, significantly increases the singlet state energy. On the other hand, for the triplet state, the orbital function is antisymmetric with respect to exchange of the two electrons, which then have a zero probability of being at the same point in space. The average value of the electrostatic repulsion is then smaller. Therefore, the energy difference between the singlet and triplet states arises from the fact that the correlations between the orbital variables of the two electrons depend, because of the symmetrization postulate, on the value of the total spin.

δ. Analysis of the role played by the symmetrization postulate

At this point in the discussion, it might be thought that the degeneracy of a configuration is removed by the symmetrization postulate. We now show[5] that this is not the case. This postulate merely fixes the value of the total spin of the terms arising from a given configuration (because of the inter-electron electrostatic repulsion).

To see this, imagine for a moment that we do not need to apply the symmetrization postulate. Suppose, for example, that the two electrons are replaced by two particles (fictitious, of course) of the same mass, the same charge and the same spin as the electrons but with another intrinsic property that permits us to distinguish between them [without, however, changing the Hamiltonian H of the problem, which is still given by formula (1)]. Since H is not spin-dependent and we do not have to apply the symmetrization postulate, we can ignore the spins completely until the end of the calculations, and then multiply the degeneracies obtained by 4. The energy level of H_0 corresponding to

[5] See also comment (*i*) of § C-4-a-β of Chapter XIV, p. 1442.

the $1s$, $2s$ configuration is two-fold degenerate from the orbital point of view, because two orthogonal states $|1:1s;2:2s>$ and $|1:2s;2:1s>$ correspond to it (they are different physical states since the two particles are of different natures). To study the effect of W, we must diagonalize W in the two-dimensional space spanned by these two kets. The corresponding matrix can be written:

$$\begin{pmatrix} K & J \\ J & K \end{pmatrix} \tag{36}$$

where J and K are given by (30) and (31) [the two diagonal elements of (36) are equal because W is invariant under permutation of the two particles]. Matrix (36) can be diagonalized immediately. The eigenvalues found are $K+J$ and $K-J$, associated respectively with the symmetric and antisymmetric linear combinations of the two kets $|1:1s;2:2s>$ and $|1:2s;2:1s>$. The fact that these orbital eigenstates have well-defined symmetries relative to exchange of the two particles has nothing to do with Pauli's principle. It arises only from the fact that W commutes with $P_{21}^{(0)}$ (common eigenstates of W and $P_{21}^{(0)}$ can therefore be found).

When the two particles are not identical, we obtain the same arrangement of levels and the same orbital symmetry as before. On the other hand, the degeneracy of the levels is obviously different: the lower level, with energy $K-J$, can have a total spin of either $S=0$ or $S=1$, as can the upper level.

If we return to the real helium atom, we now see very clearly the role played by Pauli's principle. It is not responsible for the splitting of the initial level $1s$, $2s$ into the two energy levels $K+J$ and $K-J$, since this splitting would also appear for two particles of different natures. Similarly, the symmetric or antisymmetric character of the orbital part of the eigenvectors is related to the invariance of the electrostatic interaction under permutation of the two electrons. Pauli's principle merely forbids the lower state to have a total spin $S=0$ and the upper state to have a total spin $S=1$, since the corresponding states would be globally symmetric, which is unacceptable for fermions.

ϵ. *The effective spin-dependent Hamiltonian*

We replace W by the operator:

$$\widetilde{W} = \alpha + \beta \mathbf{S}_1 \cdot \mathbf{S}_2 \tag{37}$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ denote the two electron spins. We also have:

$$\widetilde{W} = \alpha - \frac{3\beta\hbar^2}{4} + \frac{\beta}{2}\mathbf{S}^2 \tag{38}$$

so that the eigenstates of $\widetilde{W}$ are the triplet states, with the eigenvalue $\alpha + \beta\hbar^2/4$, and the singlet state, with the eigenvalue $\alpha - 3\beta\hbar^2/4$. Therefore, if we set:

$$\begin{cases} \alpha = K - \dfrac{J}{2} \\ \beta = -\dfrac{2J}{\hbar^2} \end{cases} \tag{39}$$

we obtain, by diagonalizing $\widetilde{W}$, the same eigenstates and eigenvalues we found above[6]. We can then consider that it is as if the perturbation responsible for the appearance of the terms were $\widetilde{W}$ (the "effective" Hamiltonian), which is of the same form as the magnetic interaction between two spins. However, one should not conclude that the coupling energy between the electrons, which is responsible for the appearance of the two terms, is of magnetic origin: two magnetic moments equal to that of the electron and placed at a distance of the order of 1 Å from each other would have an interaction energy much smaller than J. However, because of the very simple form of $\widetilde{W}$, this effective Hamiltonian is often used instead of W.

An analogous situation arises in the study of ferromagnetic materials. In these substances, the electron spins tend to align themselves parallel to each other. Since the spin state is then completely symmetric, Pauli's principle requires the orbital state to be completely antisymmetric. For the same reasons as for the helium atom, the electronic repulsion energy is then minimal. When we study such phenomena, we often use effective Hamiltonians of the same type as (37). However, it must be noted that the physical interaction which is at the origin of the coupling is again electrostatic and not magnetic.

Comments:

(i) The $1s$, $2p$ configuration can be treated in the same way. We then have $L = 1$, so that $M_L = +1$, 0 or -1. As for the $1s$, $2s$ configuration, the shells occupied by the two electrons are different, so that the two terms ^3P and ^1P exist simultaneously. The first one is nine-fold degenerate, and the second, three-fold. It can be shown, as above, that the ^3P term has an energy lower than that of the ^1P term, and the difference between the two energies is proportional to an exchange integral which is analogous to the one written in (35). We would proceed in the same way for all other configurations of the type $1s$, $n'l'$.

(ii) We have treated W like a perturbation of H_0. For this approach to be coherent, the energy shifts associated with W [for example, the exchange integral written in (35)] must be much smaller than the energy differences between configurations. Actually, this is not the case. For the $1s$, $2s$ and $1s$, $2p$ configurations, for example, while the energy difference $\Delta E(^1S - ^3S)$ in the $1s$, $2s$ configuration is of the order of 0.8 eV, the minimum distance between levels is $\Delta E[(1s, 2p)^3P - (1s, 2s)^1S] \simeq 0.35$ eV. We might therefore believe that it is not valid to treat W like a perturbation of H_0.

However, the approach we have given is correct. This is due to the fact that, for all configurations of the type $1s$, $n'l'$, we have $L = l'$. Therefore W, which according to (10) commutes with $\mathbf{L}$, has zero matrix elements between the states of the $1s$, $2s$ configuration and those of the $1s$, $2p$ configuration, since they correspond to different values of L. The operator W couples a $1s$, $n'l'$ configuration only to configurations with distinctly higher energies, of the $1s$, $n''l''$ type with $l'' = l'$ (only the values of n are different) or of the nl, $n''l''$ type, with n and n'' different from 1 (the angular momenta l and l'' can be added to give l').

[6] We must, obviously, keep only the eigenvectors of $\widetilde{W}$ that belong to $\mathcal{E}_A$.

3. Fine-structure levels; multiplets

Thus far, we have taken into account in the Hamiltonian only interactions of purely electrostatic origin; we have neglected all effects of relativistic and magnetic origin. Actually, such effects exist, and we have already studied them in the case of the hydrogen atom (*cf.* Chap. XII, § B-1), where they arise from the variation of the electron mass with the velocity, from the $\mathbf{L} \cdot \mathbf{S}$ spin-orbit coupling, and from the Darwin term. For helium, the situation is more complicated because of the simultaneous presence of two electrons. For example, there is a spin-spin magnetic coupling term in the Hamiltonian (*cf.* Complement B_{XI}) which acts in both the spin state space and the orbital state space of the two electrons[7]. Nevertheless, a great simplification arises from the fact that the energy differences associated with these couplings of relativistic and magnetic origin are much weaker than those which exist between two different spectral terms. This enables us to treat the corresponding Hamiltonian (the fine-structure Hamiltonian) like a perturbation.

The detailed study of the fine structure levels of helium falls outside the domain of this complement. We shall confine ourselves to describing the symmetries of the problem and indicating how to distinguish between the different energy levels. We shall use the fact that the fine-structure Hamiltonian H_{SF} is invariant under a simultaneous rotation of all the orbital and spin variables. This means (*cf.* Complement B_{VI}, § 6) that, if $\mathbf{J}$ denotes the total angular momentum of the electrons:

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{40}$$

we have:

$$[H_{SF}, \mathbf{J}] = \mathbf{0} \tag{41}$$

On the other hand, the fine-structure Hamiltonian changes if the rotation acts only on the orbital variables or only on the spins:

$$[H_{SF}, \mathbf{L}] = -[H_{SF}, \mathbf{S}] \neq \mathbf{0} \tag{42}$$

These properties can easily be seen for the operators $\sum_i \xi(r_i)\mathbf{L}_i \cdot \mathbf{S}_i$, for example, or for the dipole-dipole magnetic interaction Hamiltonian (*cf.* Complement B_{XI}).

The state space associated with a term is spanned by the ensemble of states $|n, l; n', l'; L, M_L; S, M_S >$ written in (19), where L and S are fixed, and where:

$$\begin{cases} -L \leqslant M_L \leqslant +L \\ -S \leqslant M_S \leqslant +S \end{cases} \tag{43}$$

In this subspace, it can be shown that $\mathbf{J}^2$ and J_z form a C.S.C.O. which, according to (41), commutes with H_{SF}. The eigenvectors $|J, M_J >$ common to $\mathbf{J}^2$ [eigenvalue $J(J+1)\hbar^2$] and J_z (eigenvalue $M_J\hbar$) are therefore necessarily eigenvectors of H_{SF}, with an eigenvalue that depends on J but not on M_J (this last property arises from the fact

[7] See for example § 19.6 in Sobel'man (11.12) for an explicit expression of the different terms of the fine structure Hamiltonian (Breit Hamiltonian).

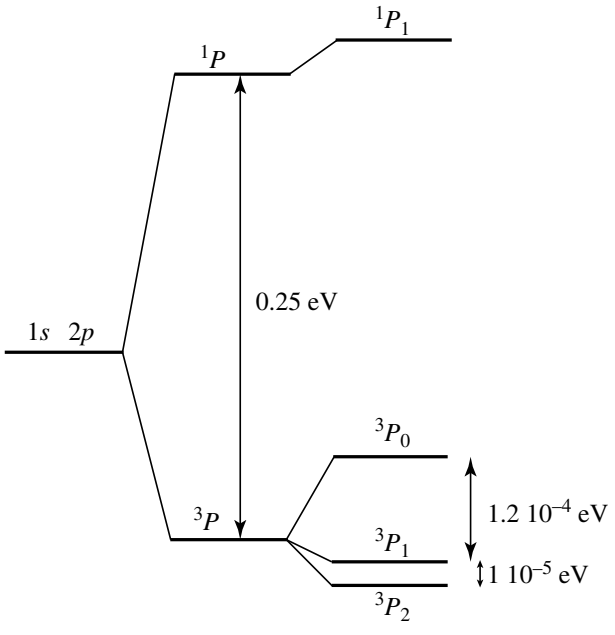


Figure 3: The relative position of the spectral terms and multiplets arising from the $1s$, $2p$ *configuration of the helium atom (the splitting of the three multiplets* $^3P_0$, $^3P_1$, $^3P_2$ *has been greatly exaggerated in order to make the figure clearer).*

---

that H_{SF} commutes with J_+ and J_-). According to the general theory of addition of angular momenta, the possible values of J are:

$$J = L + S, L + S - 1, L + S - 2, \ldots, |L - S| \tag{44}$$

The effect of H_{SF} is therefore a partial removal of the degeneracy. For each "term", there appear as many distinct levels as there are different values of J, according to relation (44). Each of these levels is $(2J + 1)$-fold degenerate and is called a "multiplet". The usual spectroscopic notation consists of denoting a multiplet by adding a right lower index equal to the value of J to the symbol representing the term from which it arises. For example, the ground state of the helium atom gives a single multiplet, $^1S_0$. Similarly, each of the terms ^1S and ^3S of the $1s$, $2s$ configuration leads to a single multiplet: $^1S_0$ and $^3S_1$, respectively. On the other hand, the ^3P term arising from $1s$, $2p$ yields three multiplets, $^3P_2$, $^3P_1$ and $^3P_0$ (*cf.* Fig. 3), and so on. We point out that the measurement and theoretical calculation of the fine structure of the ^3P level of the $1s$, $2p$ configuration is of great fundamental interest, since it can lead to the very precise knowledge of the "fine structure constant", $\alpha = e^2/\hbar c$.

Comments:

(*i*) For many atoms, the fine-structure Hamiltonian is essentially given by:

$$H_{SF} \simeq \sum_{i=1}^{N} \xi(R_i) \mathbf{L}_i . \mathbf{S}_i \tag{45}$$

where $\mathbf{R}_i$, $\mathbf{L}_i$ and $\mathbf{S}_i$ denote the positions, angular momenta and spins of each of the N electrons. It can then be shown, using the Wigner-Eckart theorem (*cf.* Complement D_X), that the energy of the J multiplet is proportional to $J(J+1) - L(L+1) - S(S+1)$. This result is sometimes called the "Landé interval rule".

For helium, the $^3P_1$ and $^3P_2$ levels arising from the $1s$, $2p$ configuration are much closer than would be predicted by this rule. This arises from the importance of the dipole-dipole magnetic coupling of the spins of the two electrons.

(*ii*) In this complement, we have neglected the "hyperfine effects" related to nuclear spin (*cf.* Chap. XII, § B-2). Such effects actually exist only for the ^{3}He isotope, whose nucleus has a spin $I = 1/2$ (the nucleus of the ^{4}He isotope has a zero spin). Each multiplet of electronic angular momentum J splits, in the case of ^{3}He, into two hyperfine levels of total angular momentum $F = J \pm 1/2$, $(2F+1)$-fold degenerate (unless, of course, $J = 0$).

References and suggestions for further reading:

Kuhn (11.1), Chap. III-B; Slater (11.8), Chap. 18; Bethe and Salpeter (11.10).

Multiplet theory and the Pauli principle: Landau and Lifshitz (1.19), §§ 64 and 65; Slater (1.6), Chap. 7 and (11.8), Chap. 13; Kuhn (11.1), Chap. V, § A; Sobel'man (11.12), Chap. 2, § 5.3.

Complement C_{XIV}

Physical properties of an electron gas. Application to solids

1 Free electrons enclosed in a box 1481
1-a Ground state of an electron gas; Fermi energy E_F 1481
1-b Importance of the electrons with energies close to E_F 1484
1-c Periodic boundary conditions 1489
2 Electrons in solids 1491
2-a Allowed bands 1491
2-b Position of the Fermi level and electric conductivity 1492

In Complements A_{XIV} and B_{XIV}, we studied, taking the symmetrization postulate into account, the energy levels of a small number of independent electrons placed in a central potential (the shell model of many-electron atoms). Now, we shall consider systems composed of a much larger number of electrons, and we shall show that Pauli's exclusion principle has an equally spectacular effect on their behavior.

To simplify the discussion, we shall neglect interactions between electrons. Moreover, we shall assume, at first (§ 1), that they are subjected to no external potential other than the one that restricts them to a given volume and which exists only in the immediate vicinity of the boundary (a free-electron gas enclosed in a "box"). We shall introduce the important concept of the *Fermi energy* E_F, which depends only on the number of electrons per unit volume. We shall also show that the physical properties of the electron gas (specific heat, magnetic susceptibility, ...) are essentially determined by the electrons whose energy is close to E_F.

A free-electron model describes the principal properties of certain metals rather well. However, the electrons of a solid are actually subjected to the periodic potential created by the ions of the crystal. We know that the energy levels of each electron are then grouped into allowed energy bands, separated by forbidden bands (*cf.* Complements F_{XI} and O_{III}). We shall show qualitatively in § 2 that the electric conductivity of a solid is essentially determined by the position of the Fermi level of the electron system relative to the allowed energy bands. Depending on this position, the solid is an insulator or a conductor.

1. Free electrons enclosed in a box

1-a. Ground state of an electron gas; Fermi energy E_F

Consider a system of N electrons, whose mutual interactions we shall neglect, and which, furthermore, are subjected to no external potential. These N electrons, however, are enclosed in a box, which, for simplicity, we shall choose to be a cube with edges of length L.

If the electrons cannot pass through the walls of the box, it is because the walls constitute practically infinite potential barriers. Since the potential energy of the electrons is zero inside the box, the problem is reduced to that of the three-dimensional

infinite square well (*cf.* Complements G_{II} and H_I). The stationary states of a particle in such a well are described by the wave functions:

$$\varphi_{n_x,n_y,n_z}(\mathbf{r}) = \left(\frac{2}{L}\right)^{3/2} \sin\left(n_x\frac{\pi x}{L}\right)\sin\left(n_y\frac{\pi y}{L}\right)\sin\left(n_z\frac{\pi z}{L}\right) \tag{1a}$$

$$n_x, n_y, n_z = 1, 2, 3\ldots \tag{1b}$$

[expression (1a) is valid for $0 \leqslant x, y, z \leqslant L$, since the wave function is zero outside this region]. The energy associated with φ_{n_x,n_y,n_z} is equal to:

$$E_{n_x,n_y,n_z} = \frac{\pi^2\hbar^2}{2m_eL^2}(n_x^2 + n_y^2 + n_z^2) \tag{2}$$

Of course, the electron spin must be taken into account: each of the wave functions (1) describes the spatial part of two distinct stationary states which differ by their spin orientation; these two states correspond to the same energy, since the Hamiltonian of the problem is spin-independent.

The set of these stationary states constitutes a discrete basis, enabling us to construct any state of an electron enclosed in this box (that is, whose wave function goes to zero at the walls). Note that, by increasing the dimensions of the box, we can make the interval between two consecutive individual energies as small as we wish, since this interval is inversely proportional to L^2. If L is sufficiently large, therefore, we cannot, in practice, distinguish between the discrete spectrum (2) and a continuous spectrum containing all the positive values of the energy.

The ground state of the system of the N independent electrons can be obtained by antisymmetrizing the tensor product of the N individual states associated with the lowest energies compatible with Pauli's principle. If N is small, it is thus simple to fill the first individual levels (2) and to find the ground state of the system, as well as its degree of degeneracy and the antisymmetrized kets that correspond to it. However, when N is much larger than 1 (in a macroscopic solid, N is of the order of 10^{23}), this method cannot be used in practice, and we must follow a more global reasoning.

We shall begin by evaluating the number $n(E)$ of individual stationary states whose energies are lower than a given value E. To do so, we shall write expression (2) for the possible energies in the form:

$$E_{n_x,n_y,n_z} = \frac{\hbar^2}{2m_e}\mathbf{k}^2_{n_x,n_y,n_z} \tag{3}$$

with:

$$\begin{aligned}(\mathbf{k}_{n_x,n_y,n_z})_x &= n_x\frac{\pi}{L}\\ (\mathbf{k}_{n_x,n_y,n_z})_y &= n_y\frac{\pi}{L}\\ (\mathbf{k}_{n_x,n_y,n_z})_z &= n_z\frac{\pi}{L}\end{aligned} \tag{4}$$

According to (1), a vector $\mathbf{k}_{n_x,n_y,n_z}$ corresponds to each function $\varphi_{n_x,n_y,n_z}(\mathbf{r})$. Conversely, to each of these vectors, there corresponds one and only one function φ_{n_x,n_y,n_z}. The number of states $n(E)$ can then be obtained by multiplying by 2 the number of vectors $\mathbf{k}_{n_x,n_y,n_z}$ whose modulus is smaller than $\sqrt{2m_eE/\hbar^2}$ (the factor 2 arises, of course,

from the existence of electron spin). The tips of the vectors $\mathbf{k}_{n_x,n_y,n_z}$ divide $\mathbf{k}$-space into elementary cubes of edge π/L (see Figure 1, in which, for simplicity, a two-dimensional rather than a three-dimensional space is shown). Each of these tips is common to eight neighboring cubes, and each cube has eight corners. Consequently, if the elementary cubes are sufficiently small (that is, if L is sufficiently large), there can be considered to be one vector $\mathbf{k}_{n_x,n_y,n_z}$ per volume element $(\pi/L)^3$ of $\mathbf{k}$-space.

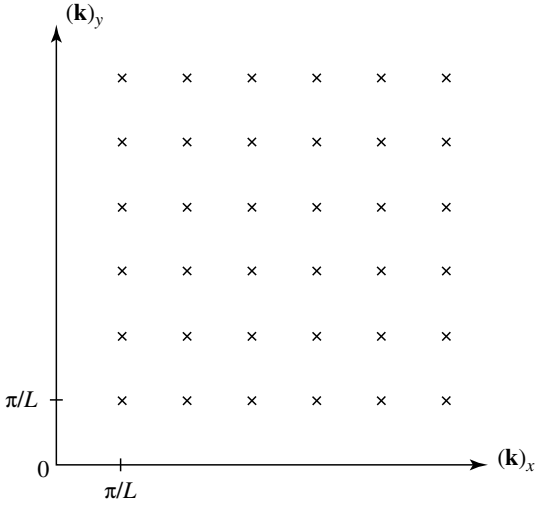


Figure 1: Tips of the vectors $\mathbf{k}_{n_x,n_y}$ characterizing the stationary wave functions in a two-dimensional infinite square well.

The value E of the energy which we have chosen defines, in $\mathbf{k}$-space, a sphere centered at the origin, of radius $\sqrt{2m_eE/\hbar^2}$. Only one-eigth of the volume of this sphere is involved, since the components of $\mathbf{k}$ are positive [*cf.* (1b) and (4)]. If we divide it by the volume element $(\pi/L)^3$ associated with each stationary state, and if we take into account the factor 2 due to the spin, we obtain:

$$n(E) = 2\frac{1}{8}\frac{4}{3}\pi\left(\frac{2m_e}{\hbar^2}E\right)^{3/2}\frac{1}{(\pi/L)^3} = \frac{L^3}{3\pi^2}\left(\frac{2m_e}{\hbar^2}E\right)^{3/2} \tag{5}$$

This result enables us to calculate immediately the maximal individual energy of an electron in the ground state of the system, that is, the *Fermi energy* E_F of the electron gas. This energy E_F satisfies:

$$n(E_F) = N \tag{6}$$

which gives:

$$E_F = \frac{\hbar^2}{2m_e}\left(3\pi^2\frac{N}{L^3}\right)^{2/3} \tag{7}$$

Note that, as might be expected, the Fermi energy depends only on the number N/L^3 of electrons per unit volume. At absolute zero, all the individual states of energy less than E_F are occupied, and all those whose energies are greater than E_F are empty. We shall see in § 1-b what happens at non-zero temperatures.

We can also deduce the *density of states* $\rho(E)$ from (5); by definition, $\rho(E)\mathrm{d}E$ is the number of states whose energies are included between E and $E + \mathrm{d}E$. This density of states, as we shall see later, is of considerable physical importance. It can be obtained simply by differentiating $n(E)$ with respect to E:

$$\rho(E) = \frac{\mathrm{d}n(E)}{\mathrm{d}E} = \frac{L^3}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2} E^{1/2} \tag{8}$$

$\rho(E)$ therefore varies like $\sqrt{E}$. At absolute zero, the number of electrons with a given energy between E and $E + \mathrm{d}E$ (less than E_F, of course) is equal to $\rho(E)\mathrm{d}E$. By using the value (7) of the Fermi energy E_F, we can put $\rho(E)$ in the form:

$$\rho(E) = \frac{3}{2}N\frac{E^{1/2}}{E_F^{3/2}} \tag{9}$$

Comment:

It can be seen from (5) that the dimensions of the box are involved only through the intermediary of the volume element $(\pi/L)^3$ associated, in **k**-space, with each stationary state. If, instead of choosing a cubic box of edge L, we had considered a parallelepiped of edges L_1, L_2, L_3, we would have obtained a volume element of $\pi^3/L_1L_2L_3$: only the volume $L_1L_2L_3$ of the box, therefore, enters into the density of states. This result can be shown to remain valid, whatever the exact form of the box, provided it is sufficiently large.

1-b. Importance of the electrons with energies close to E_F

The results obtained in the preceding section make it possible to understand the physical properties of a free electron gas. We shall give two simple examples here, that of the specific heat and that of the magnetic susceptibility of the system. We shall confine ourselves, however, to semi-quantitative arguments which simply illustrate the fundamental importance of Pauli's exclusion principle.

α. *Specific heat*

At absolute zero, the electron gas is in its ground state: all the individual levels of energy less than E_F are occupied, and all the others are empty. Taking into account the form (8) of the density of states $\rho(E)$, we can represent the situation schematically as in Figure 2a: the number $\nu(E)\ \mathrm{d}E$ of electrons with an energy between E and $E + \mathrm{d}E$ is $\rho(E)\ \mathrm{d}E$ for $E < E_F$ and zero for $E > E_F$. What happens if the temperature T is low but not strictly zero?

If the electrons obeyed classical mechanics, each of them, in going from absolute zero to the temperature T, would gain an energy of the order of kT (where k is the

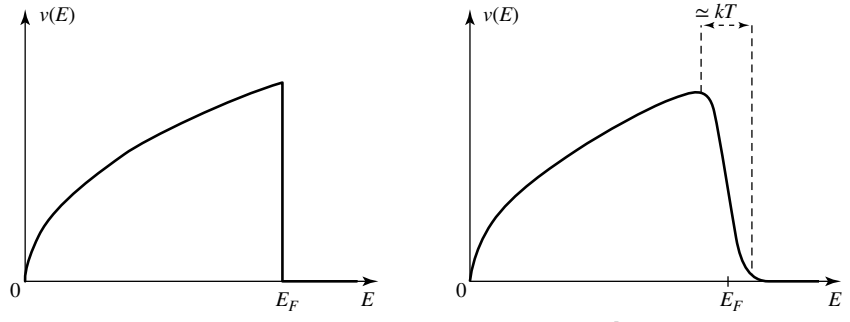


Figure 2: Variation of $\nu(E)$ with respect to E [$\nu(E)\mathrm{d}E$ is the number of electrons with energy between E and $E + \mathrm{d}E$]. At absolute zero, all the levels whose energies are less than the Fermi energy E_F are occupied (fig. a). At a slightly higher temperature T, the transition between empty and occupied levels occurs over an energy interval of a few kT (fig. b).

Boltzmann constant). The total energy per unit volume of the electron gas would then be approximately:

$$U_{cl}(T) \simeq \frac{N}{L^3} kT \tag{10}$$

This would lead to a specific heat at constant volume $\partial U_{cl}/\partial T$ that is independent of the temperature.

In reality, the physical phenomena are totally different, since Pauli's principle prevents most of the electrons from gaining energy. For an electron whose initial energy E is much less than E_F (more precisely, if $E_F - E \gg kT$), the states to which it could go if its energy increased by kT are already occupied and are therefore forbidden to it. Only electrons having an initial energy E close to E_F ($E_F - E \simeq kT$) can "heat up", as shown by Figure 2b. The number of these electrons is approximately:

$$\Delta N \simeq \rho(E_F) kT = \frac{3}{2} N \frac{kT}{E_F} \tag{11}$$

[according to (9)]. Since the energy of each one increases by about kT, the total energy per unit volume can be written:

$$U(T) \simeq \frac{N}{L^3} \frac{kT}{E_F} kT \tag{12}$$

instead of the classical expression (10). Consequently, the constant volume specific heat is proportional to the absolute temperature T:

$$c_V = \frac{\partial U}{\partial T} \simeq \frac{Nk}{L^3} \frac{kT}{E_F} \tag{13}$$

For a metal, to which the free-electron model can be applied, E_F is typically on the order of a few eV. Since kT is about 0.03 eV at ordinary temperatures, we see that in this case the factor kT/E_F introduced by Pauli's principle is of the order of 1/100.

Comments:

(*i*) In order to calculate the specific heat of the electron gas quantitatively, we must know the probability $f(E,T)$ for an individual state of energy E to be occupied when the system is at thermodynamic equilibrium at the temperature T. The number $\nu(E)\ \mathrm{d}E$ of electrons whose energies are included between E and $E + \mathrm{d}E$ is then:

$$\nu(E)\ \mathrm{d}E = f(E,T)\ \rho(E)\ \mathrm{d}E \tag{14}$$

It is shown in statistical mechanics (Complement B_{XV}, § 2-a) that, for fermions, the function $f(E,T)$ can be written:

$$f(E,T) = \frac{1}{\mathrm{e}^{(E-\mu)/kT} + 1} \tag{15}$$

where μ is the *chemical potential*(Appendix VI), also called the *Fermi level* of the system. This is the *Fermi-Dirac distribution*. The Fermi level is determined by the condition that the total number of electrons must be equal to N:

$$\int_0^{+\infty} \frac{\rho(E)\mathrm{d}E}{\mathrm{e}^{(E-\mu)/kT} + 1} = N \tag{16}$$

μ depends on the temperature, but it can be shown that it varies very slowly for small T. The shape of the function $f(E,T)$ is shown in Figure 3. At absolute zero, $f(E,0)$ is equal to 1 for $E < \mu$ and to 0 for $E > \mu$ ("step" function). At non-zero temperatures, $f(E,T)$ has the form of a rounded "step" (the energy interval over which it varies is of the order of a few kT as long as $kT \ll \mu$).

For a free electron gas, it is clear that the Fermi level μ at absolute zero coincides with the Fermi energy E_F calculated in § 1-a. According to (14) and the form that $f(E,T)$ takes for $T = 0$ (Fig. 3), μ then characterizes, like E_F, the highest individual energy.

On the other hand, for a system with a discrete spectrum of energies (E_1, E_2, $\ldots E_i, \ldots$), the Fermi level μ obtained from formula (16) does not coincide with the highest individual energy E_m in the ground state at absolute zero. In this case, the density of states is composed of a series of "delta functions" centered at $E_1, E_2 \ldots E_i, \ldots$ Consequently, at absolute zero, μ can take on any value between E_m and E_{m+1}, since, according to (14), all these possibilities lead to the same value of $\nu(E)$. We choose to define μ at absolute zero as the limit of $\mu(T)$ as T approaches zero. Since at non-zero temperatures the level E_m empties a little, and E_{m+1} begins to fill, the limit of $\mu(T)$ is found to be a value between E_m and E_{m+1} (halfway between these two values if the two states E_m and E_{m+1} have the same degree of degeneracy).

Similarly, for a system containing a series of allowed energy bands separated by forbidden bands (electrons of a solid; *cf.* Complement F_{XI}), the Fermi level μ is in a forbidden band when the highest individual energy at absolute zero coincides with the upper limit of an allowed band. On the other hand, the Fermi level μ is equal to E_F when E_F falls in the middle of an allowed band.

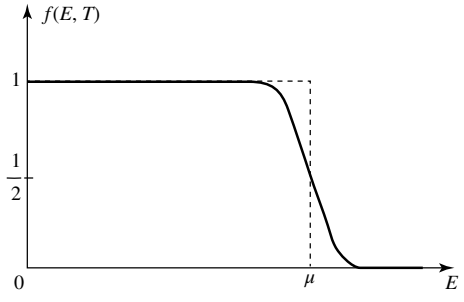


Figure 3: Plot of the Fermi-Dirac distribution at absolute zero (dashed line) and at low temperatures (solid line).
For an electron gas at absolute zero, the Fermi level μ coincides with the Fermi energy E_F. The curves in Figure 2 can be obtained by multiplying the density of states $\rho(E)$ by $f(E,T)$.

(*ii*) The preceding results explain the behavior of the specific heat of metals at very low temperatures. At ordinary temperatures, the specific heat is essentially due to vibrations of the ionic lattice (*cf.* Complement $\mathrm{L_V}$), since that of the electron gas is practically negligible. However, the specific heat of the lattice approaches zero as T^3 for small T. Therefore, that of the electron gas becomes preponderant at low temperatures (around 1 K) where, for metals, a decrease that is linear with respect to T is actually observed.

β. Magnetic susceptibility

Now suppose that a free electron gas is placed in a uniform magnetic field **B** parallel to Oz. The energy of an individual stationary state then depends on the corresponding spin state, since the Hamiltonian contains a paramagnetic spin term (*cf.* Chap. IX, § A-2):

$$W = -2\frac{\mu_B}{\hbar} B\, S_z \tag{17}$$

where μ_B is the Bohr magneton:

$$\mu_B = \frac{q\hbar}{2m_e} \tag{18}$$

and **S** is the electron spin operator. For the sake of simplicity, we shall treat (17) as the only additional term in the Hamiltonian (the behavior of the spatial wave functions was studied in detail in Complement $\mathrm{E_{VI}}$). Under these conditions, the stationary states remain the same as in the absence of a magnetic field, and the corresponding energy is increased or decreased by $\mu_B B$ depending on the spin state. The densities of states $\rho_+(E)$ and $\rho_-(E)$ corresponding respectively to the spin states $|+>$ and $|->$ can therefore be obtained very simply from the density $\rho(E)$ calculated in § 1-a:

$$\rho_\pm(E) = \frac{1}{2}\,\rho(E \pm \mu_B B) \tag{19}$$

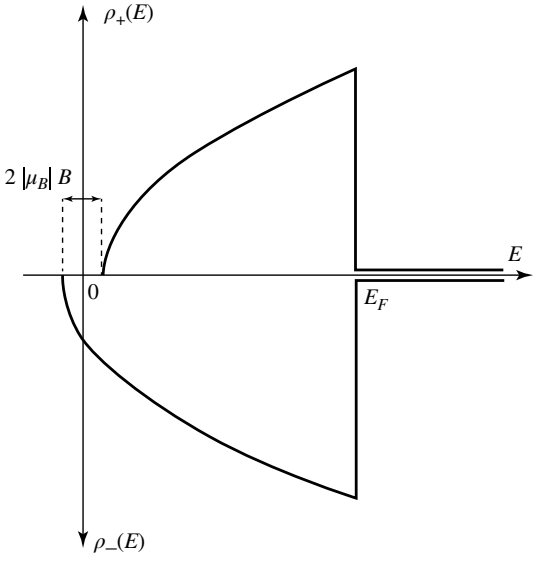


Figure 4: The densities of states $\rho_+(E)$ and $\rho_-(E)$ corresponding respectively to the spin states $|+>$ and $|->$ (μ_B is negative). At absolute zero, only the states whose energies are less than E_F are occupied.

Thus, at absolute zero, we arrive at the situation shown in Figure 4.

Since the magnetic energy $|\mu_B|$ B is much smaller than E_F, the difference between the number of electrons whose spins are antiparallel to the magnetic field and the number whose spins are parallel to **B** is practically, at absolute zero:

$$N_- - N_+ \simeq \frac{1}{2}\rho(E_F)2|\mu_B|B \tag{20}$$

The magnetic moment M per unit volume can therefore be written:

$$\begin{aligned} M &= |\mu_B|\frac{1}{L^3}(N_- - N_+) \\ &= \mu_B^2 B\frac{1}{L^3}\rho(E_F) \end{aligned} \tag{21}$$

This magnetic moment is proportional to the applied field, so that the magnetic susceptibility per unit volume is equal to:

$$\chi = \frac{M}{B} = \mu_B^2\frac{1}{L^3}\rho(E_F) \tag{22}$$

or, using expression (9) for $\rho(E)$:

$$\chi = \frac{3}{2}\frac{N}{L^3}\frac{\mu_B^2}{E_F} \tag{23}$$

Comments:

(*i*) We have assumed the system to be at absolute zero, but result (23) remains valid at low temperatures, since the modifications of the number of occupied states (Fig. 2b) are practically the same for both spin orientations. We therefore find a temperature-independent magnetic susceptibility. This is indeed what is observed for metals.

(*ii*) As in the preceding section, we see that the system behavior in the presence of a magnetic field is essentially determined by the electrons whose energies are close to E_F. This is another manifestation of Pauli's principle. When the magnetic field is applied, the electrons in the $|+>$ spin state tend to go into the $|->$ state, which is energetically more favorable. But most of them are prevented from doing so by the exclusion principle, since all the neighbouring $|->$ states are already occupied.

1-c. Periodic boundary conditions

α. Introduction

The functions φ_{n_x,n_y,n_z} given by formula (1a) have a completely different structure from that of the plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}$ which usually describe the stationary states of free electrons. This difference arises solely from the boundary conditions imposed by the walls of the box, since, inside the box, the plane waves satisfy the same equation as the φ_{n_x,n_y,n_z}:

$$-\frac{\hbar^2}{2m_e}\Delta\varphi(\mathbf{r}) = E\varphi(\mathbf{r}) \tag{24}$$

The functions (1a) are less convenient to handle than plane waves; this is why the latter are preferably used. To do so, we impose on the solutions of equation (24) new, artificial, boundary conditions which do not exclude plane waves. Of course, since these conditions are different from those actually created by the walls of the box, this changes the physical problem. However, we shall show in this section that we can find the most important physical properties of the initial system in this way. For this to be true, it is necessary for the new boundary conditions to lead to a discrete set of possible values of $\mathbf{k}$ such that:

(*i*) The system of plane waves corresponding to these values of $\mathbf{k}$ constitutes a basis on which can be expanded any function whose domain is inside the box.

(*ii*) The density of states $\rho'(E)$ associated with this set of values of $\mathbf{k}$ is identical to the density of states $\rho(E)$ calculated in § 1-a from the true stationary states.

Of course, the fact that the new boundary conditions are different from the real conditions means that the plane waves cannot correctly describe what happens near the walls (surface effects). However, it is clear that they can, because of condition (*ii*), lead to a very simple explanation of the volume effects, which, according to what we have seen in § 1-b, depend only on the density of states $\rho(E)$. Moreover, because of condition (*i*), the motion of any wave packet far from the walls can be correctly described by superposing plane waves, since, between two collisions with the walls, the wave packet propagates freely.

β. The Born-von Karman conditions

We shall no longer require the individual wave functions to go to zero at the walls of the box, but, rather, to be periodic with a period L:

$$\varphi(x+L,y,z)=\varphi(x,y,z) \tag{25}$$

with analogous relations in y and z. Wave functions of the form $e^{i\mathbf{k}\cdot\mathbf{r}}$ satisfy these conditions if the components of the vector $\mathbf{k}$ satisfy:

$$\begin{cases} k_x = n'_x \frac{2\pi}{L} \\ k_y = n'_y \frac{2\pi}{L} \\ k_z = n'_z \frac{2\pi}{L} \end{cases} \tag{26}$$

where, now, n'_x, n'_y *and* n'_z *are positive or negative integers or zero.* We therefore introduce a new system of wave functions:

$$\varphi'_{n'_x,n'_y,n'_z}(\mathbf{r}) = \frac{1}{L^{3/2}}\, e^{i\frac{2\pi}{L}(n'_x x, n'_y y, n'_z z)} \tag{27}$$

which are normalized inside the volume of the box. The corresponding energy, according to (24), can be written:

$$E_{n'_x,n'_y,n'_z} = \frac{\hbar^2}{2m_e}\,\frac{4\pi^2}{L^2}(n_x'^2 + n_y'^2 + n_z'^2) \tag{28}$$

Any wave function defined inside the box can be extended into a periodic function in x, y, z, of period L. Since this periodic function can always be expanded in a Fourier series (*cf.* Appendix I, § 1-b), the $\{\varphi'_{n'_x,n'_y,n'_z}(\mathbf{r})\}$ system constitutes a basis for wave functions with a domain inside the box. To each vector $\mathbf{k}_{n'_x,n'_y,n'_z}$, whose components are given by (26), there corresponds a well-defined value of the energy $E_{n'_x,n'_y,n'_z}$, given by (28). Note, however, that the vectors $\mathbf{k}_{n'_x,n'_y,n'_z}$ can now have positive, negative or zero components, and that their tips divide space into elementary cubes whose edges are twice that found in § 1-a.

In order to show that boundary conditions (25) lead to the same physical results (as far as the volume effects are concerned) as those of § 1-a, it suffices to calculate the number $n'(E)$ of stationary states of energy less than E, and find the value (5) [the Fermi energy E_F and the density of states $\rho(E)$ can be derived directly from $n(E)$]. We evaluate $n'(E)$ in the same way as in § 1-a, taking into account the new characteristics of the vectors $\mathbf{k}_{n'_x,n'_y,n'_z}$. Since the components of $\mathbf{k}$ can now have arbitrary signs, the volume of the sphere of radius $\sqrt{2m_eE/\hbar^2}$ must no longer be divided by 8. However, this modification is compensated by the fact that the volume element $(2\pi/L)^3$ associated with each of the states (27) is eight times larger than the one corresponding to the boundary conditions of § 1-a. Consequently, $n'(E)$ is the same as expression (5) for $n(E)$.

The periodic boundary conditions (25) therefore permit us to meet conditions (*i*) and (*ii*) of the preceding section. They are usually called the Born-Von Karman conditions ("B.V.K. conditions").

Comment:

Consider a truly free electron (not enclosed in a box). The eigenfunctions of the three components of the momentum $\mathbf{P}$ (and, consequently, those of the Hamiltonian $H = \mathbf{P}^2/2m_e$) form a "continuous basis":

$$\left\{ \left(\frac{1}{2\pi\hbar} \right)^{3/2} \mathrm{e}^{i\mathbf{P}\cdot\mathbf{r}/\hbar} \right\} \tag{29}$$

We have already indicated several times that the states for which the form (29) is valid in all space are not physical states, but can be used as mathematical intermediaries in studying the physical states, which are wave packets.

We sometimes prefer to use the discrete basis (27) rather than the continuous basis (29). To do so, we consider the electron to be enclosed in a fictitious box of edge L, much larger than any dimension involved in the problem, and we impose the B.V.K. conditions. Any wave packet, which will always be inside the box for sufficiently large L, can be as well expanded on the discrete basis (27) as on the continuous basis (29). The states (27) can therefore, like the states (29), be considered to be intermediaries of the calculation; however, they present the advantage of being normalized inside the box. We must, of course, check, at the end of the calculations, that the various physical quantities obtained (transition probabilities, cross sections,...) do not depend on L, provided that L is sufficiently large.

Obviously, for a truly free electron, L has no physical meaning and can be arbitrary, as long as it is sufficiently large for the states (27) to form a basis on which the wave packets involved in the problem can be expanded [condition (i) of § 1-c-α]. On the other hand, in the physical problem which we are studying here, L^3 is the volume inside which the N electrons are actually confined and has, consequently, a definite value.

2. Electrons in solids

2-a. Allowed bands

The model of a free electron gas enclosed in a box can be applied rather well to the conduction electrons of a metal. These electrons can be considered to move freely inside the metal, the electrostatic attraction of the crystalline lattice preventing them from escaping when they approach the surface of the metal. However, this model does not explain why some solids are good electrical conductors while others are insulators. This is a remarkable experimental fact: the electric properties of crystals are due to the electrons of the atoms of which they are composed; yet, the intrinsic conductivity can vary by a factor of 10^{30} between a good insulator and a pure metal. We shall see, in a very qualitative way, how this can be explained by Pauli's principle and by the existence of energy bands arising from the periodic nature of the potential created by the ions (*cf.* Complements $\mathrm{O_{III}}$ and $\mathrm{F_{XI}}$).

We showed in Complement $\mathrm{F_{XI}}$ that if, in a first approximation, we consider the electrons of a solid to be independent, their possible individual energies are grouped into *allowed bands*, separated by *forbidden bands*. Assuming that each electron is subjected to the influence of a linear chain of regularly spaced positive ions, we found, in the strong-bond approximation, a series of bands, each one containing $2\mathcal{N}$ levels, where $\mathcal{N}$ is the number of ions (the factor 2 arises from the spin).

The situation, of course, is more complex in a real crystal, in which the positive ions occupy the nodes of a three-dimensional lattice. The theoretical understanding of the properties of a solid requires a detailed study of the energy bands, a study which is based on the spatial characteristics of the crystalline lattice. We shall not treat in detail these specific problems of solid state physics. We shall content ourselves with a qualitative discussion of the phenomena.

2-b. Position of the Fermi level and electric conductivity

Knowing the band structure and the number of states per band, we obtain the ground state of the electron system of a solid by successively "filling" the individual states of the various allowed bands, beginning, of course, with the lowest energies. The electron system is really in the ground state only at absolute zero. However, as we pointed out in § 1-b-α, the characteristics of this ground state permit the semi-quantitative understanding of the behavior of the system at non-zero temperatures – often, up to ordinary temperatures. Like the thermal and magnetic properties (*cf.* § 1-b), the electrical properties of the system are principally determined by the electrons whose individual energies are very close to the highest value E_F. If we place the solid in an electric field, an electron whose initial energy is much lower than E_F cannot gain energy by being accelerated, since the states it would reach in this way are already occupied. It is therefore essential to know the position of E_F relative to the allowed energy bands.

First of all, we shall assume (Fig. 5a) that E_F falls in the middle of an allowed band. The Fermi level μ is then equal to E_F [*cf.* comment (i) of § 1-b-α]. The electrons whose energies are close to E_F can easily be accelerated, in this case, since the slightly higher energy states are empty and accessible. Consequently, *a solid for which the Fermi level falls in the middle of an allowed band is a conductor.* The electrons with the highest energies then behave approximately like free parlicles.

Consider, on the other hand, a solid for which the ground state is composed of entirely occupied allowed bands (Fig. 5b). E_F is then equal to the upper limit of an allowed band, and the Fermi level μ falls inside the adjacent forbidden band [*cf.* comment (i) of § 1-b-α]. In this case, no electrons can be accelerated, since the energy states immediately above theirs are forbidden. Therefore, *a solid for which the Fermi level falls inside a forbidden band is an insulator.* The larger the interval ΔE between the last occupied band and the first empty allowed band, the better the insulator. We shall return to this point later.

The deep allowed bands, completely occupied by electrons and, consequently, inert from an electrical and thermal point of view, are called *valence bands.* They are generally narrow. In a "strong-bond" model (*cf.* Complement F_{XI}, § 2), these bands arise from the atomic levels of lowest energies, which are only slightly affected by the presence of the other atoms in the crystal. On the other hand, the higher bands are wider; a partially occupied band is called a *conduction band.*

For a solid to be a good insulator, the last occupied band must not only be entirely full in the ground state, but also, separated from the immediately higher allowed band by a sufficiently wide forbidden band. As we have indicated (§ 1-b-α), at non-zero temperatures, some states of energy lower than E_F can empty, while some higher energy states fill (Fig. 2b). For the solid to remain an insulator at the temperature T, the width ΔE of the forbidden band, which prevents this excitation of electrons, must be much larger than kT. If ΔE is less than or of the order of kT, a certain number of electrons leave

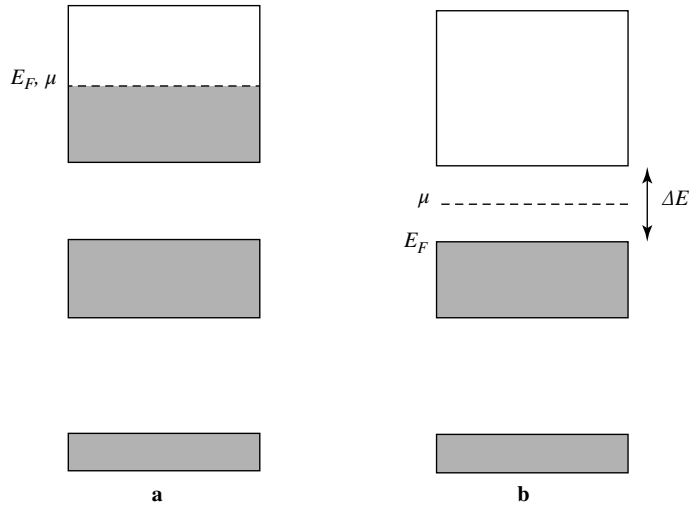


Figure 5: Schematic representation of the individual levels occupied by the electrons at absolute zero (in grey). E_F is the highest individual energy. In a conductor (fig. a), E_F (which then coincides with the Fermi level μ) falls inside an allowed band, called the "conduction band". The electrons whose energies are near E_F can then be accelerated easily, since the slightly higher energy states are accessible to them. In an insulator (fig. b), E_F falls on the upper boundary of an allowed band called the "valence band" (the Fermi level μ is then situated in the adjacent forbidden band). The electrons can be excited only by crossing the forbidden band. This requires an energy at least equal to the width ΔE of this band.

---

the last valence band to occupy states of the immediately higher allowed band (which would be completely empty at absolute zero). The crystal then possesses conduction electrons, but in restricted numbers: it is a *semiconductor* (such a semiconductor is called *intrinsic*; see comment below). For example, diamond, for which ΔE is close to 5 eV, remains an insulator at ordinary temperatures, while silicon and germanium, although quite similar to diamond, are semiconductors: their forbidden bands have a width ΔE less than 1 eV. These considerations, while very qualitative, enable us to understand why the electrical conductivity of a semiconductor increases very rapidly with the temperature; with more quantitative arguments, we indeed find a dependence of the form $\mathrm{e}^{-\Delta E/2kT}$.

The properties of semiconductors also reveal an apparently paradoxical phenomenon. It is as if, in addition to the electrons which have crossed the forbidden band ΔE at a temperature T, there existed in the crystal an equal number of particles with a positive charge. These particles also contribute to the electric current, but their contribution to the Hall effect[1], for example, is opposite in sign to what would be expected for electrons.

[1]Recall what the Hall effect is: in a sample carrying a current and placed in a magnetic field perpendicular to this current, the moving charges are subjected to the Lorentz force. In the steady state, this causes a transverse electric field to appear (perpendicular to the current and to the magnetic

This can be explained very well by band theory, and constitutes a spectacular demonstration of Pauli's principle. To understand this qualitatively, we must recall that the last valence band, when it is completely full in the vicinity of absolute zero, does not conduct any current (Pauli's principle forbids the corresponding electrons from being accelerated). When, by thermal excitation, certain electrons move into the conduction band, they free the states they had occupied in the valence band. These empty states in an almost full band are called "holes". Holes behave like particles of charge opposite to that of the electron. If an electric field is applied to the system, the electrons remaining in the valence band can move, without leaving this band, and occupy the empty states. In this way, they "fill holes" but also "leave new holes behind them". Holes therefore move in the direction opposite to that of the electrons, that is, as if they had a positive charge. This very rough argument can be made more precise, and it can indeed be shown that holes are in every way equivalent to positive charge carriers.

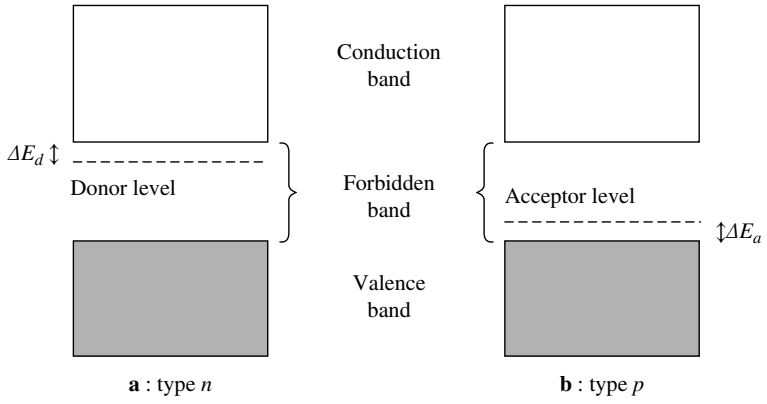


Figure 6: Extrinsic semiconductors: donor atoms (fig. a) bring in electrons which move easily into the conduction band, since their ground states are separated from it only by an energy interval ΔE_d which is much smaller than the width of the forbidden band. Acceptor atoms (fig. b) easily capture valence band electrons, since, for this to happen, these electrons need only an excitation energy ΔE_d which is much smaller than that needed to reach the conduction band. This process creates, in the valence band, holes which can conduct current.

Comment:

We have been speaking only of chemically pure and geometrically perfect crystals. However, in practice, all solids have imperfections and impurities, which often play an important role, particularly in semiconductors. Consider, for example, a quadrivalent silicon or germanium crystal, in which certain atoms are replaced by pentavalent impurity atoms, such as phosphorus, arsenic or antimony (this often happens, without any important

field).

change in the crystal structure). An atom of such an impurity possesses one too many outer electrons relative to the neighboring silicon or germanium atoms: it is called an electron *donor*. The binding energy ΔE_d of the additional electron is considerably lower in the crystal than in the free atom (it is of the order of a few hundredths of an eV); this is due essentially to the large dielectric constant of the crystal, which reduces the Coulomb force (*cf.* Complement A_{VII}, § 1-a-δ). The result is that the excess electrons brought in by the donor atoms move more easily into the conduction band than do the "normal" electrons which occupy the valence band (Fig. 6a). The crystal thus becomes a conductor at a temperature much lower than would pure silicon or germanium. This conductivity due to impurities is called *extrinsic*. Analogously, a trivalent impurity (like boron, aluminium or gallium) behaves in silicon or germanium like an electron *acceptor*: it can easily capture a valence band electron (Fig. 6b), leaving a hole which can conduct the current. In a pure (intrinsic) semiconductor, the number of conduction electrons is always equal to the number of holes in the valence band. An extrinsic semiconductor, on the other hand, can, depending on the relative proportion of donor and acceptor atoms, contain more conduction electrons than holes (it is then said to be of the *n-type*, since the majority of charge carriers are negative), or more holes than conduction electrons (*p-type* semiconductors with a majority of positive charge carriers). These properties serve as the foundation of numerous technological applications (transistors, rectifiers, photoelectric cells, etc.). This is why impurities are often intentionally added to a semiconductor to modify its characteristics: this is called "doping".

References and suggestions for further reading:

See section 8 of the bibliography, especially Kittel (8.2) and Reif (8.4).

For the solid state physics part, see Feynman III (1.2), Chap. 14 and section 13 of the bibliography.

Complement D_{XIV}

Exercises

1. Let h_0 be the Hamiltonian of a particle. Assume that the operator h_0 acts only on the orbital variables and has three equidistant levels of energies 0, $\hbar\omega_0$, $2\hbar\omega_0$ (where ω_0 is a real positive constant) which are non-degenerate in the orbital state space $\mathcal{E}_r$ (in the total state space, the degeneracy of each of these levels is equal to $2s+1$, where s is the spin of the particle). From the point of view of the orbital variables, we are concerned only with the subspace of $\mathcal{E}_r$ spanned by the three corresponding eigenstates of h_0.

a. Consider a system of three independent electrons whose Hamiltonian can be written:

$$H = h_0(1) + h_o(2) + h_0(3)$$

Find the energy levels of H and their degrees of degeneracy.

b. Same question for a system of three identical bosons of spin 0.

2. Consider a system of two identical bosons of spin $s = 1$ placed in the same central potential $V(r)$. What are the spectral terms (*cf.* Complement B_{XIV}, § 2-b) corresponding to the $1s^2$, $1s2p$, $2p^2$ configurations?

3. Consider the state space of an electron, spanned by the two vectors $|\varphi_{p_x}>$ and $|\varphi_{p_y}>$ which represent two atomic orbitals, p_x and p_y, of wave functions $\varphi_{p_x}(\mathbf{r})$ and $\varphi_{p_y}(\mathbf{r})$ (*cf.* Complement E_{VII}, § 2-b):

$$\varphi_{p_x}(\mathbf{r}) = xf(r) = \sin\theta\cos\varphi(r)f(r)$$
$$\varphi_{p_y}(\mathbf{r}) = yf(r) = \sin\theta\cos\varphi(r)f(r)$$

a. Write, in terms of $|\varphi_{p_x}>$ and $|\varphi_{p_y}>$, the state $|\varphi_{p_\alpha}>$ that represents the p_α orbital pointing in the direction of the xOy plane that makes an angle α with Ox.

b. Consider two electrons whose spins are both in the $|+>$ state, the eigenstate of S_z of eigenvalue $+\hbar/2$.

Write the normalized state vector $|\psi>$ which represents the system of these two electrons, one of which is in the state $|\varphi_{p_x}>$ and the other, in the state $|\varphi_{p_y}>$.

c. Same question, with one of the electrons in the state $|\varphi_{p_\alpha}>$ and the other one in the state $|\varphi_{p_\beta}>$, where α and β are two arbitrary angles. Show that the state vector $|\psi>$ obtained is the same.

d. The system is in the state $|\psi>$ of question *b*. Calculate the probability density $\mathcal{P}(r,\ \theta,\ \varphi;\ r',\ \theta',\ \varphi')$ of finding one electron at (r,θ,φ) and the other one at (r',θ',φ'). Show that the electronic density $\rho(r,\theta,\varphi)$ [the probability density of finding any electron at (r,θ,φ)] is symmetric with respect to revolution about the Oz axis. Determine the probability density of having $\varphi - \varphi' = \varphi_0$, where φ_0 is given. Discuss the variation of this probability density with respect to φ_0.

4. Collision between two identical particles

The notation used is that of § D-2-a-β of Chapter XIV.

a. Consider two particles (1) and (2), with the same mass m, assumed for the moment to have no spin and to be distinguishable. These two particles interact through a potential $V(r)$ that depends only on the distance r between them. At the initial time t_0, the system is in the state $|1 : p\mathbf{e}_z; 2 : -p\mathbf{e}_z >$. Let $U(t, t_0)$ be the evolution operator of the system. The probability amplitude of finding it in the state $|1 : p\mathbf{n}; 2 : -p\mathbf{n} >$ at time t_1 is:

$$F(\mathbf{n}) =< 1 : p\mathbf{n}; 2 : -p\mathbf{n}|U(t_1, t_0)|1 : p\mathbf{e}_z; 2 : -p\mathbf{e}_z >$$

Let θ and φ be the polar angles of the unit vector $\mathbf{n}$ in a system of orthonormal axes $Oxyz$. Show that $F(\mathbf{n})$ does not depend on φ. Calculate in terms of $F(\mathbf{n})$ the probability of finding any one of the particles (without specifying which one) with the momentum $p\mathbf{n}$ and the other one with the momentum $-p\mathbf{n}$. What happens to this probability if θ is changed to $\pi - \theta$?

b. Consider the same problem [with the same spin-independent interaction potential $V(r)$], but now with two identical particles, one of which is initially in the state $|p\mathbf{e}_z, m_s >$, and the other, in the state $| - p\mathbf{e}_z, m'_s >$ (the quantum numbers m_s and m'_s refer to the eigenvalues $m_s\hbar$ and $m'_s\hbar$ of the spin component along Oz). Assume that $m_s \neq m'_s$. Express in terms of $F(\mathbf{n})$ the probability of finding, at time t_1, one particle with momentum $p\mathbf{n}$ and spin m_s and the other one with momentum $-p\mathbf{n}$ and spin m'_s. If the spins are not measured, what is the probability of finding one particle with momentum $p\mathbf{n}$ and the other one with momentum $-p\mathbf{n}$? What happens to these probabilities when θ is changed to $\pi - \theta$?

c. Treat problem *b* for the case $m_s = m'_s$. In particular, examine the $\theta = \pi/2$ direction, distinguishing between two possibilites, depending on whether the particles are bosons or fermions. Show that, again, the scattering probability is the same in the θ and $\pi - \theta$ directions.

5. Collision between two identical unpolarized particles

Consider two identical particles, of spin s, which collide. Assume that their initial spin states are not known: each of the two particles has the same probability of being in the $2s+1$ possible orthogonal spin states. Show that, with the notation of the preceding exercise, the probability of observing scattering in the $\mathbf{n}$ direction is:

$$|F(\mathbf{n})|^2 + |F(-\mathbf{n})|^2 + \frac{\varepsilon}{2s+1}[F^*(\mathbf{n})F(-\mathbf{n}) + c.c.]$$

($\varepsilon = +1$ for bosons, -1 for fermions).

6. Possible values of the relative angular momentum of two identical particles

Consider a system of two identical particles interacting by means of a potential that depends only on their relative distance, so that the Hamiltonian of the system can be written:

$$H = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_1^2}{2m} + V(|\mathbf{R}_1 - \mathbf{R}_2|)$$

As in § B of Chapter VII, we set:

$$\mathbf{R}_G = \frac{1}{2}(\mathbf{R}_1 + \mathbf{R}_2) \qquad \mathbf{P}_G = \mathbf{P}_1 + \mathbf{P}_2$$

$$\mathbf{R} = \mathbf{R}_1 - \mathbf{R}_2 \qquad \mathbf{P} = \frac{1}{2}(\mathbf{P}_1 + P_2)$$

H then becomes:

$$H = H_G + H_r$$

with:

$$H_G = \frac{\mathbf{P}_G^2}{4m}$$

$$H_r = \frac{\mathbf{P}^2}{m} + V(R)$$

a. First, we assume that the two particles are identical bosons of zero spin (π mesons, for example).

α. We use the $\{|\mathbf{r}_G, \mathbf{r}>\}$ basis of the state space $\mathcal{E}$ of the system, composed of common eigenvectors of the observables $\mathbf{R}_G$ and $\mathbf{R}$. Show that, if P_{21} is the permutation operator of the two particles:

$$P_{21}|\mathbf{r}_G, \mathbf{r}> = |\mathbf{r}_G, -\mathbf{r}>$$

β. We now go to the $\{|\mathbf{p}_G; E_n, l, m>\}$ basis of common eigenvectors of $\mathbf{P}_G, H_r, \mathbf{L}^2$ and L_z ($\mathbf{L} = \mathbf{R} \times \mathbf{P}$ is the relative angular momentum of the two particles). Show that these new basis vectors are given by expressions of the form:

$$|\mathbf{p}_G; E_n, l, m> = \frac{1}{(2\pi\hbar)^{3/2}} \int \mathrm{d}^3 r_G \ \mathrm{e}^{i\mathbf{p}_G \cdot \mathbf{r}_G/\hbar} \times \int \mathrm{d}^3 r \ R_{n,l}(r) \ Y_l^m(\theta, \varphi) \ |\mathbf{r}_G, \mathbf{r}>$$

Show that:

$$P_{21}|\mathbf{p}_G; E_n, l, m> = (-1)^l |\mathbf{p}_G; E_n, l, m>$$

γ. What values of l are allowed by the symmetrization postulate?

b. The two particles under consideration are now identical fermions of spin 1/2 (electrons or protons).

α. In the state space of the system, we first use the $\{|\mathbf{r_G}, \mathbf{r}; \mathbf{S}, \mathbf{M}>\}$ basis of common eigenstates of $\mathbf{R}_G, \mathbf{R}, \mathbf{S}^2$ and S_Z, where $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$ is the total spin of the system (the kets $|S, M>$ of the spin state space were determined in § B of Chapter X). Show that:

$$P_{21}|\mathbf{r}_G, \mathbf{r}; S, M> = (-1)^{S+1} |\mathbf{r}_G, -\mathbf{r}; S, M>$$

β. We now go to the $\{|\mathbf{p}_G; E_n, l, m; S, M>\}$ basis of common eigenstates of $\mathbf{P}_G$, H_r, $\mathbf{L}^2$, L_z, $\mathbf{S}^2$ and S_z.

As in question a-β, show that:

$$P_{21}|\mathbf{p}_G; E_n, l, m; S, M >= (-1)^{S+1}(-1)^l|\mathbf{p}_G; E_n, l, m; S, M >$$

γ. Derive the values of l allowed by the symmetrization postulate for each of the values of S (triplet and singlet).

c. (more difficult)

Recall that the total scattering cross section in the center of mass system of two distinguishable particles interacting through the potential $V(r)$ can be written:

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l$$

where the δ_l are the phase shifts associated with $V(r)$ [*cf.* Chap. VIII, formula (C-58)].

α. What happens if the measurement device is equally sensitive to both particles (the two particles have the same mass)?

β. Show that, in the case envisaged in question a, the expression for σ becomes:

$$\sigma = \frac{16\pi}{k^2}\sum_{l\ \text{even}}(2l+1)\sin^2\delta_l$$

γ. For two unpolarized identical fermions of spin 1/2 (the case of question b), prove that:

$$\sigma = \frac{4\pi}{k^2}\left\{\sum_{l\ \text{even}}(2l+1)\sin^2\delta_l + 3\sum_{l\ \text{odd}}(2l+1)\sin^2\delta_l\right\}$$

7. Position probability densities for a system of two identical particles

Let $|\varphi >$ and $|\chi >$ be two normalized orthogonal states belonging to the orbital state space $\mathcal{E}_\mathbf{r}$ of an electron, and let $|+ >$ and $|- >$ be the two eigenvectors, in the spin state space $\mathcal{E}_s$, of the S_z component of its spin.

a. Consider a system of two electrons, one in the state $|\varphi, + >$ and the other, in the state $|\chi, - >$. Let $\rho_{II}(\mathbf{r}, \mathbf{r}')\ \mathrm{d}^3r\ \mathrm{d}^3r'$ be the probability of finding one of them in a volume $\mathrm{d}^3\mathrm{r}$ centered at point $\mathbf{r}$, and the other in a volume $\mathrm{d}^3\mathrm{r}'$ centered at $\mathbf{r}'$ (two-particle density function). Similarly, let $\rho_I(\mathbf{r})\ \mathrm{d}^3\mathrm{r}$ be the probability of finding one of the electrons in a volume $\mathrm{d}^3\mathrm{r}$ centered at point $\mathbf{r}$ (one-particule density function). Show that:

$$\rho_{II}(\mathbf{r}, \mathbf{r}') = |\varphi(\mathbf{r})|^2|\chi(\mathbf{r}')|^2 + |\varphi(\mathbf{r}')|^2|\chi(\mathbf{r})|^2$$
$$\rho_I(\mathbf{r}) = |\varphi(\mathbf{r})|^2 + |\chi(\mathbf{r})|^2$$

Show that these expressions remain valid even if $|\varphi >$ and $|\chi >$ are not orthogonal in $\mathcal{E}_\mathbf{r}$.

Calculate the integrals over all space of $\rho_I(\mathbf{r})$ and $\rho_{II}(\mathbf{r}, \mathbf{r}')$. Are they equal to 1?

Compare these results with those which would be obtained for a system of two distinguishable particles (both spin 1/2), one in the state $|\varphi, + >$ and the other in the

state $|\chi, - >$; the device which measures their positions is assumed to be unable to distinguish between the two particles.

b. Now assume that one electron is in the state $|\varphi, + >$ and the other one, in the state $|\chi, + >$. Show that we then have:

$$\rho_{II}(\mathbf{r}, \mathbf{r}') = |\varphi(\mathbf{r})\chi(\mathbf{r}') - \varphi(\mathbf{r}')\chi(\mathbf{r})|^2$$
$$\rho_I(\mathbf{r}) = |\varphi(\mathbf{r})|^2 + |\chi(\mathbf{r})|^2$$

Calculate the integrals over all space of $\rho_I(\mathbf{r})$ and $\rho_{II}(\mathbf{r}, \mathbf{r}')$.
What happens to ρ_I and ρ_{II} if $|\varphi >$ and $|\chi >$ are no longer orthogonal in $\mathcal{E}_r$?

c. Same questions for two identical bosons, either in the same spin state or in two orthogonal spin states.

8. The aim of this exercise is to demonstrate the following point: once the state vector of a system of N identical bosons (or fermions) has been suitably symmetrized (or antisymmetrized), it is not indispensable, in order to calculate the probability of any measurement result, to perform another symmetrization (or antisymmetrization) of the kets associated with the measurement. More precisely, provided that the state vector belongs to $\mathcal{E}_S$ (or $\mathcal{E}_A$), the physical predictions can be calculated as if we were confronted with a system of distinguishable particles studied by imperfect measurement devices unable to distinguish between them.

Let $|\chi >$ be the state vector of a system of N identical bosons (all of the following reasoning is equally valid for fermions). We have:

$$S|\psi >= |\psi > \tag{1}$$

I.

a. Let $|\chi >$ be the normalized physical ket associated with a measurement in which the N bosons are found to be in the different and orthonormal individual states $|u_\alpha >$, $|u_\beta >, \ldots, |u_\nu >$. Show that:

$$|\chi >= \sqrt{N!}S|1 : u_\alpha; 2 : u_\beta; \ldots; N : u_\nu > \tag{2}$$

b. Show that, because of the symmetry properties of $|\psi >$:

$$| < 1 : u_\alpha; 2 : u_\beta; \ldots; N : u_\nu|\psi > |^2 = | < i : u_\alpha; j : u_\beta; \ldots; l : u_\nu|\psi > |^2$$

where $i, j, \ldots, l$ is an arbitrary permutation of the numbers 1, 2, ..., N.

c. Show that the probability of finding the system in the state $|\chi >$ can be written:

$$| < \chi|\psi > |^2 = N!| < 1 : u_\alpha; 2 : u_\beta; \ldots; N : u_\nu|\psi > |^2$$
$$= \sum_{i,j,\ldots,l} | < i : u_\alpha; j : u_\beta; \ldots; l : u_\nu|\psi > |^2 \tag{3}$$

where the summation is performed over all permutations of the numbers 1, 2, ..., N.

d. Now assume that the particles are distinguishable, and that their state is described by the ket $|\psi>$. What would be the probability of finding any one of them in the state $|u_\alpha>$, another one in the state $|u_\beta>$, ..., and the last one in the state $|u_\nu>$?

Conclude, by comparison with the results of *c*, that, for identical particles, it is sufficient to apply the symmetrization postulate to the state vector $|\psi>$ of the system.

e. How would the preceding argument be modified if several of the individual states constituting the state $|\chi>$ were identical? (For the sake of simplicity, consider only the case where $N = 3$).

II. *(more difficult)*

Now, consider the general case, in which the measurement result being considered is not necessarily defined by the specification of individual states, since the measurement may no longer be complete. According to the postulates of Chapter XIV, we must proceed in the following way in order to calculate the corresponding probability:

- first of all, we treat the particles as distinguishable, and we number them: their state space is then $\mathcal{E}$. Then let $\mathcal{E}_m$ be the subspace of $\mathcal{E}$ associated with the measurement result envisaged and the measurement being performed with devices incapable of distinguishing between the particles;
- with $|\psi_m>$ denoting an arbitrary ket of $\mathcal{E}_m$, we construct the set of kets $S\,|\psi_m>$ which constitutes a vector space $\mathcal{E}_m^S$ ($\mathcal{E}_m^S$ is the projection of $\mathcal{E}_m$ onto $\mathcal{E}_S$); if the dimension of $\mathcal{E}_m^S$ is greater than 1, the measurement is not complete;
- the desired probability is then equal to the square of the norm of the orthogonal projection onto $\mathcal{E}_m^S$ of the ket $|\psi>$ describing the state of the N identical particles.

a. If P_α is an arbitrary permutation operator of the N particles, show that, by construction of $\mathcal{E}_m$:

$$P_\alpha|\psi_m> \in \mathcal{E}_m$$

Show that $\mathcal{E}_m$ is globally invariant under the action of S and that $\mathcal{E}_m^S$ is simply the intersection of $\mathcal{E}_S$ and $\mathcal{E}_m$.

b. We construct an orthonormal basis in $\mathcal{E}_m$:

$$\{|\varphi_m^1>, |\varphi_m^2>, \ldots, |\varphi_m^k>, |\varphi_m^{k+1}>, \ldots, |\varphi_m^p>\}$$

the first k vectors of which constitute a basis of $\mathcal{E}_m^S$. Show that the kets $S|\varphi_m^n>$, where $k+1 \leqslant n \leqslant p$, must be linear combinations of the first k vectors of this basis. Show, by taking their scalar products with the bras $<\varphi_m^1|, <\varphi_m^2|, \ldots, <\varphi_m^k|$, that these kets $S|\varphi_m^n>$ (with $n \geqslant k+1$) are necessarily zero.

c. Show from the preceding results that the symmetric nature of $|\psi>$ implies that:

$$\sum_{n=1}^{p} |<\varphi_m^n|\psi>|^2 = \sum_{n=1}^{k} |<\varphi_m^n|\psi>|^2$$

that is:

$$<\psi|P_m^S|\psi> = <\psi|P_m|\psi>$$

where P_m^S and P_m denote respectively the projectors onto $\mathcal{E}_m^S$ and $\mathcal{E}_m$.

Conclusion: The probabilities of the measurement results can be calculated from the projection of the ket $|\psi>$ (belonging to $\mathcal{E}_S$) onto an eigensubspace $\mathcal{E}_m$ whose kets do not all belong to $\mathcal{E}_S$, but in which all the particles play equivalent roles.

9. One- and two-particle density functions in an electron gas at absolute zero

I.

a. Consider a system of N particles $1, 2, \ldots, i, \ldots, N$ with the same spin s. First of all, assume that they are not identical. In the state space $\mathcal{E}(i)$ of particle (i), the ket $|i : \mathbf{r}_0, m>$ represents a state in which particle (i) is localized at the point $\mathbf{r}_0$ in the spin state $|m>$ ($m\hbar$: the eigenvalue of S_z).

Consider the operator:

$$F_m(\mathbf{r}_0) = \sum_{i=1}^{N} \left\{ |i : \mathbf{r}_0, m><i : \mathbf{r}_0, m| \otimes \prod_{j \neq i} I(j) \right\}$$

where $I(j)$ is the identity operator in the space $\mathcal{E}(j)$.

Let $|\psi>$ be the state of the N-particle system. Show that $<\psi|F_m(\mathbf{r}_0)|\psi> \mathrm{d}\tau$ represents the probability of finding any one of the particles in the infinitesimal volume element $\mathrm{d}\tau$ centered at $\mathbf{r}_0$, the component of its spin being equal to $m\hbar$.

b. Consider the operator:

$$G_{mm'}(\mathbf{r}_0, \mathbf{r}_0') = \sum_{i=1}^{N} \sum_{j \neq i} \left\{ |i : \mathbf{r}_0, m; j : \mathbf{r}_0', m'><i : \mathbf{r}_0, m; j : \mathbf{r}_0', m'| \otimes \prod_{k \neq i,j} I(k) \right\}$$

What is the physical meaning of the quantity $<\psi|G_{mm'}(\mathbf{r}_0, \mathbf{r}_0')|\psi> \mathrm{d}\tau\, \mathrm{d}\tau'$, where $\mathrm{d}\tau$ and $\mathrm{d}\tau'$ are infinitesimal volumes?

The average values $<\psi|F_m(\mathbf{r}_0)|\psi>$ and $<\psi|G_{mm'}(\mathbf{r}_0, \mathbf{r}_0')|\psi>$ will be written, respectively, $\rho_m^I(\mathbf{r}_0)$ and $\rho_{mm'}^{II}(\mathbf{r}_0, \mathbf{r}_0')$ and will be called the one- and two-particle density functions of the N-particle system.

The preceding expressions remain valid when the particles are identical, provided that $|\psi>$ is the suitably symmetrized or antisymmetrized state vector of the system (*cf.* preceding exercise).

II.

Consider a system of N particles in the normalized and orthogonal individual states $|u_1>, |u_2>, \ldots, |u_N>$. The normalized state vector of the system is:

$$|\psi> = \sqrt{N!}\, T\, |1:u_1; 2:u_2; \ldots; N:u_N>$$

where T is the symmetrizer for bosons and the antisymmetrizer for fermions. In this part, we want to calculate the average values in the state $|\psi>$ of symmetric one-particle operators of the type:

$$F = \sum_{i=1}^{N} \left\{ f(i) \otimes \prod_{j \neq i} I(j) \right\}$$

or of symmetric two-particle operators of the type:

$$G = \sum_{i=1}^{N} \sum_{j \neq i} \left\{ g(i,j) \otimes \sum_{k \neq i,j} I(k) \right\}$$

a. Show that:

$$<\psi|F|\psi> = <1:u_1; 2:u_2; \ldots; N:u_N| \left[\sum_{\alpha} \varepsilon_\alpha P_\alpha \right] F |1:u_1; 2:u_2; \ldots; N:u_N>$$

where $\varepsilon_\alpha = +1$ for bosons, and $+1$ or -1 for fermions, depending on whether the permutation P_α is even or odd.

Show that the same expression is valid for the operator G.

b. Derive the relations:

$$<\psi|F|\psi> = \sum_{i=1}^{N} <i:u_i|f(i)|i:u_i>$$

$$<\psi|G|\psi> = \sum_{i=1}^{N} \sum_{j \neq i} \{ <i:u_i; j:u_j|g(i,j)|i:u_i; j:u_j> + \varepsilon <i:u_j; j:u_i|g(i,j)|i:u_i; j:u_j> \}$$

with $\varepsilon = +1$ for bosons, $\varepsilon = -1$ for fermions.

III.

We now want to apply the results of part II to the operators $F_m(\mathbf{r}_0)$ and $G_{mm'}(\mathbf{r}_0, \mathbf{r}'_0)$ introduced in part I. The physical system under study is a gas of N free electrons enclosed in a cubic box of edge L at absolute zero (Complement C_{XIV}, § 1). By applying periodic boundary conditions, we obtain individual states of the form $|\varphi_{\mathbf{k}}> |\pm>$, where the wave function associated with $|\varphi_{\mathbf{k}}>$ is a plane wave $\frac{1}{L^{3/2}} e^{i\mathbf{k}.\mathbf{r}}$, and the components of $\mathbf{k}$ satisfy relations (26) of Complement C_{XIV}. We shall call $E_F = \hbar^2 k_F^2 / 2m$ the Fermi energy of the system and $\lambda_F = 2\pi / k_F$, the Fermi wavelength.

a. Show that the two one-particle density functions $\rho_+^I(\mathbf{r}_0)$ and $\rho_-^I(\mathbf{r}_0)$ are both equal to:

$$\rho_+^I(\mathbf{r}_0) = \rho_-^I(\mathbf{r}_0) = \sum_k |\varphi_{\mathbf{k}}(\mathbf{r}_0)|^2$$

where the summation over $\mathbf{k}$ is performed over all values of $\mathbf{k}$ of modulus less then k_F, satisfying the periodic boundary conditions. By using § 1 of Complement C_{XIV}, show that $\rho_+^I(\mathbf{r}_0) = \rho_-^I(\mathbf{r}_0) = k_F^3/6\pi^2 = N/2L^3$. Could this result have been predicted simply?

b. Show that the two two-particle density functions $\rho_{+-}^{II}(\mathbf{r}_0, \mathbf{r}_0')$ and $\rho_{-+}^{II}(\mathbf{r}_0, \mathbf{r}_0')$ are both equal to:

$$\sum_{\mathbf{k}} \sum_{\mathbf{k}'} |\varphi_{\mathbf{k}}(\mathbf{r}_0)\varphi_{\mathbf{k}'}(\mathbf{r}_0')|^2 = \frac{N^2}{4L^6}$$

where the summations over $\mathbf{k}$ and $\mathbf{k}'$ are defined as above. Give a physical interpretation.

c. Finally, consider the two two-particle density functions $\rho_{++}^{II}(\mathbf{r}_0, \mathbf{r}_0')$ and $\rho_{--}^{II}(\mathbf{r}_0, \mathbf{r}_0')$. Prove that they are both equal to:

$$\sum_{\mathbf{k}} \sum_{k' \neq \mathbf{k}} \{|\varphi_{\mathbf{k}}(\mathbf{r}_0)\varphi_{\mathbf{k}'}(\mathbf{r}_0')|^2 - \varphi_{\mathbf{k}}^*(\mathbf{r}_0')\varphi_{\mathbf{k}'}^*(\mathbf{r}_0)\varphi_{\mathbf{k}}(\mathbf{r}_0)\varphi_{\mathbf{k}'}(\mathbf{r}_0')\}$$

Show that the restriction $\mathbf{k}' \neq \mathbf{k}$ can be omitted, and show that the two two-particle density functions are equal to:

$$\frac{N^2}{4L^6} - \left| \sum_{\mathbf{k}} \varphi_{\mathbf{k}}^*(\mathbf{r}_0)\varphi_{\mathbf{k}}(\mathbf{r}_0') \right|^2 = \frac{N^2}{4L^6}[1 - C^2(k_F d)]$$

with $d = |\mathbf{r}_0 - \mathbf{r}_0'|$, where the function $C(x)$ is defined by:

$$C(x) = \frac{3}{x^3}[\sin x - x \cos x]$$

($\sum_k$ can be replaced by an integral over $\mathbf{k}$).

How do the two-particle density functions $\rho_{++}^{II}(\mathbf{r}_0, \mathbf{r}_0')$ and $\rho_{--}^{II}(\mathbf{r}_0, \mathbf{r}_0')$ vary with respect to the distance d between $\mathbf{r}_0$ and $\mathbf{r}_0'$? Show that it is practically impossible to find two electrons with the same spin separated by a distance much smaller than λ_F.

Appendix I

Fourier series and Fourier transforms

1 Fourier series 1505
1-a Periodic functions 1505
1-b Expansion of a periodic function in a Fourier series 1506
1-c The Bessel-Parseval relation 1507
2 Fourier transforms 1508
2-a Definitions 1508
2-b Simple properties 1509
2-c The Parseval-Plancherel formula 1511
2-d Examples 1511
2-e Fourier transforms in three-dimensional space 1512

In this appendix, we shall review a certain number of definitions, formulas and properties which are useful in quantum mechanics. We do not intend to enter into the details of the derivations, nor shall we give rigorous proofs of the mathematical theorems.

1. Fourier series

1-a. Periodic functions

A function $f(x)$ of a variable is said to be *periodic* if there exists a real non-zero number L such that, for all x:

$$\boxed{f(x+L) = f(x)} \tag{1}$$

L is called the *period* of the function $f(x)$.

If $f(x)$ is periodic with a period of L, all numbers nL, where n is a positive or negative integer, are also periods of $f(x)$. The *fundamental period* L_0 of such a function is defined as being its smallest positive period (the term "period" is often used in physics to denote what is actually the fundamental period of a function).

Comment:

We can take a function $f(x)$ defined only on a finite interval $[a, b]$ of the real axis and construct a function $f_p(x)$ which is equal to $f(x)$ inside $[a, b]$ and is periodic, with a period $(b - a)$. The function $f_p(x)$ is continuous if $f(x)$ is and if:

$$f(b) = f(a) \tag{2}$$

We know that the *trigonometric functions* are periodic. In particular:

$$\cos 2\pi \frac{x}{L} \quad \text{and} \quad \sin 2\pi \frac{x}{L} \tag{3}$$

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

have fundamental periods equal to L.

Other particularly important examples of periodic functions are the *periodic exponentials*. For an exponential $e^{\alpha x}$ to have a period of L, it is necessary and sufficient, according to definition (1), that:

$$e^{\alpha L} = 1 \tag{4}$$

that is:

$$\alpha L = 2in\pi \tag{5}$$

where n is an integer. There are therefore two exponentials of fundamental period L:

$$e^{\pm 2i\pi x/L} \tag{6}$$

which are, furthermore, related to the trigonometric functions (3) which have the same period:

$$e^{\pm 2i\pi x/L} = \cos 2\pi \frac{x}{L} \pm i \sin 2\pi \frac{x}{L} \tag{7}$$

The exponential $e^{2in\pi x/L}$ also has a period of L, but its fundamental period is L/n.

1-b. Expansion of a periodic function in a Fourier series

Let $f(x)$ be a periodic function with a fundamental period of L. If it satisfies certain mathematical conditions (as is practically always the case in physics), it can be expanded in a series of imaginary exponentials or trigonometric functions.

α. Series of imaginary exponentials

We can write $f(x)$ in the form:

$$\boxed{f(x) = \sum_{n=-\infty}^{+\infty} c_n \, e^{ik_n x}} \tag{8}$$

with:

$$k_n = n\frac{2\pi}{L} \tag{9}$$

The coefficients c_n of the Fourier series (8) are given by the formula:

$$\boxed{c_n = \frac{1}{L}\int_{x_0}^{x_0+L} \mathrm{d}x \, e^{-ik_n x} f(x)} \tag{10}$$

where x_0 is an arbitrary real number.

To prove (10), we multiply (8) by $e^{-ik_p x}$ and integrate between x_0 and $x_0 + L$:

$$\int_{x_0}^{x_0+L} \mathrm{d}x \, e^{-ik_p x} f(x) = \sum_{n=-\infty}^{+\infty} c_n \int_{x_0}^{x_0+L} \mathrm{d}x \, e^{i(k_n - k_p)x} \tag{11}$$

The integral of the right-hand side is zero for $n \neq p$ and equal to L for $n = p$. Hence formula (10). It can easily be shown that the value obtained for c_n is independent of the number x_0 chosen.

The set of values $|c_n|$ is called the *Fourier spectrum* of $f(x)$. Note that $f(x)$ is real if and only if:

$$c_{-n} = c_n^* \tag{12}$$

β. Cosine and sine series

If, in the series (8), we group the terms corresponding to opposite values of n, we obtain:

$$f(x) = c_0 + \sum_{n=1}^{\infty} \left(c_n \mathrm{e}^{ik_n x} + c_{-n}\, \mathrm{e}^{-ik_n x} \right) \tag{13}$$

that is, according to (7):

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos k_n x + b_n \sin k_n x) \tag{14}$$

with:

$$\left.\begin{aligned} a_0 &= c_0 \\ a_n &= c_n + c_{-n} \\ b_n &= i(c_n - c_{-n}) \end{aligned}\right\} n > 0 \tag{15}$$

The formulas giving the coefficients a_n and b_n can therefore be derived from (10):

$$\begin{aligned} a_0 &= \frac{1}{L} \int_{x_0}^{x_0+L} \mathrm{d}x\, f(x) \\ a_n &= \frac{2}{L} \int_{x_0}^{x_0+L} \mathrm{d}x\, f(x)\, \cos k_n x \\ b_n &= \frac{2}{L} \int_{x_0}^{x_0+L} \mathrm{d}x\, f(x)\, \sin k_n x \end{aligned} \tag{16}$$

If $f(x)$ has a definite parity, expansion (14) is particularly convenient, since:

$$\begin{aligned} b_n &= 0 \quad \text{if} \quad f(x) \text{ is even} \\ a_n &= 0 \quad \text{if} \quad f(x) \text{ is odd} \end{aligned} \tag{17}$$

Moreover, if $f(x)$ is real, the coefficients a_n and b_n are real.

1-c. The Bessel-Parseval relation

It can easily be shown from the Fourier series (8) that:

$$\boxed{\frac{1}{L} \int_{x_0}^{x_0+L} \mathrm{d}x\, |f(x)|^2 = \sum_{n=-\infty}^{+\infty} |c_n|^2} \tag{18}$$

This can be shown using equation (8):

$$\frac{1}{L}\int_{x_0}^{x_0+L} \mathrm{d}x\ |f(x)|^2 = \sum_{n,p} c_p^*\ c_n \frac{1}{L}\int_{x_0}^{x_0+L} \mathrm{d}x\ \mathrm{e}^{i(k_n-k_p)x} \tag{19}$$

As in (11), the integral of the right-hand side is equal to $L\,\delta_{np}$. This proves (18).

When expansion (14) is used, the Bessel-Parseval relation (18) can also be written:

$$\frac{1}{L}\int_{x_0}^{x_0+L} \mathrm{d}x\ |f(x)|^2 = |a_0|^2 + \frac{1}{2}\sum_{n=1}^{\infty}\left[|a_n|^2 + |b_n|^2\right] \tag{20}$$

If we have two functions, $f(x)$ and $g(x)$, with the same period L, whose Fourier coefficients are, respectively. c_n and d_n, we can generalize relation (18) to the form:

$$\frac{1}{L}\int_{x_0}^{x_0+L} \mathrm{d}x\ g^*(x)\ f(x) = \sum_{n=-\infty}^{+\infty} d_n^*\ c_n \tag{21}$$

2. Fourier transforms

2-a. Definitions

α. The Fourier integral as the limit of a Fourier series

Now, consider a function $f(x)$ which is not necessarily periodic. We define $f_L(x)$ to be the periodic function of period L which is equal to $f(x)$ inside the interval $[-L/2,\ L/2]$. The function $f_L(x)$ can be expanded in a Fourier series:

$$f_L(x) = \sum_{n=-\infty}^{+\infty} c_n\,\mathrm{e}^{ik_n x} \tag{22}$$

where k_n is defined by formula (9), and:

$$c_n = \frac{1}{L}\int_{x_0}^{x_0+L} \mathrm{d}x\,\mathrm{e}^{-ik_n x} f_L(x) = \frac{1}{L}\int_{-\frac{L}{2}}^{+\frac{L}{2}} \mathrm{d}x\,\mathrm{e}^{-ik_n x} f(x) \tag{23}$$

When L approaches infinity, $f_L(x)$ becomes the same as $f(x)$. We shall therefore let L approach infinity in the expressions above.

Definition (9) of k_n then yields:

$$k_{n+1} - k_n = \frac{2\pi}{L} \tag{24}$$

We shall now replace $1/L$ by its expression in terms of $(k_{n+1}-k_n)$ in (23), and substitute this value of c_n into the series (22):

$$f_L(x) = \sum_{n=-\infty}^{+\infty} \frac{k_{n+1}-k_n}{2\pi}\,\mathrm{e}^{ik_n x}\int_{-\frac{L}{2}}^{+\frac{L}{2}} \mathrm{d}\xi\,\mathrm{e}^{-ik_n\xi} f(\xi) \tag{25}$$

When $L \longrightarrow \infty$, $k_{n+1} - k_n$ approaches zero [*cf.* (24)], so that the sum over n is transformed into a definite integral; $f_L(x)$ approaches $f(x)$. The integral appearing in (25) becomes a function of the continuous variable k. If we set:

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}x \, \mathrm{e}^{-ikx} f(x) \tag{26}$$

relation (25) can be written in the limit of infinite L:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}k \, \mathrm{e}^{ikx} \tilde{f}(k) \tag{27}$$

$f(x)$ and $\tilde{f}(k)$ are called *Fourier transforms* of each other.

β. Fourier transforms in quantum mechanics

In quantum mechanics, we actually use a slightly different convention. If $\psi(x)$ is a (one-dimensional) wave function, its Fourier transform $\overline{\psi}(p)$ is defined by:

$$\boxed{\overline{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}x \, \mathrm{e}^{-ipx} \, \psi(x)} \tag{28}$$

and the inverse formula is:

$$\boxed{\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}p \, \mathrm{e}^{ipx/\hbar} \, \overline{\psi}(p)} \tag{29}$$

To go from (26) and (27) to (28) and (29), we set:

$$p = \hbar k \tag{30}$$

(p has the dimensions of a momentum if x is a length), and:

$$\overline{\psi}(p) = \frac{1}{\sqrt{\hbar}} \tilde{\psi}(k) = \frac{1}{\sqrt{\hbar}} \tilde{\psi}\left(\frac{p}{\hbar}\right) \tag{31}$$

In this appendix, as is usual in quantum mechanics, we shall use definition (28) of the Fourier transform instead of the traditional definition, (26). To return to the latter definition, furthermore, all we need to do is replace $\hbar$ by 1 and p by k in all the following expressions.

2-b. Simple properties

We shall state (28) and (29) in the condensed notation:

$$\overline{\psi}(p) = \mathcal{F}[\psi(x)] \tag{32a}$$

$$\psi(x) = \overline{\mathcal{F}}[\overline{\psi}(p)] \tag{32b}$$

The following properties can easily be demonstrated:

(i)

$$\begin{aligned} &\overline{\psi}(p - p_0) = \mathcal{F}[\mathrm{e}^{ip_0 x/\hbar} \, \psi(x)] \\ &\mathrm{e}^{-ipx_0/\hbar} \overline{\psi}(p) = \mathcal{F}[\psi(x - x_0)] \end{aligned} \tag{33}$$

This follows directly from definition (28).

$$(ii) \qquad \overline{\psi}(p) = \mathcal{F}[\psi(x)] \Longrightarrow \mathcal{F}[\psi(cx)] = \frac{1}{|c|}\overline{\psi}\left(\frac{p}{c}\right) \tag{34}$$

To see this, all we need to do is change the integration variable:

$$u = cx \tag{35}$$

In particular:

$$\mathcal{F}[\psi(-x)] = \overline{\psi}(-p) \tag{36}$$

Therefore, if the function $\psi(x)$ has a definite parity, its Fourier transform has the same parity.

$$(iii) \qquad \psi(x) \text{ real} \qquad \leftrightarrow [\overline{\psi}(p)]^* = \overline{\psi}(-p) \tag{37a}$$

$$\psi(x) \text{ pure imaginary} \leftrightarrow [\overline{\psi}(p)]^* = -\overline{\psi}(-p) \tag{37b}$$

The same expressions are valid if the functions ψ and $\overline{\psi}$ are inverted.

(iv) If $f^{(n)}$ denotes the nth derivative of the function f, successive differentiations inside the summation yield, according to (28) and (29):

$$\mathcal{F}[\psi^{(n)}(x)] = \left(\frac{ip}{\hbar}\right)^n \overline{\psi}(p) \tag{38a}$$

$$\overline{\psi}^{(n)}(p) = \mathcal{F}\left[\left(-\frac{ix}{\hbar}\right)^n \psi(x)\right] \tag{38b}$$

(v) The *convolution* of two functions $\psi_1(x)$ and $\psi_2(x)$ is, by definition, the function $\psi(x)$ equal to:

$$\psi(x) = \int_{-\infty}^{+\infty} \mathrm{d}y \; \psi_1(y)\, \psi_2(x-y) \tag{39}$$

Its Fourier transform is proportional to the ordinary product of the transforms of $\psi_1(x)$ and $\psi_2(x)$:

$$\overline{\psi}(p) = \sqrt{2\pi\hbar}\, \overline{\psi}_1(p)\, \overline{\psi}_2(p) \tag{40}$$

This can be shown as follows.
We take the Fourier transform of expression (39):

$$\overline{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}x\, \mathrm{e}^{-ipx/\hbar} \int_{-\infty}^{+\infty} \mathrm{d}y \; \psi_1(y)\, \psi_2(x-y) \tag{41}$$

and perform the change of integration variables:

$$\{x, y\} \Longrightarrow \{u = x - y, y\} \tag{42}$$

If we multiply and divide by $\mathrm{e}^{ipy/\hbar}$ we obtain:

$$\overline{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}y\, \mathrm{e}^{-ipy/\hbar} \psi_1(y) \int_{-\infty}^{+\infty} \mathrm{d}u\, \mathrm{e}^{-ipu/\hbar} \psi_2(u) \tag{43}$$

which proves (40).

(*vi*) When $\psi(x)$ is a peaked function of width Δx, the width Δp of $\overline{\psi}(p)$ satisfies:

$$\Delta x \cdot \Delta p \gtrsim \hbar \tag{44}$$

(see § C-2 of Chapter I, where this inequality is analyzed, and Complement C_{III}).

2-c. The Parseval-Plancherel formula

A function and its Fourier transform have the same norm:

$$\boxed{\int_{-\infty}^{+\infty} \mathrm{d}x\, |\psi(x)|^2 = \int_{-\infty}^{+\infty} \mathrm{d}p\, |\overline{\psi}(p)|^2} \tag{45}$$

To prove this, all we need to do is use (28) and (29) in the following way:

$$\begin{aligned}\int_{-\infty}^{+\infty} \mathrm{d}x\, |\psi(x)|^2 &= \int_{-\infty}^{+\infty} \mathrm{d}x\, \psi^*(x) \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}p\, \mathrm{e}^{ipx/\hbar} \overline{\psi}(p) \\ &= \int_{-\infty}^{+\infty} \mathrm{d}p\, \overline{\psi}(p) \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}x\, \mathrm{e}^{ipx/\hbar} \psi^*(x) \\ &= \int_{-\infty}^{+\infty} \mathrm{d}p\, \overline{\psi}^*(p)\, \overline{\psi}(p)\end{aligned} \tag{46}$$

As in § 1-c, the Parseval-Plancherel formula can be generalized:

$$\boxed{\int_{-\infty}^{+\infty} \mathrm{d}x\, \varphi^*(x)\, \psi(x) = \int_{-\infty}^{+\infty} \mathrm{d}p\, \overline{\varphi}^*(p)\, \overline{\psi}(p)} \tag{47}$$

2-d. Examples

We shall confine ourselves to three examples of Fourier transforms, for which the calculations are straightforward.

(*i*) Square function

$$\left.\begin{aligned}\psi(x) &= \frac{1}{a} \quad \text{for} \quad -\frac{a}{2} < x < \frac{a}{2} \\ &= 0 \quad \text{for} \quad |x| > \frac{a}{2}\end{aligned}\right\} \Longleftrightarrow \overline{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \frac{\sin(pa/2\hbar)}{pa/2\hbar} \tag{48}$$

(*ii*) Decreasing exponential

$$\psi(x) = \mathrm{e}^{-|x|/a} \Longleftrightarrow \overline{\psi}(p) = \sqrt{\frac{2}{\pi\hbar}} \frac{1/a}{(p^2/\hbar^2) + (1/a^2)} \tag{49}$$

(*iii*) Gaussian function

$$\psi(x) = \mathrm{e}^{-x^2/a^2} \Longleftrightarrow \overline{\psi}(p) = \frac{a}{\sqrt{2\hbar}} \mathrm{e}^{-p^2a^2/4\hbar^2} \tag{50}$$

(note the remarkable fact that the Gaussian form is conserved by the Fourier transform).

Comment:

In each of these three cases, the widths Δx and Δp can be defined for $\psi(x)$ and $\overline{\psi}(p)$ respectively, and they verify inequality (44).

2-e. Fourier transforms in three-dimensional space

For wave functions $\psi(\mathbf{r})$ which depend on the three spatial variables x, y, z, (28) and (29) are replaced by:

$$\overline{\psi}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int \mathrm{d}^3 r \; \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} \; \psi(\mathbf{r}) \tag{51a}$$

$$\psi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int \mathrm{d}^3 p \; \mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \; \overline{\psi}(\mathbf{p}) \tag{51b}$$

The properties stated above (§§ 2-b and 2-c) can easily be generalized to three dimensions.

If ψ depends only on the modulus r of the radius-vector $\mathbf{r}$, $\overline{\psi}$ depends only on the modulus p of the momentum $\mathbf{p}$ and can be calculated from the expression:

$$\overline{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \frac{2}{p} \int_0^\infty \mathrm{r}\,\mathrm{d}r \; \sin\frac{pr}{\hbar} \; \psi(r) \tag{52}$$

Proof:

First, we shall find using (51a) the value of $\overline{\psi}$ for a vector $\mathbf{p}'$ obtained from $\mathbf{p}$ by an arbitrary rotation $\mathcal{R}$:

$$\mathbf{p}' = \mathcal{R}\mathbf{p} \tag{53}$$

$$\overline{\psi}(\mathbf{p}') = \frac{1}{(2\pi\hbar)^{3/2}} \int \mathrm{d}^3 r \, \mathrm{e}^{-i\mathbf{p}'\cdot\mathbf{r}/\hbar} \psi(r) \tag{54}$$

In this integral, we replace the variable $\mathbf{r}$ by $\mathbf{r}'$ and set:

$$\mathbf{r}' = \mathcal{R}\mathbf{r} \tag{55}$$

Since the volume element is conserved under rotation, we have:

$$\mathrm{d}^3 r' = \mathrm{d}^3 r \tag{56}$$

In addition, the function ψ is unchanged, since the modulus of $\mathbf{r}'$ remains equal to r; finally:

$$\mathbf{p}' \cdot \mathbf{r}' = \mathbf{p} \cdot \mathbf{r} \tag{57}$$

since the scalar product is rotation-invariant. We thus find:

$$\overline{\psi}(\mathbf{p}') = \overline{\psi}(\mathbf{p}) \tag{58}$$

that is, $\overline{\psi}$ depends only on the modulus of $\mathbf{p}$ and not on its direction.

We can then choose $\mathbf{p}$ along Oz to evaluate $\overline{\psi}(p)$:

$$\begin{aligned}\overline{\psi}(p) &= \frac{1}{(2\pi\hbar)^{3/2}} \int \mathrm{d}^3 r\, \mathrm{e}^{-ipz/\hbar}\, \psi(r) \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int_0^\infty r^2\, \mathrm{d}r\, \psi(r) \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\theta\, \sin\theta\, \mathrm{e}^{-ipr\cos\theta/\hbar} \\ &= \frac{1}{(2\pi\hbar)^{3/2}} \int_0^\infty r^2\, \mathrm{d}r\, \psi(r)\, 2\pi \frac{2\hbar}{pr} \sin\frac{pr}{\hbar} \\ &= \frac{1}{\sqrt{2\pi\hbar}} \frac{2}{p} \int_0^\infty r\, \mathrm{d}r\, \psi(r) \sin\frac{pr}{\hbar} \end{aligned} \tag{59}$$

This proves (52).

Assume for instance that $\Psi(r)$ is given by the following (non normalized) function:

$$\Psi(r) = \frac{\mathrm{e}^{-\varepsilon r}}{r} \tag{60}$$

where ε is positive. Relation (52) then becomes:

$$\begin{aligned}\overline{\psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \frac{1}{ip} \int_0^\infty \mathrm{d}r\, \mathrm{e}^{-\varepsilon r} \big[\mathrm{e}^{ipr/\hbar} - \mathrm{e}^{-ipr/\hbar}\big] = \frac{1}{\sqrt{2\pi\hbar}} \frac{1}{ip} \big[\frac{-1}{-\varepsilon + ip/\hbar} - \frac{-1}{-\varepsilon - ip/\hbar}\big] \\ &= \sqrt{\frac{2}{\pi\hbar^3}}\ \frac{1}{\varepsilon^2 + p^2/\hbar^2} \end{aligned} \tag{61}$$

A central potential that varies with r as the right-hand side of (60) is called a "Yukawa potential". When $\varepsilon = 0$, it becomes a Coulomb potential, whose gradient gives an electric field. If we take the gradient of (51b), we obtain the Fourier transformation correspondence between two following vector functions (we now use variable $\boldsymbol{k}$ instead of $\mathbf{p}$, and therefore set $\hbar = 1$):

$$\big[-\varepsilon - \frac{1}{r}\big] \frac{\mathrm{e}^{-\varepsilon r}}{r} \frac{\mathbf{r}}{r} \quad \underset{\mathrm{FT}}{\leftrightarrow} \quad \sqrt{\frac{2}{\pi}} \frac{i\boldsymbol{k}}{\varepsilon^2 + k^2} \tag{62}$$

The limit $\varepsilon \to 0$ then provides:

$$-\frac{\mathbf{r}}{r^3} \ \underset{\mathrm{FT}}{\leftrightarrow}\ \sqrt{\frac{2}{\pi}} \frac{i\boldsymbol{k}}{k^2} \tag{63}$$

References and suggestions for further reading:

See, for example, Arfken (10.4), Chaps. 14 and 15, or Butkov (10.8), Chaps. 4 and 7; Bass (10.1), vol. I, Chaps. XVIII through XX: section 10 of the bibliography, especially the subsection "Fourier transforms; distributions".

Appendix II

The Dirac δ-"function"

1 Introduction; principal properties 1515
1-a Introduction of the δ-"function" 1515
1-b Functions that approach δ 1517
1-c Properties of δ 1518
2 The δ-"function" and the Fourier transform 1520
2-a The Fourier transform of δ 1520
2-b Applications 1521
3 Integral and derivatives of the δ-"function" 1521
3-a δ is the derivative of the "unit step-function" 1521
3-b Derivatives of δ 1523
4 The δ-"function" in three-dimensional space 1524

The δ-"function" is actually a distribution. However, like most physicists, we shall treat it like an ordinary function. This approach, although not mathematically rigorous, is sufficient for quantum mechanical applications.

1. Introduction; principal properties

1-a. Introduction of the δ-"function"

Consider the function $\delta^{(\varepsilon)}(x)$ given by (*cf.* Fig. 1):

$$\begin{aligned} \delta^{(\varepsilon)}(x) &= \frac{1}{\varepsilon} \quad \text{for} \quad -\frac{\varepsilon}{2} < x < \frac{\varepsilon}{2} \\ &= 0 \quad \text{for} \quad |x| > \frac{\varepsilon}{2} \end{aligned} \tag{1}$$

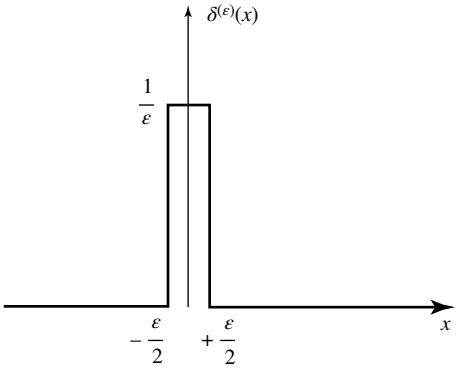


Figure 1: The function $\delta^{(\varepsilon)}(x)$: a square function of width ε and height $1/\varepsilon$, centered at $x = 0$.

where ε is a positive number. We shall evaluate the integral:

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, \delta^{(\varepsilon)}(x)\, f(x) \tag{2}$$

where $f(x)$ is an arbitrary function, well-defined for $x = 0$. If ε is sufficiently small, the variation of $f(x)$ over the effective integration interval $[-\varepsilon/2, \varepsilon/2]$ is negligible, and $f(x)$ remains practically equal to $f(0)$. Therefore:

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, \delta^{(\varepsilon)}(x)\, f(x) \simeq f(0) \int_{-\infty}^{+\infty} \mathrm{d}x\, \delta^{(\varepsilon)}(x) = f(0) \tag{3}$$

The smaller ε, the better the approximation. We therefore examine the limit $\varepsilon = 0$ and define the δ-"function" by the relation:

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, \delta(x)\, f(x) = f(0) \tag{4}$$

which is valid for any function $f(x)$ defined at the origin. More generally, $\delta(x - x_0)$ is defined by:

$$\boxed{\int_{-\infty}^{+\infty} \mathrm{d}x\, \delta(x - x_0)\, f(x) = f(x_0)} \tag{5}$$

Comments:

(i) Actually, the integral notation in (5) is not mathematically justified. δ is defined rigorously not as a function but as a distribution. Physically, this distinction is not an essential one as it becomes impossible to distinguish between $\delta^{(\varepsilon)}(x)$ and $\delta(x)$ as soon as ε becomes negligible compared to all the distances involved in a given physical problem[1]: any function $f(x)$ which we might have to consider does not vary significantly over an interval of length ε. Whenever a mathematical difficulty might arise, all we need to do is assume that $\delta(x)$ is actually $\delta^{(\varepsilon)}(x)$ [or an analogous but more regular function, for example, one of those given in (7), (8), (9), (10), (11)], with ε extremely small but not strictly zero.

(ii) For arbitrary integration limits a and b, we have:

$$\begin{aligned} \int_a^b \mathrm{d}x\, \delta(x)\, f(x) &= f(0) \quad \text{if} \quad 0 \in [a, b] \\ &= 0 \quad \text{if} \quad 0 \notin [a, b] \end{aligned} \tag{6}$$

[1]The accuracy of present-day physical measurements does not, in any case, allow us to investigate phenomena on a scale of less than a fraction of a Fermi (1 Fermi = 10^{15} m).

1-b. Functions that approach δ

It can easily be shown that, in addition to $\delta^{(\varepsilon)}(x)$ defined by (1), the following functions approach $\delta(x)$, that is, satisfy (5), when the parameter ε approaches zero from the positive side:

$$(i) \qquad \frac{1}{2\varepsilon}\mathrm{e}^{-|x|/\varepsilon} \tag{7}$$

$$(ii) \qquad \frac{1}{\pi}\frac{\varepsilon}{x^2+\varepsilon^2} \tag{8}$$

$$(iii) \qquad \frac{1}{\varepsilon\sqrt{\pi}}\,\mathrm{e}^{-x^2/\varepsilon^2} \tag{9}$$

$$(iv) \qquad \frac{1}{\pi}\frac{\sin(x/\varepsilon)}{x} \tag{10}$$

$$(v) \qquad \frac{\varepsilon}{\pi}\frac{\sin^2(x/\varepsilon)}{x^2} \tag{11}$$

We shall also mention an identity which is often useful in quantum mechanics (particularly in collision theory):

$$\operatorname*{Lim}_{\varepsilon\to 0_+}\frac{1}{x\pm i\varepsilon} = \mathcal{P}\frac{1}{x} \mp i\pi\,\delta(x) \tag{12}$$

where $\mathcal{P}$ denotes the Cauchy principal part, defined by[2] [$f(x)$ is a regular function at $x=0$]:

$$\mathcal{P}\int_{-A}^{+B}\frac{\mathrm{d}x}{x}f(x) = \operatorname*{Lim}_{\eta\to 0_+}\left[\int_{-A}^{-\eta}+\int_{+\eta}^{+B}\right]\frac{\mathrm{d}x}{x}f(x); \quad A,B>0 \tag{13}$$

To prove (12), we separate the real and imaginary parts of $1/(x\pm i\varepsilon)$:

$$\frac{1}{x\pm i\varepsilon} = \frac{x\mp i\varepsilon}{x^2+\varepsilon^2} \tag{14}$$

Since the imaginary part is proportional to the function (8), we have:

$$\operatorname*{Lim}_{\varepsilon\to 0_+}\mp i\frac{\varepsilon}{x^2+\varepsilon^2} = \mp i\pi\,\delta(x) \tag{15}$$

As for the real part, we shall multiply it by a function $f(x)$ that is regular at the origin, and integrate over x:

$$\operatorname*{Lim}_{\varepsilon\to 0_+}\int_{-\infty}^{+\infty}\frac{x\,\mathrm{d}x}{x^2+\varepsilon^2}f(x) = \operatorname*{Lim}_{\varepsilon\to 0_+}\operatorname*{Lim}_{\eta\to 0_+}\left[\int_{-\infty}^{-\eta}+\int_{-\eta}^{+\eta}+\int_{+\eta}^{+\infty}\right]\frac{x\,\mathrm{d}x}{x^2+\varepsilon^2}f(x) \tag{16}$$

[2] One often uses one of the following relations:

$$\mathcal{P}\int_{-A}^{+B}\frac{\mathrm{d}x}{x}f(x) = \int_{-B}^{+B}\mathrm{d}x\frac{f_-(x)}{x} + \int_{-A}^{-B}\frac{\mathrm{d}x}{x}f(x) = \int_{-A}^{+B}\mathrm{d}x\frac{f(x)-f(0)}{x} + f(0)\,\operatorname{Log}\frac{B}{A}$$

where $f_-(x) = [f(x)-f(-x)]/2$ is the odd part of $f(x)$. These formulas allow us to explicitly eliminate the divergence at the origin.

The second integral is zero:

$$\lim_{\eta\to 0_+}\int_{-\eta}^{+\eta}\frac{x\,\mathrm{d}x}{x^2+\varepsilon^2}f(x)=f(0)\lim_{\eta\to 0_+}\frac{1}{2}\left[\mathrm{Log}(x^2+\varepsilon^2)\right]_{-\eta}^{+\eta}=0 \tag{17}$$

If we now reverse the order of the evaluation of the limits in (16), the $\varepsilon\longrightarrow 0$ limit presents no difficulties in the two other integrals. Thus:

$$\lim_{\varepsilon\to 0_+}\int_{-\infty}^{+\infty}\frac{x\,\mathrm{d}x}{x^2+\varepsilon^2}f(x)=\lim_{\eta\to 0_+}\left[\int_{-\infty}^{-\eta}+\int_{+\eta}^{+\infty}\right]\frac{\mathrm{d}x}{x}f(x) \tag{18}$$

This establishes identity (12).

1-c. Properties of δ

The properties we shall now state can be demonstrated using (5). Multiplying both sides of the equations below by a function $f(x)$ and integrating, we see that the results obtained are indeed equal.

$$(i)\qquad \delta(-x)=\delta(x) \tag{19}$$

$$(ii)\qquad \delta(cx)=\frac{1}{|c|}\delta(x) \tag{20}$$

and, more generally:

$$\delta[g(x)]=\sum_j\frac{1}{|g'(x_j)|}\delta(x-x_j) \tag{21}$$

where $g'(x)$ is the derivative of $g(x)$ and the x_j are the simple zeros of the function $g(x)$:

$$\begin{aligned} g(x_j)&=0\\ g'(x_j)&\neq 0\end{aligned} \tag{22}$$

The summation is performed over all the simple zeros of $g(x)$. If $g(x)$ has zeros of multiple order [that is, for which $g'(x_j)$ is zero], the expression $\delta[g(x)]$ makes no sense.

$$(iii)\qquad x\,\delta(x-x_0)=x_0\,\delta(x-x_0) \tag{23}$$

and, in particular:

$$x\,\delta(x)=0 \tag{24}$$

The converse is also true and it can be shown that the equation:

$$x\,u(x)=0 \tag{25}$$

has the general solution:

$$u(x)=c\,\delta(x) \tag{26}$$

where c is an arbitrary constant.

More generally:

$$g(x)\,\delta(x-x_0) = g(x_0)\,\delta(x-x_0) \tag{27}$$

(iv) $$\int_{-\infty}^{+\infty} \mathrm{d}x\,\delta(x-y)\,\delta(x-z) = \delta(y-z) \tag{28}$$

Equation (28) can be understood by examining functions $\delta^{(\varepsilon)}(x)$ like the one shown in Figure 1. The integral:

$$F^{(\varepsilon)}(y,z) = \int_{-\infty}^{+\infty} \mathrm{d}x\,\delta^{(\varepsilon)}(x-y)\,\delta^{(\varepsilon)}(x-z) \tag{29}$$

is zero as long as $|y-z| > \varepsilon$, that is, as long as the two square functions do not overlap (Fig. 2).

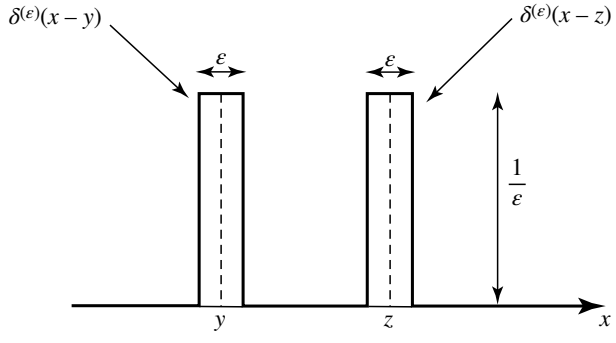


Figure 2: The functions $\delta^{(\varepsilon)}(x-y)$ and $\delta^{(\varepsilon)}(x-z)$: two square functions of width ε and height $1/\varepsilon$, centered respectively at $x=y$ and $x=z$.

The maximum value of the integral, obtained for $y=z$, is equal to $1/\varepsilon$. Between this maximum value and 0, the variation of $F^{(\varepsilon)}(y,z)$ with respect to $y-z$ is linear (Fig. 3). We see immediately that $F^{(\varepsilon)}(y,z)$ approaches $\delta(y-z)$ when $\varepsilon \longrightarrow 0$.

Comment:

A sum of regularly spaced δ-functions:

$$\sum_{q=-\infty}^{+\infty} \delta(x-qL) \tag{30}$$

can be considered to be a periodic "function" of period L. By applying formulas (8), (9) and (10) of Appendix I, we can write it in the form:

$$\sum_{q=-\infty}^{+\infty} \delta(x-qL) = \frac{1}{L} \sum_{n=-\infty}^{+\infty} \mathrm{e}^{2i\pi\frac{nx}{L}} \tag{31}$$

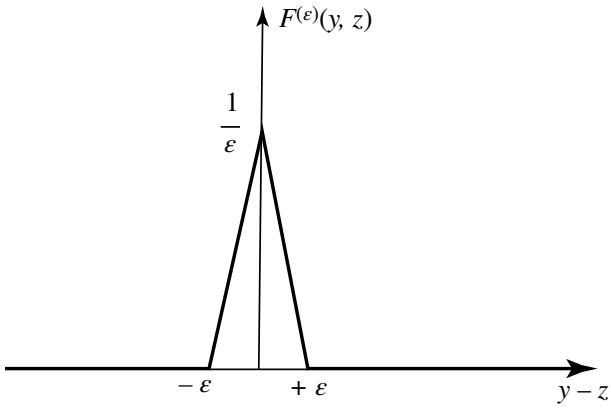


Figure 3: The variation with respect to $y - z$ of the scalar product $F^{(\varepsilon)}(y, z)$ of the two square functions shown in Figure 2. This scalar product is zero when the two functions do not overlap ($|y - z| \geq \varepsilon$), and maximal when they coincide. $F^{(\varepsilon)}(y, z)$ approaches $\delta(y - z)$ when $\varepsilon \longrightarrow 0$.

2. The δ-"function" and the Fourier transform

2-a. The Fourier transform of δ

Definition (28) of Appendix I and equation (5) enable us to calculate directly the Fourier transform $\overline{\delta}_{x_0}(p)$ of $\delta(x - x_0)$:

$$\overline{\delta}_{x_0}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}x \, \mathrm{e}^{-ipx/\hbar} \, \delta(x - x_0) = \frac{1}{\sqrt{2\pi\hbar}} \, \mathrm{e}^{-ipx_0/\hbar} \tag{32}$$

In particular, that of $\delta(x)$ is a constant:

$$\overline{\delta}_0(p) = \frac{1}{\sqrt{2\pi\hbar}} \tag{33}$$

The inverse Fourier transform [formula (29) of Appendix I] then yields:

$$\boxed{\delta(x - x_0) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \mathrm{d}p \, \mathrm{e}^{ip(x-x_0)/\hbar} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k \, \mathrm{e}^{ik(x-x_0)}} \tag{34}$$

This result can also be found by using the function $\delta^{(\varepsilon)}(x)$ defined by (1) or any of the functions given in § 1-b. For example, (48) of Appendix I enables us to write:

$$\delta^{(\varepsilon)}(x) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \mathrm{d}p \, \mathrm{e}^{ipx/\hbar} \, \frac{\sin(p\varepsilon/2\hbar)}{p\varepsilon/2\hbar} \tag{35}$$

If we let ε approach zero, we indeed obtain (34).

2-b. Applications

Expression (34) for the δ-function is often very convenient. We shall show, for example, how it simplifies finding the inverse Fourier transform and the Parseval-Plancherel relation [formulas (29) and (45) of Appendix I].

Starting with:

$$\overline{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}x \; \mathrm{e}^{-ipx/\hbar}\, \psi(x) \tag{36}$$

we calculate:

$$\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}p \; \mathrm{e}^{ipx/\hbar}\; \overline{\psi}(p) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \mathrm{d}\xi\, \psi(\xi) \int_{-\infty}^{+\infty} \mathrm{d}p \; \mathrm{e}^{ip(x-\xi)/\hbar} \tag{37}$$

In the second integral, we recognize $\delta(x-\xi)$, so that:

$$\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}p \; \mathrm{e}^{ipx/\hbar}\; \overline{\psi}(p) = \int_{-\infty}^{+\infty} \mathrm{d}\xi\, \psi(\xi)\; \delta(x-\xi) = \psi(x) \tag{38}$$

which is the inversion formula of the Fourier transform.

Similarly:

$$|\overline{\psi}(p)|^2 = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \mathrm{d}x \; \mathrm{e}^{ipx/\hbar}\; \psi^*(x) \int_{-\infty}^{+\infty} \mathrm{d}x' \, \mathrm{e}^{-ipx'/\hbar}\; \psi(x') \tag{39}$$

If we integrate this expression over p, we find:

$$\int_{-\infty}^{+\infty} \mathrm{d}p \; |\overline{\psi}(p)|^2 = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \mathrm{d}x \; \psi^*(x) \int_{-\infty}^{+\infty} \mathrm{d}x' \, \psi(x') \int_{-\infty}^{+\infty} \mathrm{d}p \; \mathrm{e}^{ip(x-x')/\hbar} \tag{40}$$

that is, according to (34):

$$\int_{-\infty}^{+\infty} \mathrm{d}p \; |\overline{\psi}(p)|^2 = \int_{-\infty}^{+\infty} \mathrm{d}x \; \psi^*(x) \int_{-\infty}^{+\infty} \mathrm{d}x' \psi(x')\, \delta(x-x') = \int_{-\infty}^{+\infty} \mathrm{d}x \; |\psi(x)|^2 \tag{41}$$

which is none other than the Parseval-Plancherel formula.

We can obtain the Fourier transform of a convolution product in an analogous way [*cf.* formulas (39) and (40) of Appendix I].

3. Integral and derivatives of the δ-"function"

3-a. δ is the derivative of the "unit step-function"

We shall evaluate the integral:

$$\theta^{(\varepsilon)}(x) = \int_{-\infty}^{x} \delta^{(\varepsilon)}(x')\, \mathrm{d}x' \tag{42}$$

where the function $\delta^{(\varepsilon)}(x)$ is defined in (1). It can easily be seen that $\theta^{(\varepsilon)}(x)$ is equal to 0 for $x \leq -\dfrac{\varepsilon}{2}$, to 1 for $x \geq \dfrac{\varepsilon}{2}$, and to $\dfrac{1}{\varepsilon}\left(x + \dfrac{\varepsilon}{2}\right)$ for $-\dfrac{\varepsilon}{2} \leq x \leq \dfrac{\varepsilon}{2}$. The variation of $\theta^{(\varepsilon)}(x)$

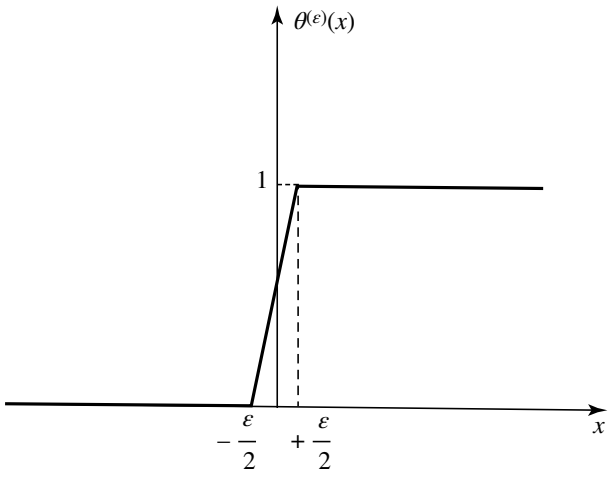


Figure 4: Variation of the function $\theta^{(\varepsilon)}(x)$, whose derivative $\delta^{(\varepsilon)}(x)$ is shown in Figure 1. When $\varepsilon \longrightarrow 0$, $\theta^{(\varepsilon)}(x)$ approaches the Heaviside step-function $\theta(x)$.

with respect to x is shown in Figure 4. When $\varepsilon \to 0$, $\theta^{(\varepsilon)}(x)$ approaches the *Heaviside* "*step-function*" $\theta(x)$, which, by definition, is equal to:

$$\begin{aligned} &\theta(x) = 1 \quad \text{if} \quad x > 0 \\ &\theta(x) = 0 \quad \text{if} \quad x < 0 \end{aligned} \tag{43}$$

$\delta^{(\varepsilon)}(x)$ is the derivative of $\theta^{(\varepsilon)}(x)$. By considering the limit $\varepsilon \longrightarrow 0$, we see that $\delta(x)$ is the derivative of $\theta(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\theta(x) = \delta(x) \tag{44}$$

Now, consider a function $g(x)$ which has a discontinuity σ_0 at $x = 0$:

$$\operatorname*{Lim}_{x\to 0_+} g(x) - \operatorname*{Lim}_{x\to 0_-} g(x) = \sigma_0 \tag{45}$$

Such a function can be written in the form $g(x) = g_1(x)\theta(x) + g_2(x)\theta(-x)$, where $g_1(x)$ and $g_2(x)$ are continuous functions which satisfy $g_1(0) - g_2(0) = \sigma_0$. If we differentiate this expression, using (44), we obtain:

$$\begin{aligned} g'(x) &= g_1'(x)\theta(x) + g_2'(x)\theta(-x) + g_1(x)\delta(x) - g_2(x)\delta(-x) \\ &= g_1'(x)\theta(x) + g_2'(x)\theta(-x) + \sigma_0\delta(x) \end{aligned} \tag{46}$$

according to properties (19) and (27) of δ. For a discontinuous function, there is then added to the ordinary derivative [the first two terms of (46)] a term proportional to the δ-function, the proportionality coefficient being the magnitude of the function's discontinuity[3].

[3]Of course, if the function is discontinuous at $x = x_0$, the additional term is of the form: $[g_1(x_0) - g_2(x_0)]\delta(x - x_0)$.

Comment:

The Fourier transform of the step-function $\theta(k)$ can be found simply from (12). We get:

$$\int_{-\infty}^{+\infty} \theta(k)\, \mathrm{e}^{ikx}\, \mathrm{d}k = \lim_{\varepsilon \to 0_+} \int_0^{\infty} \mathrm{d}k\, \mathrm{e}^{ik(x+i\varepsilon)} = \lim_{\varepsilon \to 0_+} \frac{i}{x + i\varepsilon} = i\mathcal{P}\frac{1}{x} + \pi\delta(x) \quad (47)$$

3-b. Derivatives of δ

By analogy with the expression for integration by parts, the derivative $\delta'(x)$ of the δ-function is defined by the relation[4]:

$$\int_{-\infty}^{+\infty} \mathrm{d}x\ \delta'(x)\ f(x) = -\int_{-\infty}^{+\infty} \mathrm{d}x\ \delta(x)\ f'(x) = -f'(0) \quad (48)$$

From this definition, we immediately get:

$$\delta'(-x) = -\delta'(x) \quad (49)$$

and:

$$x\ \delta'(x) = -\delta(x) \quad (50)$$

Conversely it can be shown that the general solution of the equation:

$$x\ u(x) = \delta(x) \quad (51)$$

can be written:

$$u(x) = -\delta'(x) + c\ \delta(x) \quad (52)$$

where the second term arises from the homogeneous equation [*cf.* formulas (25) and (26)].

Equation (34) allows us to write $\delta'(x)$ in the form:

$$\delta'(x) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \mathrm{d}p \left(\frac{ip}{\hbar}\right) \mathrm{e}^{ipx/\hbar} = \frac{i}{2\pi} \int_{-\infty}^{+\infty} k\, \mathrm{d}k\, \mathrm{e}^{ikx} \quad (53)$$

The nth-order derivative $\delta^{(n)}(x)$ can be defined in the same way:

$$\int_{-\infty}^{+\infty} \mathrm{d}x\, \delta^{(n)}(x)\, f(x) = (-1)^n\, f^{(n)}(0) \quad (54)$$

Relations (49) and (50) can then be generalized to the forms:

$$\delta^{(n)}(-x) = (-1)^n\, \delta^{(n)}(x) \quad (55)$$

and:

$$x\ \delta^{(n)}(x) = -n\, \delta^{(n-1)}(x) \quad (56)$$

[4] $\delta'(x)$ can be considered to be the limit, for $\varepsilon \longrightarrow 0$, of the derivative of one of the functions given in § 1-b.

4. The δ-"function" in three-dimensional space

The δ-"fonction" in three-dimensional space, which we shall write simply as $\delta(\mathbf{r})$, is defined by an expression analogous to (4):

$$\int \mathrm{d}^3 r\, \delta(\mathbf{r})\, f(\mathbf{r}) = f(\mathbf{0}) \tag{57}$$

and, more generally:

$$\int \mathrm{d}^3 r\, \delta(\mathbf{r} - \mathbf{r}_0)\, f(\mathbf{r}) = f(\mathbf{r}_0) \tag{58}$$

$\delta(\mathbf{r} - \mathbf{r}_0)$ can be broken down into a product of three one-dimensional functions:

$$\delta(\mathbf{r} - \mathbf{r}_0) = \delta(x - x_0)\ \delta(y - y_0)\ \delta(z - z_0) \tag{59}$$

or, if we use polar coordinates:

$$\begin{aligned} \delta(\mathbf{r} - \mathbf{r}_0) &= \frac{1}{r^2 \sin\theta}\ \delta(r - r_0)\ \delta(\theta - \theta_0)\ \delta(\varphi - \varphi_0) \\ &= \frac{1}{r^2}\ \delta(r - r_0)\ \delta(\cos\theta - \cos\theta_0)\ \delta(\varphi - \varphi_0) \end{aligned} \tag{60}$$

The properties stated above for $\delta(x)$ can therefore easily be generalized to $\delta(\mathbf{r})$. We shall mention, in addition, the important relation:

$$\Delta\left(\frac{1}{r}\right) = -4\pi\, \delta(\mathbf{r}) \tag{61}$$

where Δ is the Laplacian operator.

Equation (61) can easily be understood if it is recalled that in electrostatics, an electrical point charge q placed at the origin can be described by a volume density $\rho(\mathbf{r})$ equal to:

$$\rho(\mathbf{r}) = q\, \delta(\mathbf{r}) \tag{62}$$

We know that the expression for the electrostatic potential produced by this charge is:

$$U(\mathbf{r}) = \frac{q}{4\pi\varepsilon_0} \frac{1}{r} \tag{63}$$

Equation (61) is thus simply the Poisson equation for this special case:

$$\Delta U(\mathbf{r}) = -\frac{1}{\varepsilon_0} \rho(\mathbf{r}) \tag{64}$$

To prove (61) rigorously, it is necessary to use mathematical distribution theory. We shall confine ourselves here to an elementary "proof".

First of all, note that the Laplacian of $1/r$ is everywhere zero, except, perhaps, at the origin, which is a singular point:

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)\frac{1}{r} = 0 \quad \text{for} \quad r \neq 0 \tag{65}$$

Let $g_\varepsilon(\mathbf{r})$ be a function equal to $1/r$ when $\mathbf{r}$ is outside the sphere S_ε, centered at O and of a radius ε, and which takes on values (of the order of $1/\varepsilon$) inside this sphere such that $g_\varepsilon(\mathbf{r})$ is sufficiently regular (continuous, differentiable, etc.). Let $f(\mathbf{r})$ be an arbitrary function of $\mathbf{r}$ which is also regular at all points in space. We now find the limit of the integral:

$$I(\varepsilon) = \int \mathrm{d}^3 r\, f(\mathbf{r})\, \Delta g_\varepsilon(\mathbf{r}) \tag{66}$$

for $\varepsilon \to 0$. According to (65), this integral can receive contributions only from inside the sphere S_ε, and:

$$I(\varepsilon) = \int_{r \leq \varepsilon} \mathrm{d}^3 r\, f(\mathbf{r})\, \Delta g_\varepsilon(\mathbf{r}) \tag{67}$$

We choose ε small enough for the variation of $f(\mathbf{r})$ inside S_ε to be negligible. Then:

$$I(\varepsilon) \simeq f(\mathbf{0}) \int_{r \leq \varepsilon} \mathrm{d}^3 r\, \Delta g_\varepsilon(\mathbf{r}) \tag{68}$$

Transforming the integral so obtained into an integral over the surface $\mathscr{S}_\varepsilon$ of S_ε, we obtain:

$$I(\varepsilon) \simeq f(\mathbf{0}) \int_{\mathscr{S}_\varepsilon} \boldsymbol{\nabla} g_\varepsilon(\mathbf{r}) \cdot \mathrm{d}\mathbf{n} \tag{69}$$

Now, since $g_\varepsilon(\mathbf{r})$ is continuous on the surface $\mathscr{S}_\varepsilon$, we get:

$$[\boldsymbol{\nabla} g_\varepsilon(\mathbf{r})]_{r=\varepsilon} = \left[-\frac{1}{r^2}\right]_{r=\varepsilon} \mathbf{e}_r = -\frac{1}{\varepsilon^2}\ \mathbf{e}_r \tag{70}$$

(where $\mathbf{e}_r$ is the unit vector $\mathbf{r}/r$). This yields:

$$\begin{aligned} I(\varepsilon) &\simeq f(\mathbf{0}) \times 4\pi\varepsilon^2 \times \left[-\frac{1}{\varepsilon^2}\right] \\ &\simeq -4\pi f(\mathbf{0}) \end{aligned} \tag{71}$$

that is:

$$\operatorname*{Lim}_{\varepsilon \to 0} \int \mathrm{d}^3 r\, \Delta g_\varepsilon(\mathbf{r})\, f(\mathbf{r}) = -4\pi f(\mathbf{0}) \tag{72}$$

According to definition (57), this is simply (61).

Equation (61) can be used, for example, to derive an expression which is useful in collision theory (*cf.* Chap. VIII):

$$(\Delta + k^2)\frac{\mathrm{e}^{\pm ikr}}{r} = -4\pi\, \delta(\mathbf{r}) \tag{73}$$

To do so, it is sufficient to consider $\mathrm{e}^{\pm ikr}/r$ as a product:

$$\Delta\left[\frac{\mathrm{e}^{\pm ikr}}{r}\right] = \frac{1}{r}\ \Delta(\mathrm{e}^{\pm ikr}) + \mathrm{e}^{\pm ikr}\Delta\left(\frac{1}{r}\right) + 2\boldsymbol{\nabla}\left(\frac{1}{r}\right) \cdot \boldsymbol{\nabla}(\mathrm{e}^{\pm ikr}) \tag{74}$$

Now:

$$\begin{aligned} \boldsymbol{\nabla}(\mathrm{e}^{\pm ikr}) &= \pm ik\, \mathrm{e}^{\pm ikr}\frac{\mathbf{r}}{r} \\ \Delta(\mathrm{e}^{\pm ikr}) &= -k^2\mathrm{e}^{\pm ikr} \pm \frac{2ik}{r}\mathrm{e}^{\pm ikr} \end{aligned} \tag{75}$$

We therefore find, finally:

$$(\Delta + k^2)\frac{\mathrm{e}^{\pm ikr}}{r} = \left[-\frac{k^2}{r} \pm \frac{2ik}{r^2} - 4\pi\delta(\mathbf{r}) - \frac{2}{r^2} \times (\pm ik) + \frac{k^2}{r}\right] \mathrm{e}^{\pm ikr}$$
$$= -4\pi\, \mathrm{e}^{\pm ikr}\delta(\mathbf{r})$$
$$= -4\pi\, \delta(\mathbf{r}) \tag{76}$$

according to (27).

Equation (61) can, furthermore, be generalized: the Laplacian of the function $Y_l^m(\theta,\varphi)/r^{l+1}$ involves lth-order derivatives of $\delta(\mathbf{r})$. Consider, for example $\cos\theta/r^2$. We know that the expression for the electrostatic potential created at a distant point by an electric dipole of moment D directed along Oz is $\dfrac{D}{4\pi\varepsilon_0}\dfrac{\cos\theta}{r^2}$. If q is the absolute value of each of the two charges which make up the dipole and a is the distance between them, the modulus D of the dipole moment is the product qa, and the corresponding charge density can be written:

$$\rho(\mathbf{r}) = q\,\delta\left(\mathbf{r} - \frac{a}{2}\mathbf{e}_z\right) - q\,\delta\left(\mathbf{r} + \frac{a}{2}\mathbf{e}_z\right) \tag{77}$$

(where $\mathbf{e}_z$ denotes the unit vector of the Oz axis). If we let a approach zero, while maintaining $D = qa$ finite, this charge density becomes:

$$\rho(\mathbf{r}) \underset{a\to 0}{\longrightarrow} D\,\frac{\partial}{\partial z}\,\delta(\mathbf{r}) \tag{78}$$

Therefore, in the limit where $a \longrightarrow 0$, the Poisson equation (64) yields:

$$\Delta\left(\frac{\cos\theta}{r^2}\right) = -4\pi\frac{\partial}{\partial z}\,\delta(\mathbf{r}) \tag{79}$$

Of course, this formula could be justified as (61) was above, or proven by distribution theory. Analogous reasoning could be applied to the function $Y_l^m(\theta,\varphi)/r^{l+1}$ which gives the potential created by an electric multipole moment $\mathcal{Q}_l^m$ located at the origin (complement $\mathrm{E_X}$).

References and suggestions for further reading:

See Dirac (1.13) § 15, and, for example, Butkov (10.8), Chap. 6, or Bass (10.1), vol. I, §§ 21.7 and 21.8; section 10 of the bibliography, especially the subsection "Fourier transforms; distributions".

Appendix III

Lagrangian and Hamiltonian in classical mechanics

1 Review of Newton's laws 1527
1-a Dynamics of a point particle 1527
1-b Systems of point particles 1528
1-c Fundamental theorems 1528
2 The Lagrangian and Lagrange's equations 1530
3 The classical Hamiltonian and the canonical equations . . . 1531
3-a The conjugate momenta of the coordinates 1531
3-b The Hamilton-Jacobi canonical equations 1532
4 Applications of the Hamiltonian formalism 1533
4-a A particle in a central potential 1533
4-b A charged particle placed in an electromagnetic field 1536
5 The principle of least action 1539
5-a Geometrical representation of the motion of a system 1539
5-b The principle of least action 1540
5-c Lagrange's equations as a consequence of the principle of least action 1541

We shall review the definition and principal properties of the Lagrangian and the Hamiltonian in classical mechanics. This appendix is not meant to be a course in analytical mechanics. Its goal is simply to indicate the classical basis for applying the quantization rules (*cf.* Chap. III) to a physical system. In particular, we shall concern ourselves essentially with systems of point particles.

1. Review of Newton's laws

1-a. Dynamics of a point particle

Non-relativistic classical mechanics is based on the hypothesis that there exists at least one geometrical frame, called the *Galilean* or *inertial frame*, in which the following law is valid:

The fundamental law of dynamics: a point particle has, at all times, an acceleration γ which is proportional to the resultant $\mathbf{F}$ of the forces acting on it:

$$\mathbf{F} = m\gamma \tag{1}$$

The constant m is an intrinsic property of the particle, called its *inertial mass.*

It can easily be shown that if a Galilean frame exists, all frames which are in uniform translational motion with respect to it are also Galilean frames. This leads us to the *Galilean relativity principle*: there is no absolute frame; there is no experiment which can give one inertial frame a privileged role with respect to all others.

1-b. Systems of point particles

If we are dealing with a system composed of n point particles, we apply the fundamental law to each of them[1]:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i \quad ; \quad i = 1, 2, \ldots n \tag{2}$$

The forces that act on the particles can be classed in two categories: *internal forces* represent the interactions between the particles of the system, and *external forces* originate outside the system. The internal forces are postulated to satisfy the *principle of action and reaction*: the force exerted by particle (i) on particle (j) is equal and opposite to the one exerted by (j) on (i). This principle is true for gravitational forces (Newton's law) and electrostatic forces, but not for magnetic forces (whose origin is relativistic).

If all the forces can be derived from a potential, the equations of motion (2) can be written:

$$m_i \ddot{\mathbf{r}}_i = -\boldsymbol{\nabla}_i V \tag{3}$$

where $\boldsymbol{\nabla}_i$ denotes the gradient with respect to the $\mathbf{r}_i$ coordinates, and the potential energy V is of the form:

$$V = \sum_{i=1}^{n} V_i(\mathbf{r}_i) + \sum_{i<j} V_{ij}(\mathbf{r}_i - \mathbf{r}_j) \tag{4}$$

(the first term in this expression corresponds to the external forces, and the second one to the internal forces). *In cartesian coordinates*, the motion of the system is therefore described by the $3n$ differential equations:

$$\left.\begin{aligned} m_i \ddot{x}_i &= -\frac{\partial V}{\partial x_i} \\ m_i \ddot{y}_i &= -\frac{\partial V}{\partial y_i} \\ m_i \ddot{z}_i &= -\frac{\partial V}{\partial z_i} \end{aligned}\right\} i = 1, 2, \ldots n \tag{5}$$

1-c. Fundamental theorems

We shall first review a few definitions. The *center of mass* or *center of gravity* of a system is the point G whose coordinates are:

$$\mathbf{r}_G = \frac{\sum_{i=1}^{n} m_i \mathbf{r}_i}{\sum_{i=1}^{n} m_i} \tag{6}$$

[1] In mechanics, a simplified notation is generally used for the time-derivatives; by definition, $\dot{u} = \dfrac{\mathrm{d}u}{\mathrm{d}t}$, $\ddot{u} = \dfrac{\mathrm{d}^2 u}{\mathrm{d}t^2}$, etc...

The total *kinetic energy* of the system is equal to:

$$T = \sum_{i=1}^{n} \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 \tag{7}$$

where $\dot{\mathbf{r}}_i$ is the velocity of particle (i). The *angular momentum* with respect to the origin is the vector:

$$\boldsymbol{\mathscr{L}} = \sum_{i=1}^{n} \mathbf{r}_i \times m_i \dot{\mathbf{r}}_i \tag{8}$$

The following theorems can then be easily proven:

(i) The center of mass of a system moves like a point particle with a mass equal to the total mass of the system, subject to a force equal to the resultant of all the forces involved in the system:

$$\left[\sum_{i=1}^{n} m_i \right] \ddot{\mathbf{r}}_G = \sum_{i=1}^{n} \mathbf{F}_i \tag{9}$$

(ii) The time-derivative of the angular momentum evaluated at a fixed point is equal to the moment of the forces with respect to this point:

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\mathscr{L}} = \sum_{i=1}^{n} \mathbf{r}_i \times \mathbf{F}_i \tag{10}$$

(iii) The variation of the kinetic energy between time t_1 and t_2 is equal to the work performed by all the forces during the motion between these two times:

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \sum_{i=1}^{n} \mathbf{F}_i \cdot \dot{\mathbf{r}}_i \mathrm{d}t \tag{11}$$

If the internal forces satisfy the principle of action and reaction, and if they are directed along the straight lines joining the interacting particles, their contribution to the resultant [equation (9)] and to the moment with respect to the origin [equation (10)] is zero. If, in addition, the system is isolated (that is, if it is not subject to any external forces), the total angular momentum $\boldsymbol{\mathscr{L}}$ is constant, and the center of mass is in uniform rectilinear motion. This means that the total mechanical momentum:

$$\sum_{i=1}^{n} m_i \dot{\mathbf{r}}_i \tag{12}$$

is also a constant of the motion.

2. The Lagrangian and Lagrange's equations

Consider a system of n particles in which the forces are derived from a potential energy [*cf.* formula (4)], which we shall write simply $V(\mathbf{r}_i)$. The *Lagrangian*, or *Lagrange's function*, of this system is the function of $6n$ variables $\{x_i, y_i, z_i; \dot{x}_i, \dot{y}_i, \dot{z}_i \ ; \ i = 1, 2, ..., n\}$ given by:

$$\begin{aligned}\mathcal{L}(\mathbf{r}_i, \dot{\mathbf{r}}_i) &= T - V \\ &= \frac{1}{2}\sum_{i=1}^{n} m_i \dot{\mathbf{r}}_i^2 - V(\mathbf{r}_i)\end{aligned} \tag{13}$$

It can immediately be shown that the equations of motion written in (5) are identical to *Lagrange's equations*:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} &= 0 \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{y}_i} - \frac{\partial \mathcal{L}}{\partial y_i} &= 0 \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{z}_i} - \frac{\partial \mathcal{L}}{\partial z_i} &= 0\end{aligned} \tag{14}$$

A very interesting feature of Lagrange's equations is that they always have the same form, independent of the type of coordinates used (whether they are cartesian or not). In addition, they can be applied to systems which are more general than particle systems. Many physical systems (including for example one or several solid bodies) can be described at a given time t by a set of N independent parameters $q_i (i = 1, 2, ..., N)$, called *generalized coordinates*. Knowledge of the q_i permits the calculation of the position in space of any point of the system. The motion of this system is therefore characterized by specifying the N functions of time $q_i(t)$. The time-derivatives $\dot{q}_i(t)$ are called the *generalized velocities*. The state of the system at a given instant t_0 is therefore defined by the set of $q_i(t_0)$ and $\dot{q}_i(t_0)$. If the forces acting on the system can be derived from a potential energy $V(q_1, q_2, ..., q_N)$, the Lagrangian $\mathcal{L}(q_1, q_2, ..., q_N; \dot{q}_1, \dot{q}_2, ..., \dot{q}_N)$ is again the difference between the total kinetic energy T and the potential energy V. It can be shown that, for any choice of the coordinates q_i, the equations of motion can always be written:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0} \tag{15}$$

where $\dfrac{\mathrm{d}}{\mathrm{d}t}$ denotes the *total time-derivative*

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \sum_{i=1}^{N} \dot{q}_i \frac{\partial}{\partial q_i} + \sum_{i=1}^{N} \ddot{q}_i \frac{\partial}{\partial \dot{q}_i} \tag{16}$$

Furthermore, it is not really necessary for the forces to be derived from a potential for us to be able to define a Lagrangian and use Lagrange's equations (we shall see an example

of this situation in § 4-b). In the general case, the Lagrangian is a function of the coordinates q_i and the velocities $\dot{q}_i$, and can also be explicitly time-dependent[2]. We shall then write it:

$$\mathcal{L}(q_i, \dot{q}_i; t) \tag{17}$$

Lagrange's equations are important in classical mechanics for several reasons. For one thing, as we have just indicated, they always have the same form, independent of the coordinates which are used. Furthermore, they are more convenient than Newton's equations when the system is complex. Finally, they are of considerable theoretical interest, since they form the foundation of the Hamiltonian formalism (*cf.* § 3 below), and since they can be derived from a variational principle (§ 5). The first two points are secondary as far as quantum mechanics is concerned, since quantum mechanics treats particle systems almost exclusively and since the quantization rules are stated in cartesian coordinates (*cf.* Chap. III, § B-5). However, the last point is an essential one, since the Hamiltonian formalism constitutes the point of departure for the quantization of physical systems.

3. The classical Hamiltonian and the canonical equations

For a physical system described by N generalized coordinates, Lagrange's equations (15) constitute a system of N coupled second-order differential equations with N unknown functions, the $q_i(t)$. We shall see that this system can be replaced by a system of $2N$ first-order equations with $2N$ unknown functions.

3-a. The conjugate momenta of the coordinates

The conjugate momentum p_i of the generalized coordinate q_i is defined as:

$$\boxed{p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}} \tag{18}$$

p_i is also called the *generalized momentum*. In the case of a particle system for which the forces are derived from a potential energy, the conjugate momenta of the position variables $\mathbf{r}_i(x_1, y_i, z_i)$ are simply [see (13)] the mechanical momenta:

$$\mathbf{p}_i = m_i \dot{\mathbf{r}}_i \tag{19}$$

However, we shall see in § 4-b-γ that this is no longer true in the presence of a magnetic field.

Instead of defining the state of the system at a given time t by the N coordinates $q_i(t)$ and the N velocities $\dot{q}_i(t)$, we shall henceforth characterize it by the $2N$ variables:

$$\{q_i(t),\, p_i(t); i = 1, 2,\, ...,\, N\} \tag{20}$$

[2]The Lagrangian is not unique: two functions $\mathcal{L}(q_i, \dot{q}_i; t)$ and $\mathcal{L}'(q_i, \dot{q}_i; t)$ may lead, using (15), to the same equations of motion. This is true, in particular, if the difference between $\mathcal{L}$ and $\mathcal{L}'$ is the total derivative with respect to time of a function $F(q_i; t)$..

$$\mathcal{L}' - \mathcal{L} = \frac{\mathrm{d}}{\mathrm{d}t} F(q_i, t) \equiv \frac{\partial F}{\partial t} + \sum_i \dot{q}_i \frac{\partial F}{\partial q_i}$$

This amounts to assuming that from the $2N$ parameters $q_i(t)$ and $p_i(t)$, we can determine the $\dot{q}_i(t)$ uniquely. These variables may be considered as the $2N$ coordinates of a point defining the state of the system at every time, and moving in a $2N$ dimensional space called the *phase space*.

3-b. The Hamilton-Jacobi canonical equations

The *classical Hamiltonian*, or *Hamilton's function*, of the system is, by definition:

$$\mathcal{H} = \sum_{i=1}^{N} p_i \dot{q}_i - \mathcal{L} \tag{21}$$

In accordance with convention (20), we eliminate the $\dot{q}_i$ and consider the Hamiltonian to be a function of the coordinates and their conjugate momenta. Like $\mathcal{L}$, $\mathcal{H}$ may be explicitly time-dependent:

$$\mathcal{H}(q_i, p_i; t) \tag{22}$$

The total differential of the function $\mathcal{H}$:

$$\mathrm{d}\mathcal{H} = \sum_i \frac{\partial \mathcal{H}}{\partial q_i}\mathrm{d}q_i + \sum_i \frac{\partial \mathcal{H}}{\partial p_i}\,\mathrm{d}p_i + \frac{\partial \mathcal{H}}{\partial t}\mathrm{d}t \tag{23}$$

is equal to, using definitions (21) and (18):

$$\begin{aligned}\mathrm{d}\mathcal{H} &= \sum_i \left[p_i\,\mathrm{d}\dot{q}_i + \dot{q}_i\,\mathrm{d}p_i\right] - \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\mathrm{d}q_i - \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\mathrm{d}\dot{q}_i - \frac{\partial \mathcal{L}}{\partial t}\mathrm{d}t \\ &= \sum_i \dot{q}_i\,\mathrm{d}p_i - \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\mathrm{d}q_i - \frac{\partial \mathcal{L}}{\partial t}\mathrm{d}t \end{aligned} \tag{24}$$

Setting (23) and (24) equal, we see that the change from the $\{q_i, \dot{q}_i\}$ variables to the $\{q_i, p_i\}$ variables leads to:

$$\frac{\partial \mathcal{H}}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i} \tag{25a}$$

$$\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i \tag{25b}$$

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t} \tag{25c}$$

Furthermore, using (18) and (25a), we can write Lagrange's equations (15) in the form:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_i = -\frac{\partial \mathcal{H}}{\partial q_i} \tag{26}$$

By grouping terms in (25b) and (26), we obtain the equations of motion:

$$\boxed{\begin{aligned}\frac{\mathrm{d}q_i}{\mathrm{d}t} &= \frac{\partial \mathcal{H}}{\partial p_i} \\ \frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\frac{\partial \mathcal{H}}{\partial q_i}\end{aligned}} \tag{27}$$

which are called the Hamilton-Jacobi canonical equations. As we said, (27) is a system of $2N$ first-order differential equations for $2N$ unknown functions, the $q_i(t)$ and $p_i(t)$. These equations determine the motion of the point in the phase space.

For an n-particle system whose potential energy is $V(\mathbf{r}_i)$, we have, according to (13):

$$\begin{aligned}\mathcal{H} &= \sum_{i=1}^{n} \mathbf{p}_i \cdot \dot{\mathbf{r}}_i - \mathcal{L} \\ &= \sum_{i=1}^{n} \mathbf{p}_i \cdot \dot{\mathbf{r}}_i - \frac{1}{2}\sum_{i=1}^{n} m_i \dot{\mathbf{r}}_i^2 + V(\mathbf{r}_i)\end{aligned} \tag{28}$$

To express the Hamiltonian in terms of the variables $\mathbf{r}_i$ and $\mathbf{p}_i$, we use (19). This yields:

$$\mathcal{H}(\mathbf{r}_i, \mathbf{p}_i) = \sum_{i=1}^{n} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_i) \tag{29}$$

Note that the Hamiltonian is thus equal to the *total energy* of the system. The canonical equations:

$$\begin{aligned}\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} &= \frac{\mathbf{p}_i}{m_i} \\ \frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} &= -\boldsymbol{\nabla}_i V\end{aligned} \tag{30}$$

are equivalent to Newton's equations, (3).

4. Applications of the Hamiltonian formalism

4-a. A particle in a central potential

Consider a system composed of a single particle of mass m whose potential energy $V(r)$ depends only on its distance from the origin. In polar coordinates (r, θ, φ), the components of the particle's velocity on the local axes (Fig. 1) are:

$$\begin{aligned}v_r &= \dot{r} \\ v_\theta &= r\,\dot{\theta} \\ v_\varphi &= r\sin\theta\,\dot{\varphi}\end{aligned} \tag{31}$$

so that the Lagrangian, (13), can be written:

$$\mathcal{L}(r, \theta, \varphi; \dot{r}, \dot{\theta}, \dot{\varphi}) = \frac{1}{2} m \left[\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\,\dot{\varphi}^2\right] - V(r) \tag{32}$$

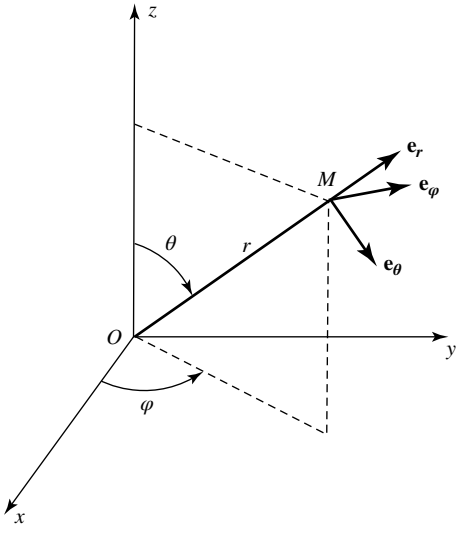


Figure 1: The unit vectors $\mathbf{e_r}$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ *of the local axes associated with point* M, *where* M *is defined by its spherical coordinates* r, θ, φ.

The conjugate momenta of the three variables r, θ, φ can then be calculated:

$$p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r} \tag{33a}$$

$$p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\ \dot{\theta} \tag{33b}$$

$$p_\varphi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = mr^2\ \sin^2\theta\ \dot{\varphi} \tag{33c}$$

To obtain the Hamiltonian of the particle, we use definition (21). This amounts to adding $V(r)$ to the kinetic energy, expressed in terms of r, θ, φ and p_r, p_θ, p_φ. We find:

$$\mathcal{H}(r, \theta, \varphi; p_r, p_\theta, p_\varphi) = \frac{p_r^2}{2m} + \frac{1}{2mr^2}\left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right) + V(r) \tag{34}$$

The system of canonical equations [formulas (27)] can be written here:

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial p_r} = \frac{p_r}{m} \tag{35a}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial p_\theta} = \frac{p_\theta}{mr^2} \tag{35b}$$

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial p_\varphi} = \frac{p_\varphi}{mr^2 \sin^2\theta} \tag{35c}$$

$$\frac{\mathrm{d}p_r}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial r} = \frac{1}{mr^3}\left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right) - \frac{\partial V}{\partial r} \tag{35d}$$

$$\frac{\mathrm{d}p_\theta}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial \theta} = \frac{p_\varphi^2 \cos\theta}{mr^2 \sin^3\theta} \tag{35e}$$

$$\frac{\mathrm{d}p_\varphi}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial \varphi} = 0 \tag{35f}$$

The first three of these equations simply give (33); the last three are the real equations of motion.

Now, consider the angular momentum of the particle with respect to the origin:

$$\boldsymbol{\mathscr{L}} = \mathbf{r} \times m\,\mathbf{v} \tag{36}$$

Its local components can easily be calculated from (31):

$$\begin{aligned} \mathscr{L}_r &= 0 \\ \mathscr{L}_\theta &= -mr\ v_\varphi = -mr^2\ \sin\theta\ \dot{\varphi} = -\frac{p_\varphi}{\sin\theta} \\ \mathscr{L}_\varphi &= mr\ v_\theta = mr^2\ \dot{\theta} = p_\theta \end{aligned} \tag{37}$$

so that:

$$\boldsymbol{\mathscr{L}}^2 = p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta} \tag{38}$$

From the angular momentum theorem [formula (10)], we know that $\boldsymbol{\mathscr{L}}$ is a vector which is constant over time, since the force derived from the potential $V(r)$ is central, that is, collinear at each instant[3] with the vector $\mathbf{r}$.

By comparing (34) and (38), we see that the Hamiltonian $\mathcal{H}$ depends on the angular variables and their conjugate momenta only through the intermediary of $\boldsymbol{\mathscr{L}}^2$:

$$\mathcal{H}(r, \theta, \varphi; p_r, p_\theta, p_\varphi) = \frac{p_r^2}{2m} + \frac{1}{2mr^2}\ \boldsymbol{\mathscr{L}}^2\,(\theta, p_\theta, p_\varphi) + V(r) \tag{39}$$

Now, assume that the initial angular momentum of the particle is $\boldsymbol{\mathscr{L}}_0$. Since the angular momentum remains constant, the Hamiltonian (39) and the equation of motion (35d)

[3]This conclusion can also be derived from (35e) and (35f) by calculating the time-derivatives of the components of $\boldsymbol{\mathscr{L}}$ on the fixed axes Ox, Oy, Oz.

are the same as they would be for a particle of mass m, in a one-dimensional problem, placed in the effective potential:

$$V_{\text{eff}}(r) = V(r) + \frac{\mathscr{L}_0^2}{2mr^2} \tag{40}$$

4-b. A charged particle placed in an electromagnetic field

Now, consider a particle of mass m and charge q placed in an electromagnetic field characterized by the electric field vector $\mathbf{E}(\mathbf{r},t)$ and the magnetic field vector $\mathbf{B}(\mathbf{r},t)$.

α. Description of the electromagnetic field. Gauges

$\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$ satisfy Maxwell's equations:

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \frac{\rho}{\varepsilon_0} \tag{41a}$$

$$\boldsymbol{\nabla}\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} \tag{41b}$$

$$\boldsymbol{\nabla}\cdot\mathbf{B} = 0 \tag{41c}$$

$$\boldsymbol{\nabla}\times\mathbf{B} = \mu_0\,\mathbf{j} + \varepsilon_0\mu_0\frac{\partial\mathbf{E}}{\partial t} \tag{41d}$$

where $\rho(\mathbf{r},t)$ and $\mathbf{j}(\mathbf{r},t)$ are the volume charge density and the current density producing the electromagnetic field. The fields $\mathbf{E}$ and $\mathbf{B}$ can be described by a scalar potential $U(\mathbf{r},t)$ and a vector potential $\mathbf{A}(\mathbf{r},t)$, since equation (41c) implies that there exists a vector field $\mathbf{A}(\mathbf{r},t)$ such that:

$$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}(\mathbf{r},t) \tag{42}$$

(41b) can thus be written:

$$\boldsymbol{\nabla}\times\left[\mathbf{E} + \frac{\partial\mathbf{A}}{\partial t}\right] = \mathbf{0} \tag{43}$$

Consequently, there exists a scalar function $U(\mathbf{r},t)$ such that:

$$\mathbf{E} + \frac{\partial\mathbf{A}}{\partial t} = -\boldsymbol{\nabla}U(\mathbf{r},t) \tag{44}$$

The set of the two potentials $\mathbf{A}(\mathbf{r},t)$ and $U(\mathbf{r},t)$ constitutes what is called a *gauge* for describing the electromagnetic field. The electric and magnetic fields can be calculated from the $\{\mathbf{A},U\}$ gauge by:

$$\mathbf{B}(\mathbf{r},t) = \boldsymbol{\nabla}\times\mathbf{A}(\mathbf{r},t) \tag{45a}$$

$$\mathbf{E}(\mathbf{r},t) = -\boldsymbol{\nabla}U(\mathbf{r},t) - \frac{\partial}{\partial t}\mathbf{A}(\mathbf{r},t) \tag{45b}$$

A given electromagnetic field, that is, a pair of fields $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$, can be described by an infinite number of gauges, which, for this reason, are said to be equivalent.

If we know one gauge, $\{\mathbf{A}, U\}$, which yields the fields $\mathbf{E}$ and $\mathbf{B}$, all the equivalent gauges, $\{\mathbf{A}', U'\}$, can be found from the *gauge transformation formulas*:

$$\mathbf{A}'(\mathbf{r},t) = \mathbf{A}(\mathbf{r},t) + \boldsymbol{\nabla}\chi(\mathbf{r},t) \tag{46a}$$

$$U'(\mathbf{r},t) = U(\mathbf{r},t) - \frac{\partial}{\partial t}\chi(\mathbf{r},t) \tag{46b}$$

where $\chi(\mathbf{r},t)$ is any scalar function.

First of all, it is easy to show from (46) that:

$$\begin{cases} \boldsymbol{\nabla}\times\mathbf{A}'(\mathbf{r},t) = \boldsymbol{\nabla}\times\mathbf{A}(\mathbf{r},t) \\ -\boldsymbol{\nabla}U'(\mathbf{r},t) - \dfrac{\partial}{\partial t}\mathbf{A}'(\mathbf{r},t) = -\boldsymbol{\nabla}U(\mathbf{r},t) - \dfrac{\partial}{\partial t}\mathbf{A}(\mathbf{r},t) \end{cases} \tag{47}$$

Any gauge, $\{\mathbf{A}', U'\}$, which satisfies (46) therefore yields the same electric and magnetic fields as $\{\mathbf{A}, U\}$.

Conversely we shall show that if two gauges, $\{\mathbf{A}, U\}$ and $\{\mathbf{A}', U'\}$, are equivalent, there must exist a function $\chi(\mathbf{r},t)$ which establishes relations (46) between them. Since, by hypothesis:

$$\mathbf{B}(\mathbf{r},t) = \boldsymbol{\nabla}\times\mathbf{A}(\mathbf{r},t) = \boldsymbol{\nabla}\times\mathbf{A}'(\mathbf{r},t) \tag{48}$$

we have:

$$\boldsymbol{\nabla}\times(\mathbf{A}'-\mathbf{A}) = \mathbf{0} \tag{49}$$

This implies that $\mathbf{A}'-\mathbf{A}$ is the gradient of a scalar function:

$$\mathbf{A}'-\mathbf{A} = \boldsymbol{\nabla}\chi(\mathbf{r},t) \tag{50}$$

$\chi(\mathbf{r},t)$ is, for the moment, determined only to within an arbitrary function of t, $f(t)$. Furthermore, the fact that the two gauges are equivalent means that:

$$\mathbf{E}(\mathbf{r},t) = -\boldsymbol{\nabla}U(\mathbf{r},t) - \frac{\partial}{\partial t}\mathbf{A}(\mathbf{r},t) = -\boldsymbol{\nabla}U'(\mathbf{r},t) - \frac{\partial}{\partial t}\mathbf{A}'(\mathbf{r},t) \tag{51}$$

that is:

$$\boldsymbol{\nabla}(U'-U) + \frac{\partial}{\partial t}(\mathbf{A}'-\mathbf{A}) = \mathbf{0} \tag{52}$$

According to (50), we must have:

$$\boldsymbol{\nabla}(U'-U) = -\boldsymbol{\nabla}\frac{\partial}{\partial t}\chi(\mathbf{r},t) \tag{53}$$

Consequently, the functions $U'-U$ and $-\dfrac{\partial}{\partial t}\chi(\mathbf{r},t)$ can differ only by a function of t; thus, we can choose $f(t)$ so as to make them equal:

$$U'-U = -\frac{\partial}{\partial t}\chi(\mathbf{r},t) \tag{54}$$

This completes the determination of the function $\chi(\mathbf{r},t)$ (to within an additive constant). Two equivalent gauges must therefore satisfy relations of the form (46).

β. Equations of motion and the Lagrangian

In the electromagnetic field, the charged particle is subject to the *Lorentz force*:

$$\mathbf{F} = q\,[\mathbf{E} + \mathbf{v} \times \mathbf{B}] \tag{55}$$

(where $\mathbf{v}$ is the velocity of the particle at the time t). Newton's law therefore gives the equations of motion in the form:

$$m\ddot{\mathbf{r}} = q\,[\mathbf{E}(\mathbf{r},t) + \dot{\mathbf{r}} \times \mathbf{B}(\mathbf{r},t)] \tag{56}$$

Projecting this equation onto Ox and using (45), we obtain:

$$\begin{aligned} m\ddot{x} &= q\,[E_x + \dot{y}B_z - \dot{z}B_y] \\ &= q\left[-\frac{\partial U}{\partial x} - \frac{\partial A_x}{\partial t} + \dot{y}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - \dot{z}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\right] \end{aligned} \tag{57}$$

It can easily be shown that these equations can be derived from the Lagrangian by using (15):

$$\mathcal{L}(\mathbf{r},\dot{\mathbf{r}},t) = \frac{1}{2}m\ \dot{\mathbf{r}}^2 + q\ \dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r},t) - qU(\mathbf{r},t) \tag{58}$$

Therefore, although the Lorentz force is not derived from a potential energy, we can find a Lagragian for the problem.

Let us show that Lagrange's equations (15) do yield the equations of motion (56), using the Lagrangian (58). To do so, we shall first calculate:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \dot{x}} &= m\dot{x} + qA_x(\mathbf{r},t) \\ \frac{\partial \mathcal{L}}{\partial x} &= q\dot{\mathbf{r}} \cdot \frac{\partial}{\partial x}\mathbf{A}(\mathbf{r},t) - q\frac{\partial}{\partial x}U(\mathbf{r},t) \end{aligned} \tag{59}$$

Lagrange's equation for the x-coordinate can therefore be written:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m\dot{x} + qA_x(\mathbf{r},t)\right] - q\dot{\mathbf{r}} \cdot \frac{\partial}{\partial x}\mathbf{A}(\mathbf{r},t) + q\frac{\partial}{\partial x}U(\mathbf{r},t) = 0 \tag{60}$$

Writing this equation explicitly and using (16), we again get (57):

$$m\ddot{x} + q\left[\frac{\partial A_x}{\partial t} + \dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_x}{\partial y} + \dot{z}\frac{\partial A_x}{\partial z}\right] - q\left[\dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_y}{\partial x} + \dot{z}\frac{\partial A_z}{\partial x}\right] + q\frac{\partial U}{\partial x} = 0 \tag{61}$$

that is:

$$m\ddot{x} = q\left[-\frac{\partial U}{\partial x} - \frac{\partial A_x}{\partial t} + \dot{y}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - \dot{z}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\right] \tag{62}$$

γ. Momentum. The classical Hamiltonian

The Lagrangian (58) enables us to calculate the conjugate momenta of the cartesian coordinates x, y, z of the particle. For example:

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} + qA_x(\mathbf{r}, t) \tag{63}$$

The *momentum of the particle*, which is, by definition, the vector whose components are (p_x, p_y, p_z), *is no longer equal*, as it was in (19), to the mechanical momentum $m\,\dot{\mathbf{r}}$:

$$\mathbf{p} = m\dot{\mathbf{r}} + q\mathbf{A}(\mathbf{r}, t) \tag{64}$$

Finally, we shall write the classical Hamiltonian:

$$\begin{aligned} \mathcal{H}(\mathbf{r}, \mathbf{p}; t) &= \mathbf{p} \cdot \dot{\mathbf{r}} - \mathcal{L} \\ &= \mathbf{p} \cdot \frac{1}{m}(\mathbf{p} - q\mathbf{A}) - \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 - \frac{q}{m}(\mathbf{p} - q\mathbf{A}) \cdot \mathbf{A} + qU \end{aligned} \tag{65}$$

that is:

$$\mathcal{H}(\mathbf{r}, \mathbf{p}; t) = \frac{1}{2m}\left[\mathbf{p} - q\mathbf{A}(\mathbf{r}, t)\right]^2 + qU(\mathbf{r}, t) \tag{66}$$

Comment:

Hamiltonian formalism therefore uses the potentials $\mathbf{A}$ and U, and not the fields $\mathbf{E}$ and $\mathbf{B}$ directly. The result is that the description of the particle depends on the gauge chosen. It is reasonable to expect, however, since the Lorentz force is expressed in terms of the fields, that predictions concerning the physical behavior of the particle must be the same for two equivalent gauges. The physical consequences of the Hamiltonian formalism are said to be *gauge-invariant*. The concept of gauge invariance is analyzed in detail in Complement $\mathrm{H_{III}}$.

5. The principle of least action

Classical mechanics can be based on a variational principle, the principle of least action. In addition to its theoretical importance, the concept of action serves as the foundation of the *Lagrangian formulation of quantum mechanics* (*cf.* Complement $\mathrm{J_{III}}$). This is why we shall now briefly discuss the principle of least action and show how it leads to Lagrange's equations.

5-a. Geometrical representation of the motion of a system

First of all, consider a particle constrained to move along the Ox axis. Its motion can be represented by tracing, in the (x, t) plane, the curve defined by the law of motion which yields $x(t)$.

More generally, let us study a physical system described by N generalized coordinates q_i (for an n-particle system in three-dimensional space, $N = 3n$). It is convenient to interpret the q_i to be the coordinates of a point Q in an N-dimensional Euclidean space R_N. There is then a one-to-one correspondence between the positions of the system and the points of R_N.

With each motion of the system is associated a motion of point Q in R_N, characterized by the N-dimensional vector function $Q(t)$ whose components are the $q_i(t)$. As in the simple case of a single particle moving in one dimension, the motion of point Q, that is, the motion of the system, can be represented by the graph of $Q(t)$, which is a curve in an $(N+1)$-dimensional space-time (the time axis is added to the N dimensions of R_N). This curve characterizes the motion being studied.

5-b. The principle of least action

The $q_i(t)$ can be fixed arbitrarily; this gives point Q and the system an arbitrary motion. But their real behavior is defined by the initial conditions and the equations of motion. Suppose that we know that, in the course of the real motion. Q is at Q_1 at time t_1 and at Q_2 at a subsequent time t_2 (as is shown schematically by Figure 2):

$$Q(t_1) = Q_1$$

$$Q(t_2) = Q_2 \tag{67}$$

There is an infinite number of *a priori* possible motions which satisfy conditions (67). They are represented by all the curves[4], or *paths in space time*, which connect the points (Q_1, t_1) and (Q_2, t_2) (*cf.* Fig. 2).

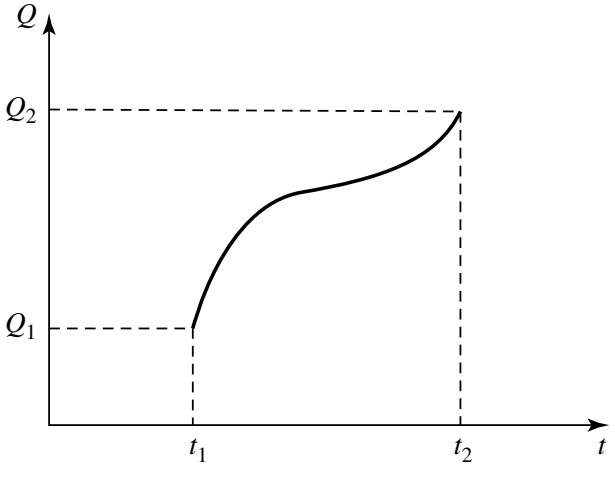


Figure 2: The path in space-time which is associated with a given motion of the physical system. The "x-axis" represents the time and the "y-axis", Q (which symbolizes the set of generalized coordinates q_i).

Consider such a path in space-time Γ, characterized by the vector function $Q(t)$ which satisfies (67). If:

$$\mathcal{L}(q_1, q_2, \ldots, q_N; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_N; t) \equiv \mathcal{L}(Q, \dot{Q}; t) \tag{68}$$

[4]Excluding, of course, the curves which "go backward", that is, which would give two distinct positions of Q for the same time t.

is the Lagrangian of the system, the *action* S_Γ which corresponds to the path Γ is, by definition:

$$S_\Gamma = \int_{t_1}^{t_2} \mathrm{d}t\, \mathcal{L}\left[Q_\Gamma(t), \dot{Q}_\Gamma(t); t\right] \tag{69}$$

[the function to be integrated depends only on t; it is obtained by replacing the q_i and $\dot{q}_i$ by the time-dependent coordinates of $Q_\Gamma(t)$ and $\dot{Q}_\Gamma(t)$ in the Lagrangian (68)].

The *principle of least action* can then be stated in the following way: of all the paths in space-time connecting (Q_1, t_1) with (Q_2, t_2), the one which is actually followed (that is, the one which characterizes the real motion of the system) is the one for which the action is minimal. In other words, when we go from the path which is actually followed to one infinitely close to it, the action does not vary to first order. Note the analogy with other variational principles, such as Fermat's principle in optics.

5-c. Lagrange's equations as a consequence of the principle of least action

In conclusion, we shall show how Lagrange's equations can be deduced from the principle of least action.

Suppose that the real motion of the system under study is characterized by the N functions of time $q_i(t)$, that is by the path in space-time Γ connecting the points (Q_1, t_1) and (Q_2, t_2). Now consider an infinitely close path, Γ' (fig. 3), for which the generalized coordinates are equal to:

$$q_i'(t) = q_i(t) + \delta q_i(t) \tag{70}$$

where the $\delta q_i(t)$ are infinitesimally small and satisfy conditions (67), that is:

$$\delta q_i(t_1) = \delta q_i(t_2) = 0 \tag{71}$$

The generalized velocities $\dot{q}_i'(t)$ corresponding to Γ' can be obtained by differentiating relations (70):

$$\dot{q}_i'(t) = \dot{q}_i(t) + \frac{\mathrm{d}}{\mathrm{d}t}\delta q_i(t) \tag{72}$$

Thus, their increments $\delta\dot{q}_i(t)$ are simply:

$$\delta\dot{q}_i(t) = \frac{\mathrm{d}}{\mathrm{d}t}\delta q_i(t) \tag{73}$$

We now calculate the variation of the action in going from the path Γ to the path Γ':

$$\begin{aligned}
\delta S &= \int_{t_1}^{t_2} \mathrm{d}t\, \delta\mathcal{L} \\
&= \int_{t_1}^{t_2} \mathrm{d}t\, \left[\sum_i \frac{\partial\mathcal{L}}{\partial q_i}\delta q_i + \sum_i \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\delta\dot{q}_i\right] \\
&= \int_{t_1}^{t_2} \mathrm{d}t\, \left[\sum_i \frac{\partial\mathcal{L}}{\partial q_i}\delta q_i + \sum_i \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\frac{\mathrm{d}}{\mathrm{d}t}\delta q_i\right]
\end{aligned} \tag{74}$$

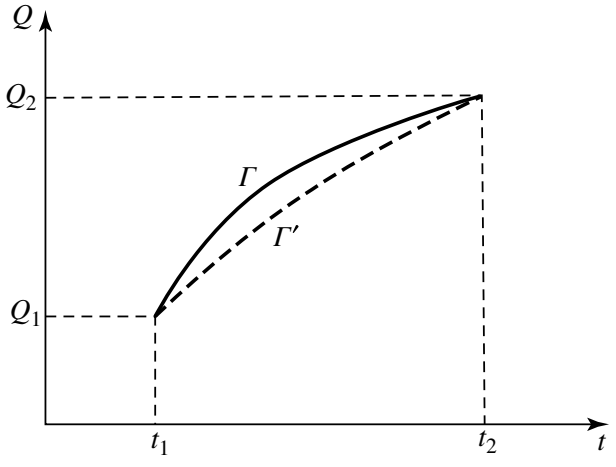


Figure 3: Two paths in space-time which pass through the points (Q_1, t_1) and (Q_2, t_2): the solid-line curve is the path associated with the real motion of the system, and the dashed-line curve is another, infinitely close, path.

according to (73). If we integrate the second term by parts, we obtain:

$$\delta S = \left[\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\delta q_i\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \mathrm{d}t \sum_i \delta q_i \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right]$$

$$= \int_{t_1}^{t_2} \mathrm{d}t \sum_i \delta q_i \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right] \tag{75}$$

since the integrated term is zero, because of conditions (71).

If Γ is the path in space-time which is actually followed during the real motion of the system, the increment δS of the action is zero, according to the principle of least action. For this to be so, it is necessary and sufficient that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0 \quad ; \quad i = 1, 2, ..., N \tag{76}$$

It is obvious that this condition is sufficient. It is also necessary, since, if there existed a time interval during which expression (76) were non-zero for a given value k of the index i, the $\delta q_i(t)$ could be chosen so as to make the corresponding increment δS different from zero. (It would suffice, for example, to choose them so as to make the product $\delta q_k \left[\frac{\partial \mathcal{L}}{\partial q_k} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_k}\right]$ always positive or zero). Consequently, the principle of least action is equivalent to Lagrange's equations.

References and suggestions for further reading:

See section 6 of the bibliography, in particular Marion (6.4). Goldstein (6.6), Landau and Lifshitz (6.7).

For a simple presentation of the use of variational principles in physics, see Feynman II (7.2), Chap. 19.

For Lagrangian formalism applied to a classical field, see Bogoliubov and Chirkov (2.15), Chap. I.

BIBLIOGRAPHY OF VOLUMES I and II

1. QUANTUM MECHANICS: GENERAL REFERENCES

A - INTRODUCTORY TEXTS

Quantum Physics

(1.1) E. H. WICHMANN, *Berkeley Physics Course, Vol. 4: Quantum Physics*, McGraw-Hill, New York (1971).

(1.2) R. P. FEYNMAN, R. B. LEIGHTON and M. SANDS, *The Feynman Lectures on Physics, Vol. III: Quantum Mechanics*, Addison-Wesley, Reading, Mass. (1965).

(1.3) R. EISBERG and R. RESNICK, *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particules*, Wiley, New York (1974).

(1.4) M. ALONSO and E. J. FINN, *Fundamental University Physics, Vol. III: Quantum and Statistical Physics*, Addison Wesley, Reading, Mass. (1968).

(1.5) U. FANO and L. FANO, *Basic Physics of Atoms and Molecules*, Wiley, New York (1959).

(1.6) J. C. SLATER, *Quantum Theory of Matter*, McGraw-Hill, New York (1968).

Quantum mechanics

(1.7) S. BOROWITZ, *Fundamentals of Quantum Mechanics*, Benjamin, New York (1967).

(1.8) S. I. TOMONAGA, *Quantum Mechanics, Vol. I: Old Quantum Theory*, North Holland, Amsterdam (1962).

(1.9) L. PAULING and E. B. WILSON JR., *Introduction to Quantum Mechanics*, McGraw-Hill, New York (1935).

(1.10) Y. AYANT et E. BELORIZKY, *Cours de Mécanique Quantique*, Dunod, Paris (1969).

(1.11) P. T. MATTHEWS, *Introduction to Quantum Mechanics*, McGraw-Hill, New York (1963).

(1.12) J. AVERY, *The Quantum Theory of Atoms, Molecules and Photons*, McGraw-Hill, London (1972).

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

B - MORE ADVANCED TEXTS:

(1.13) P. A. M. DIRAC, *The Principles of Quantum Mechanics*, Oxford University Press (1958).

(1.14) R. H. DICKE and J. P. WITTKE, *Introduction to Quantum Mechanics*, Addison-Wesley, Reading, Mass. (1966).

(1.15) D. I. BLOKHINTSEV, *Quantum Mechanics*, D. Reidel, Dordrecht (1964).

(1.16) E. MERZBACIIER, *Quantum Mechanics,* Wiley, New York (1970).

(1.17) A. MESSIAH, *Mécanique Quantique*, Vols 1 et 2, Dunod, Paris (1964).

(1.18) L. I. SCHIFF, *Quantum Mechanics*, McGraw-Hill, New York (1968).

(1.19) L. D. LANDAU and E. M. LIFSHITZ, *Quantum Mechanics, Nonrelativistic Theory,* Pergamon Press, Oxford (1965).

(1.20) A. S. DAVYDOV, *Quantum Mechanics*, Translated, edited and with additions by D. Ter HAAR, Pergamon Press, Oxford (1965).

(1.21) H. A. BETHE and R. W. JACKIW, *Intermediate Quantum Mechanics*, Benjamin, New York (1968).

(1.22) H. A. KRAMERS, *Quantum Mechanics*, North Holland, Amsterdam (1958).

C - PROBLEMS IN QUANTUM MECHANICS

(1.23) *Selected Problems in Quantum Mechanics*, Collected and edited by D. Ter HAAR, Infosearch, London (1964).

(1.24) S. FLÜGGE, *Practical Quantum Mechanics*, I and II, Springer-Verlag, Berlin (1971).

D - ARTICLES

(1.25) E. SCHRÖDINGER, "What is Matter?", *Scientific American* **189**, 52 (Sept. 1953).

(1.26) G. GAMOW, "The Principle of Uncertainty", *Scientific American* **198**, 51 (Jan. 1958).

(1.27) G. GAMOW, "The Exclusion Principle", *Scientific American* **201**, 74 (July 1959).

(1.28) M. BORN and W. BIEM, "Dualism in Quantum Theory", *Physics Today* **21**, p. 51 (Aug. 1968).

(1.29) W. E. LAMB JR., "An Operational Interpretation of Nonrelativistic Quantum Mechanics", *Physics Today* **22**, 23 (April 1969).

(1.30) M. O. SCULLY and M. SARGENT III, "The Concept of the Photon", *Physics Today* **25**, 38 (March 1972).

(1.31) A. EINSTEIN, "Zur Quantentheorie der Strahlung", *Physik. Z.* **18,** 121 (1917).

(1.32) A. GOLDBERG, H. M.. SCHEY and J. L. SCHWARTZ, "Computer-Generated Motion Pictures of One-Dimensional Quantum-Mechanical Transmission and Reflection Phenomena", *Am. J. Phys.*, **35**, 177 (1967).

(1.33) R. P. FEYNMAN, F. L. VERNON JR. and R. W. HELLWARTH, "Geometrical Representation of the Schrödinger Equation for Solving Maser Problems", *J. Appl. Phys.* **28**, 49 (1957).

(1.34) A. A. VUYLSTEKE, "Maser States in Ammonia-Inversion", *Am. J. Phys.* **27**, 554 (195

2. QUANTUM MECHANICS: MORE SPECIALIZED REFERENCES

A - COLLISIONS

(2.1) T. Y. WU and T. OHMURA, *Quantum Theory of Scattering*, Prentice Hall, Englewood Cliffs (1962).

(2.2) R. G. NEWTON, *Scattering Theory of Waves and Particles*, McGraw-Hill, New York (1966).

(2.3) P. ROMAN, *Advanced Quantum Theory*, Addison-Wesley, Reading, Mass. (1965).

(2.4) M. L. GOLDBERGER and K. M. WATSON, *Collision Theory*, Wiley, New York (1964).

(2.5) N. F. MOTT and H. S. W. MASSEY, *The Theory of Atomic Collisions*, Oxford University Press (1965).

B - RELATIVISTIC QUANTUM MECHANICS

(2.6) J. D. BJORKEN and S. D. DRELL, *Relativistic Quantum Mechanics*, McGraw-Hill, New York (1964).

(2.7) J. J. SAKURAI, *Advanced Quantum Mechanics*, Addison-Wesley, Reading, Mass. (1967).

(2.8) V. B. BERESTETSKII, E. M. LIFSHITZ and L. P. PITAEVSKII, *Relativistic Quantum Theory*, Pergamon Press, Oxford (1971).

C - FIELD THEORY. QUANTUM ELECTRODYNAMICS

(2.9) F. MANDL, *Introduction to Quantum Field Theory*, Wiley Interscience, New York (1959).

(2.10) J. D. BJORKEN and S. D. DRELL, *Relativistic Quantum Fields*, McGraw-Hill, New York (1965).

(2.11) E. A. POWER, *Introductory Quantum Electrodynamics,* Longmans, London (1964).

(2.12) R. P. FEYNMAN, *Quantum Electrodynamics*, Benjamin, New York (1961).

(2.13) W. HEITLER, *The Quantum Theory of Radiation*, Clarendon Press, Oxford (1954).

(2.14) A. I. AKHIEZER and V. B. BERESTETSKII, *Quantum Electrodynamics*, Wiley Interscience, New York (1965).

(2.15) N. N. BOGOLIUBOV and D. V. SHIRKOV, *Introduction to the Theory of Quantized Fields*, Interscience Publishers, New York (1959).

(2.16) S. S. SCHWEBER, *An Introduction to Relativistic Quantum Field Theory*, Harper and Row, New York (1961).

(2.17) M. M. STERNHEIM, "Resource Letter TQE-1 : Tests of Quantum Electrodynamics", *Am. J. Phys.* **40,** 1363 (1972).

D - ROTATIONS AND GROUP THEORY

(2.18) P. H. E. MEIJER and E. BAUER, *Group Theory*, North Holland, Amsterdam (1962).

(2.19) M. E. ROSE, *Elementary Theory of Angular Momentum*, Wiley, New York (1957).

(2.20) M. E. ROSE, *Multipole Fields*, Wiley, New York (1955).

(2.21) A. R. EDMONDS, *Angular Momentum in Quantum Mechanics*, Princeton University Press (1957).

(2.22) M. TINKHAM, *Group Theory and Quantum Mechanics*, McGraw-Hill, New York (1964).

(2.23) E. P. WIGNER, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York (1959).

(2.24) D. PARK, "Resource Letter SP-I on Symmetry in Physics", *Am. J. Phys.* **36**, 577 (1968).

E - MISCELLANEOUS

(2.25) R. P. FEYNMAN and A. R. HIBBS, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York (1965).

(2.26) J. M. ZIMAN, *Elements of Advanced Quantum Theory*, Cambridge University Press (1969).

(2.27) F. A. KAEMPFFER, *Concepts in Quantum Mechanics*, Academic Press, New York (1965).

F - ARTICLES

(2.28) P. MORRISON, "The Overthrow of Parity", *Scientific American* **196,** 45 (April 1957).

(2.29) G. FEINBERG and M. GOLDHABER, "The Conservation Laws of Physics", *Scientific American* **209**, 36 (Oct. 1963).

(2.30) E. P. WIGNER, "Violations of Symmetry in Physics", *Scientific American* **213**, 28 (Dec. 1965).

(2.31) U. FANO, "Description of States in Quantum Mechanics by Density Matrix and Operator Techniques", *Rev. Mod. Phys.* **29,** 74 (1957).

(2.32) D. Ter HAAR, "Theory and Applications of the Density Matrix", *Rept. Progr. Phys.* **24**, 304 (1961).

(2.33) V. F. WEISSKOPF and E. WIGNER, "Berechnung der Natürlichen Linienbreite auf Grund der Diracschen Lichttheorie", *Z. Physik* **63,** 54 (1930).

(2.34) A. DALGARNO and J. T. LEWIS, "The Exact Calculation of Long-Range Forces between Atoms by Perturbation Theory", *Proc. Roy. Soc.* **A 233**, 70 (1955).

(2.35) A. DALGARNO and A. L. STEWART, "On the Perturbation Theory of Small Disturbances", *Proc. Roy. Soc.* **A 238**, 269 (1957).

(2.36) C. SCHWARTZ, "Calculations in Schrödinger Perturbation Theory", *Annals of Physics* (New York), **6**, 156 (1959).

(2.37) J. O. HIRSCHFELDER, W. BYERS BROWN and S. T. EPSTEIN, "Recent Developments in Perturbation Theory", in *Advances in Quantum Chemistry*, P. O. LOWDIN ed., Vol. I, Academic Press, New York (1964).

(2.38) R. P. FEYNMAN, "Space Time Approach to Nonrelativistic Quantum Mechanics", *Rev. Mod. Phys.,* **20**, 367 (1948).

(2.39) L. VAN HOVE, "Correlations in Space and Time and Born Approximation Scattering in Systems of Interacting Particles", *Phys. Rev.* **95,** 249 (1954).

3. QUANTUM MECHANICS: FUNDAMENTAL EXPERIMENTS

Interference effects with weak light:

(3.1) G. I. TAYLOR, "Interference Fringes with Feeble Light", *Proc. Camb. Phil. Soc.* **15,** 114 (1909).

(3.2) G. T. REYNOLDS, K. SPARTALIAN and D. B. SCARL, "Interference Effects Produced by Single Photons", *Nuovo Cimento* **61 B,** 355 (1969).

Experimental verification of Einstein's law for the photoelectric effect; measurements of h

(3.3) A. L. HUGHES, "On the Emission Velocities of Photoelectrons", *Phil. Trans. Roy. Soc.* **212**, 205 (1912).

(3.4) R. A. MILLIKAN, "A Direct Photoelectric Determination of Planck's h", *Phys. Rev.* **7** 355 (1916).

The Franck-Hertz experiment:

(3.5) J. FRANCK und G. HERTZ, "Über Zusammenstösse zwischen Elecktronen und den Molekullen des Quecksilberdampfes und die Ionisierungsspannung desselben", *Verhandlungen der Deutschen Physikalischen Gesellschaft*, **16**, 457 (1914).

"Über Kinetik von Elektronen und Ionen in Gasen", *Physikalische Zeitschrift* **17**, 409 (1916).

The proportionality between the magnetic moment and the angular momentum:

(3.6) A. EINSTEIN und J. W. DE HAAS, "Experimenteller Nachweis der Ampereschen Molekularströme", *Verhandlungen der Deutschen Physikalischen Gesellschaft* **17**, 152 (1915).

(3.7) E. BECK, "Zum Experimentellen Nachweis der Ampereschen Molekularströme", *Annalen der Physik* (Leipzig) **60**, 109 (1919).

The Stern-Gerlach experiment:

(3.8) W. GERLACH und O. STERN, "Der Experimentelle Nachweis der Richtungsquantelung im Magnetfeld", *Zeitschrift für Physik* **9**, 349 (1922).

The Compton effect:

(3.9) A. H. COMPTON, "A Quantum Theory of the Scattering of X-Rays by Light Elements", *Phys. Rev.*, **21**, 483 (1923).

"Wavelength Measurements of Scattered X-Rays", *Phys. Rev.*, **21**, 715 (1923).

Electron diffration:

(3.10) C. DAVISSON and L. H. GERMER, "Diffraction of Electrons by a Crystal of Nickel", *Phys. Rev.* **30**, 705 (1927).

The Lamb shift:

(3.11) W. E. LAMB JR. and R. C. RETHERFORD, "Fine Structure of the Hydrogen Atom",

I - *Phys. Rev.* **79**, 549 (1950),

II - *Phys. Rev.* **81**, 222 (1951).

Hyperfine structure of the hydrogen atom:

(3.12) S. B. CRAMPTON, D. KLEPPNER and N. F. RAMSEY, "Hyperfine Separation of Ground State Atomic Hydrogen", *Phys. Rev. Letters* **11**, 338 (1963).

Some fundamental experiments are described in:

(3.13) O. R. FRISCH, "Molecular Beams", *Scientific American* **212**, 58 (May 1965).

4. QUANTUM MECHANICS: HISTORY

(4.1) L. DE BROGLIE, "Recherches sur la Théorie des Quanta", *Annales de Physique*(Paris), **3**, 22 (1925).

(4.2) N. BOHR, "The Solvay Meetings and the Development of Quantum Mechanics", *Essays 1958-1962 on Atomic Physics and Human Knowledge*, Vintage, New York (1966).

(4.3) W. HEISENBERG, *Physics and Beyond: Encounters and Conversations*, Harper and Row, New York (1971).

La Partie et le Tout, Albin Michel, Paris (1972).

(4.4) *Niels Bohr, His life and work as seen by his friends and colleagues*, S. ROZENTAL, ed., North Holland, Amsterdam (1967).

(4.5) A. EINSTEIN, M. and H. BORN, *Correspondance 1916-1955*, Editions du Seuil, Paris (1972). See also *La Recherche,* 3, 137 (fev. 1972).

(4.6) *Theoretical Physics in the Twentieth Century,* M. FIERZ and V. F. WEISSKOPF eds., Wiley Interscience, New York (1960).

(4.7) *Sources of Quantum Mechanics*, B. L. VAN DER WAERDEN ed., North Holland, Amsterdam (1967); Dover, New York (1968).

(4.8) M. JAMMER, *The Conceptual Development of Quantum Mechanics*, McGraw-Hill, New York (1966). This book traces the historical development of quantum mechanics. Its very numerous footnotes provide a multitude of references. See also (5.13).

ARTICLES

(4.9) K. K. DARROW, "The Quantum Theory", *Scientific American* **186,** 47 (March 1952).

(4.10) M. J. KLEIN, "Thermodynamics and Quanta in Planck's work", *Physics Today* **19,** 23 (Nov. 1966).

(4.11) H. A. MEDICUS, "Fifty Years of Matter Waves", *Physics Today* **27**, 38 (Feb. 1974).

Reference (5.12) contains a large number of references to the original articles.

5. QUANTUM MECHANICS: DISCUSSION OF ITS FOUNDATIONS

A - GENERAL PROBLEMS

(5.1) D. BOHM, *Quantum Theory*, Constable, London (1954).

(5.2) J. M. JAUCH, *Foundations of Quantum Mechanics*, Addison-Wesley, Reading, Mass. (1968).

(5.3) B. D'ESPAGNAT, *Conceptual Foundations of Quantum Mechanics*, Benjamin, New York (1971); *Conceptions de la Physique Contemporaine. Les Interprétations de la Mécanique Quantique et de la Mesure*, Hermann, Paris (1965).

(5.4) Proceedings of the International School of Physics "Enrico Fermi" (Varenna), Course IL; *Foundations of Quantum Mechanics*, B. D'ESPAGNAT ed., Academic Press, New York (1971).

(5.5) B. S. DEWITT, "Quantum Mechanics and Reality", *Physics Today* **23**, 30, (Sept. 1970).

(5.6) "Quantum Mechanics debate", *Physics Today* **24**, 36 (April 1971). See also (1.28).

(5.7) F. LALOË, *Do we really understand quantum mechanics?*, Cambridge University Press, (second edition 2019).

See also (1.28).

B - MISCELLANEOUS INTERPRETATIONS

(5.8) N. BOHR, "Discussion with Einstein on Epistemological Problems in Atomic Physics", in *A. Einstein: Philosopher-Scientist*, P. A. SCHILPP ed., Harper and Row, New York (1959).

(5.9) M. BORN, *Natural Philosophy of Cause and Chance*, Oxford University Press, London (1951); Clarendon Press, Oxford (1949).

(5.10) L. DE BROGLIE, *Une Tentative d'Interprétation Causale et Non Linéaire de la Mécanique Ondulatoire: la Théorie de la Double Solution*, Gauthier-Villars, Paris (1956); *Etude Critique des Bases de l'Interprétation Actuelle de la Mécanique Ondulatoire*, Gauthier-Villars, Paris (1963).

(5.11) *The Many-Worlds Interpretation of Quantum Mechanics*, B. S. DEWITT and N. GRAHAM eds., Princeton University Press (1973).

A very complete set of references with comments may be found in:

(5.12) B. S. DEWITT and R. N. GRAHAM, "Resource Letter IQM-1 on the Interpretation of Quantum Mechanics", *Am. J. Phys.* **39,** 724 (1971).

(5.13) M. JAMMER, *The Philosophy of Quantum Mechanics*, Wiley-interscience, New York (1974). General presentation of the different interpretations and formalisms of quantum mechanics. Contains many references.

C - MEASUREMENT THEORY

(5.14) K. GOTTFRIED, *Quantum Mechanics*, Vol. I, Benjamin, New York (1966).

(5.15) D. I. BLOKIIINTSEV, *Principes Essentiels de la Mécanique Quantique*, Dunod, Paris (1968).

(5.16) A. SHIMONY, "Role of the Observer in Quantum Theory", *Am. J. Phys.,* **31**, 755 (1963). See also (5.13), Chap. 11.

D - HIDDEN VARIABLES AND "PARADOXES"

(5.17) A. EINSTEIN, B. PODOLSKY and N. ROSEN, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", *Phys. Rev.* **47**, 777 (1935).

N. BOHR, "Can Quantum Mechanical Description of Physical Reality Be Considered Complete?", *Phys. Rev.* **48**, 696 (1935).

(5.18) *Paradigms and Paradoxes, the Philosophical Challenge of the Quantum Domain*, R. G. COLODNY ed., University of Pittsburgh Press (1972).

(5.19) J. S. BELL, "On the Problem of Hidden Variables in Quantum Mechanics", *Rev. Mod. Phys.* **38**, 447 (1966).

See also Ref. (4.8), as well as (5.12) and Chap. 7 of (5.13).

6. CLASSICAL MECHANICS

A - INTRODUCTORY LEVEL

(6.1) M. ALONSO and E. J. FINN, *Fundamental University Physics, Vol. I: Mechanics*, Addison-Wesley, Reading, Mass. (1967).

(6.2) C. KITTEL, W. D. KNIGHT and M. A. RUDERMAN, *Berkeley Physics Course, Vol. 1: Mechanics*, McGraw-Hill, New York (1962).

(6.3) R. P. FEYNMAN, R. B. LEIGHTON and M. SANDS, *The Feynman Lectures on Physics, Vol. I: Mechanics, Radiation, and Heat*, Addison-Wesley, Reading, Mass. (1966).

(6.4) J. B. MARION, *Classical Dynamics of Particles and Systems*, Academic Press, New York (1965).

B - MORE ADVANCED LEVEL

(6.5) A. SOMMERFELD, *Lectures on Theoretical Physics, Vol. I: Mechanics*, Academic Press, New York (1964).

(6.6) H. GOLDSTEIN, *Classical Mechanics*, Addison-Wesley, Reading, Mass. (1959).

(6.7) L. D. LANDAU and E. M. LIFSHITZ, *Mechanics*, Pergamon Press, Oxford (1960); *Mécanique*, 3e éd., Ed. Mir, Moscou (1969).

7. ELECTROMAGNETISM AND OPTICS

A - INTRODUCTORY LEVEL

(7.1) E. M. PURCELL, *Berkeley Physics Course, Vol. 2: Electricity and Magnetism*, McGraw-Hill, New York (1965).

F. S. CRAWFORD JR., *Berkeley Physics Course, Vol. 3: Waves*, McGraw-Hill, New York (1968).

(7.2) R. P. FEYNMAN, R. B. LEIGHTON and M. SANDS, *The Feynman Lectures on Physics, Vol. II: Electromagnetism and Matter*, Addison-Wesley, Reading, Mass. (1966).

(7.3) M. ALONSO and E. J. FINN, *Fundamental University Physics, Vol. II: Fields and Waves*, Addison-Wesley, Reading, Mass. (1967).

(7.4) E. HECHT and A. ZAJAC, *Optics*, Addison-Wesley, Reading, Mass. (1974).

B - MORE ADVANCED LEVEL

(7.5) J. D. JACKSON, *Classical Electrodynamics*, 2e edition Wiley, New York (1975).

(7.6) W. X. H. PANOFSKY and M. PHILLIPS, *Classical Electricity and Magnetism,* Addison-Wesley, Reading, Mass. (1964).

(7.7) J. A. STRATTON, *Electromagnetic Theory*, McGraw-Hill, New York (1941).

(7.8) M. BORN and E. WOLF, *Principles of Optics*, Pergamon Press, London (1964).

(7.9) A. SOMMERFELD, *Lectures on Theoretical Physics, Vol. IV: Optics*, Academic Press, New York (1964).

(7.10) G. BRUHAT, *Optique*, 5e completed by A. KASTLER, Masson, Paris (1954).

(7.11) L. LANDAU and E. LIFSHITZ, *The Classical Theory of Fields*, Addison-Wesley, Reading, Mass. (1951); Pergamon Press, London (1951); *Théorie du champ*, 2e éd., Ed. Mir, Moscou (1966).

(7.12) L. D. LANDAU and E. M. LIFSHITZ, *Electrodynamics of Continuous Media*, Pergamon Press, Oxford (1960).

(7.13) L. BRILLOUIN, *Wave Propagation and Group Velocity*, Academic Press, New York (1960).

8. THERMODYNAMICS. STATISTICAL MECHANICS

A - INTRODUCTORY LEVEL

(8.1) F. REIF, *Berkeley Physics Course, Vol. 5: Statistical Physics*, McGraw-Hill, New York (1967).

(8.2) C. KITTEL, *Thermal Physics*, Wiley, New York (1969).

(8.3) G. BRUHAT, *Thermodynamique*, 5^e édition remaniée par A. KASTLER, Masson, Paris (1962). See also (1.4), part 2, and (6.3).

B - MORE ADVANCED LEVEL

(8.4) F. REIF, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York (1965).

(8.5) R. CASTAING, *Thermodynamique Statistique*, Masson, Paris (1970).

(8.6) P. M. MORSE, *Thermal Physics*, Benjamin, New York (1964).

(8.7) R. KUBO, *Statistical Mechanics*, North Holland, Amsterdam and Wiley, New York (1965).

(8.8) L. D. LANDAU and E. M. LIFSHITZ, *Course of Theoretical Physics, Vol. 5: Statistical Physics*, Pergamon Press, London (1963).

(8.9) H. B. CALLEN, *Thermodynamics*, Wiley, New York (1961).

(8.10) A. B. PIPPARD, *The Elements of Classical Thermodynamics*, Cambridge University Press (1957).

(8.11) R. C. TOLMAN, *The Principles of Statistical Mechanics*, Oxford University Press (1950).

9. RELATIVITY

A - INTRODUCTORY LEVEL

(9.1) J. H. SMITH, *Introduction to Special Relativity*, Benjamin, New York (1965). See also references (6.2) and (6.3).

B - MORE ADVANCED LEVEL

(9.2) J. L. SYNGE, *Relativity: The Special Theory*, North Holland, Amsterdam (1965).

(9.3) R. D. SARD, *Relativistic Mechanics*, Benjamin, New York (1970).

(9.4) J. AHARONI, *The Special Theory of Relativity*, Oxford University Press, London (1959).

(9.5) C. MØLLER, *The Theory of Relativity*, Oxford University Press, London (1972).

(9.6) P. G. BERGMANN, *Introduction to the Theory of Relativity*, Prentice Hall, Englewood Cliffs (1960).

(9.7) C. W. MISNER, K. S. THORNE and J. A. WHEELER, *Gravitation*, Freeman, San Francisco (1973).

See also references on Electromagnetism, in particular (7.5) and (7.11).

Other useful references are :

(9.8) A. EINSTEIN, *Quatre Conférences sur la Théorie de la Relativité*, Gauthier-Villars, Paris (1971).

(9.9) A. EINSTEIN, *La Théorie de la Relativité Restreinte et Générale. La Relativité et le Problème de l'Espace*, Gauthier-Villars, Paris (1971).

(9.10) A. EINSTEIN, *The Meaning of Relativity*, Methuen, London (1950).

(9.11) A. EINSTEIN, *Relativity, the Special and General Theory, a Popular Exposition*, Methuen, London (1920); H. Holt, New York (1967).

A much more complete set of references can be found in:

(9.12) G. HOLTON, Resource Letter SRT-1 on Special Relativity Theory, *Am. J. Phys.* **30,** 462 (1962).

10. MATHEMATICAL METHODS

A - ELEMENTARY GENERAL TEXTS

(10.1) J. BASS, *Cours de Mathématiques*, Vols. I, II and III, Masson, Paris (1961).

(10.2) A. ANGOT, *Compléments de Mathématiques*, Revue d'Optique, Paris (1961).

(10.3) T. A. BAK and J. LICHTENBERG, *Mathematics for Scientists*, Benjamin, New York (1966).

(10.4) G. ARFKEN, *Mathematical Methods for Physicists*, Academic Press, New York (1966).

(10.5) J. D. JACKSON, *Mathematics for Quantum Mechanics*, Benjamin, New York (1962).

B - MORE ADVANCED GENERAL TEXTS

(10.6) J. MATHEWS and R. L. WALKER, *Mathematical Methods of Physics*, Benjamin, New York (1970).

(10.7) L. SCHWARTZ, *Méthodes Mathématiques pour les Sciences Physiques*, Hermann, Paris (1965). *Mathematics for the Physical Sciences,* Hermann, Paris (1968).

(10.8) E. BUTKOV, *Mathematical Physics,* Addison-Wesley, Reading, Mass (1968).

(10.9) H. CARTAN, *Théorie Elémentaire des Fonctions Analytiques d'une ou Plusieurs Variables Complexes*, Hermann, Paris (1961). *Elementary Theory of Analytic Functions of One or Several Complex Variables*, Addison-Wesley, Reading, Mass. (1966).

(10.10) J. VON NEUMANN, *Mathematical Foundations of Quantum Mechanics,* Princeton University Press (1955).

(10.11) R. COURANT and D. HILBERT, *Methods of Mathematical Physics*, Vols. I and II, Wiley, Interscience, New York (1966).

(10.12) E. T. WHITTAKER and G. N. WATSON, *A Course of Modern Analysis*, Cambridge University Press (1965).

(10.13) P. M. MORSE and H. FESHBACH, *Methods of Theoretical Physics*, McGraw-Hill, New York (1953).

C - LINEAR ALGEBRA. HILBERT SPACES

(10.14) A. C. AITKEN, *Determinants and Matrices*, Oliver and Boyd, Edinburgh (1956).

(10.15) R. K. EISENSCHITZ, *Matrix Algebra for Physicists*, Plenum Press, New York (1966).

(10.16) M. C. PEASE III, *Methods of Matrix Algebra*, Academic Press, New York (1965).

(10.17) J. L. SOULE, *Linear Operators in Hilbert Space*, Gordon and Breach, New York (1967).

(10.18) W. SCHMEIDLER, *Linear Operators in Hilbert Space*, Academic Press, New York (1965).

(10.19) N. I. AKHIEZER and I. M. GLAZMAN, *Theory of Linear Operators in Hilbert Space*, Ungar, New York (1961).

D - FOURIER TRANSFORMS. DISTRIBUTIONS

(10.20) R. STUART, *Introduction to Fourier Analysis*, Chapman and Hall, London (1969).

(10.21) M. J. LIGHTHILL, *Introduction to Fourier Analysis and Generalized Functions*, Cambridge University Press (1964).

(10.22) L. SCHWARTZ, *Théorie des Distributions*, Hermann, Paris (1967).

(10.23) I. M. GEL'FAND and G. E. SHILOV, *Generalized Functions*, Academic Press, New York (1964).

(10.24) F. OBERHETTINGER, *Tabellen zur Fourier Transformation*, Springer-Verlag, Berlin (1957).

E - PROBABILITY AND STATISTICS

(10.25) *Elements of Probability Theory*, Academic Press, New York (1966).

(10.26) P. G. HOEL, S. C. PORT and C. J. STONE, *Introduction to Probability Theory*, Houghton- Mifflin, Boston (1971).

(10.27) H. G. TUCKER, *An Introduction to Probability and Mathematical Statistics*, Academic Press, New York (1965).

(10.28) J. LAMPERTI, *Probability*, Benjamin, New York (1966).

(10.29) W. FELLER, *An Introduction to Probability Theory and its Applications*, Wiley, New York (1968).

(10.30) L. BREIMAN, *Probability*, Addison-Wesley, Reading, Mass. (1968).

F - GROUP THEORY

Applied to physics:

(10.31) H. BACRY, *Lectures on Group Theory*, Gordon and Breach, New York (1967).

(10.32) M. HAMERMESH, *Group Theory and its Application to Physical Problems*, Addison-Wesley, Reading, Mass. (1962).

See also (2.18), (2.22), (2.23) or reference (16.13), which provides a simple introduction to continuous groups in physics.

More mathematical:

(10.33) G. PAPY, *Groupes,* Presses Universitaires de Bruxelles, Bruxelles (1961); *Groups*, Macmillan, New York (1964).

(10.34) A. G. KUROSH, *The Theory of Groups*, Chelsea, New York (1960).

(10.35) L. S. PONTRYAGIN, *Topological Groups*, Gordon and Breach, New York (1966).

G - SPECIAL FUNCTIONS AND TABLES

(10.36) A. GRAY and G. B. MATHEWS, *A Treatise on Bessel Functions and their Applications to Physics*, Dover, New York (1966).

(10.37) E. D. RAINVILLE, *Special Functions*, Macmillan, New York (1965).

(10.38) W. MAGNUS, F. OBERHETTINGER and R. P. SONI, *Formulas and Theorems for the Special Functions of Mathematical Physics*, Springer-Verlag, Berlin (1966).

(10.39) BATEMAN MANUSCRIPT PROJECT, *Higher Transcendental Functions*, Vols. I, II and III, A. ERDELYI ed., McGraw-Hill, New York (1953).

(10.40) M. ABRAMOWITZ and I. A. STEGUN, *Handbook of Mathematical Functions*, Dover, New York (1965).

(10.41) L. J. COMRIE, *Chambers's Shorter Six-Figure Mathematical Tables*, Chambers, London (1966).

(10.42) E. JAHNKE and F. EMDE, *Tables of Functions,* Dover, New York (1945).

(10.43) V. S. AIZENSHTADT, V. I. KRYLOV and A. S. METEL'SKII, *Tables of Laguerre Polynomials and Functions*, Pergamon Press, Oxford (1966).

(10.44) H. B. DWIGHT, *Tables of Integrals and Other Mathematical Data*, Macmillan, New York (1965).

(10.45) D. BIERENS DE HAAN, *Nouvelles Tables d'Intégrales Définies*, Hafner, New York (1957).

(10.46) F. OBERHETTINGER and L. BADII, *Tables of Laplace Transforms*, Springer-Verlag, Berlin (1973).

(10.47) BATEMAN MANUSCRIPT PROJECT, *Tables of Integral Transforms, Vols.* I and II, A. ERDELYI ed., McGraw-Hill, New York (1954).

(10.48) M. ROTENBERG, R. BIVINS, N. METROPOLIS and J. K. WOOTEN JR., *The 3-j and 6-j symbols*, M.I.T. Technology Press (1959); Crosby Lockwood and Sons, London.

11. ATOMIC PHYSICS

A - INTRODUCTORY LEVEL

(11.1) H. G. KUHN, *Atomic Spectra*, Longman, London (1969).

(11.2) B. CAGNAC and J. C. PEBAY-PEYROULA, *Modern Atomic Physics*, Vol. 1 : *Fundamental Principles,* and 2 : *Quantum Theory and its Application*, Macmillan, London (1975).

(11.3) A. G. MITCHELL and M. W. ZEMANSKY, *Resonance Radiation and Excited Atoms*, Cambridge University Press, London (1961).

(11.4) M. BORN, *Atomic Physics*, Blackie and Son, London (1951).

(11.5) H. E. WHITE, *Introduction to Atomic Spectra*, McGraw-Hill, New York (1934).

(11.6) V. N. KONDRATIEV, *La Structure des Atomes et des Molécules*, Masson, Paris (1964). See also (1.3) and (12.1).

B - MORE ADVANCED LEVEL

(11.7) G. W. SERIES, *The Spectrum of Atomic Hydrogen*, Oxford University Press, London (1957).

(11.8) J. C. SLATER, *Quantum Theory of Atomic Structure*, Vols. I and II, McGraw-Hill, New York (1960).

(11.9) A. E. RUARK and H. C. UREY, *Atoms, Molecules and Quanta*, Vols. I and II, Dover, New York (1964).

(11.10) *Handbuch der Physik, Vols. XXXV and XXXVI, Atoms*, S. FLÜGGE ed., Springer-Verlag Berlin (1956 and 1957).

(11.11) N. F. RAMSEY, *Molecular Beams*, Oxford University Press, London (1956).

(11.12) I. I. SOBEL'MAN, *Introduction to the Theory of Atomic Spectra*, Pergamon Press, Oxford (1972).

(11.13) E. U. CONDON and G. H. SHORTLEY, *The Theory of Atomic Spectra*, Cambridge University Press (1953).

C - ARTICLES

Many references to articles and books, with comments, can be found in:

(11.14) J. C. ZORN, "Resource Letter MB-1 on Experiments with Molecular Beams", *Am. J. Phys.* **32**, 721 (1964).

See also (3.13).

(11.15) V. F. WEISSKOPF, "How Light Interacts with Matter", *Scientific American*, **219**, 60 (Sept. 1968).

(11.16) H. R. CRANE, "The g Factor of the Electron", *Scientific American* **218**, 72 (Jan. 1968).

(11.17) M. S. ROBERTS, "Hydrogen in Galaxies", *Scientific American* **208**, 94 (June 1963).

(11.18) S. A. WERNER, R. COLELLA, A. W. OVERHAUSER and C. F. EAGEN, "Observation of the Phase Shift of a Neutron due to Precession in a Magnetic Field", *Phys. Rev. Letters* **35**, 1053 (1975).

See also: H. RAUCH, A. ZEILINGER, G. BADUREK A. WILFING, W. BAUPIESS and U. BONSE, *Physics Letters* **54 A**, 425 (1975).

D - EXOTIC ATOMS

(11.19) H. C. CORBEN and S. DE BENEDETTI, "The Ultimate Atom", *Scientific American* **191**, 88 (Dec. 1954).

(11.20) V. W. HUGHES, "The Muonium Atom", *Scientific American* **214**, 93, (April 1966). "Muonium", *Physics Today* **20**, 29 (Dec. 1967).

(11.21) S. DE BENEDETTI, "Mesonic Atoms", *Scientific American* **195**, 93 (Oct. 1956).

(11.22) C. E. WIEGAND, "Exotic Atoms", *Scientific American* **227**, 102 (Nov. 1972).

(11.23) V. W. HUGHES, "Quantum Electrodynamics: Experiment", in *Atomic Physics*, B. Bederson, V. W. Cohen and F. M. Pichanick eds., Plenum Press, New York (1969).

(11.24) R. DE VOE, P. M. Mc INTYRE, A. MAGNON, D. Y. STOWELL, R. A. SWANSON and V. L. TELEGDI, "Measurement of the Muonium Hfs Splitting and of the Muon Moment by Double Resonance, and New Value of α", *Phys. Rev. Letters* **25**, 1779 (1970).

(11.25) K. F. CANTER, A. P. MILLS JR. and S. BERKO, "Observations of Positronium Lyman-Radiation", *Phys. Rev. Letters* **34**, 177 (1975). "Fine-Structure Measurement in the First Excited State of Positronium" *Phys. Rev. Letters* 34, 1541 (1975).

(11.26) V. MEYER et al., "Measurement of the 1s-2s energy interval in muonium", *Phys. Rev. Letters* **84**, 1136 (2000).

12. MOLECULAR PHYSICS

A - INTRODUCTORY LEVEL

(12.1) M. KARPLUS and R. N. PORTER, *Atoms and Molecules*, Benjamin, New York (1970).

(12.2) L. PAULING, *The Nature of the Chemical Bond*, Cornell University Press (1948).

See also (1.3), Chap. 12; (1.5) and (11.6).

B - MORE ADVANCED LEVEL

(12.3) I. N. LEVINE, *Quantum Chemistry*, Allyn and Bacon, Boston (1970).

(12.4) G. HERZBERG, *Molecular Spectra and Molecular Structure*, Vol. I: *Spectra of Diatomic Molecules* and Vol. II: *Infrared and Raman Spectra of Polyatomic Molecules*, D. Van Nostrand Company, Princeton (1963 and 1964).

(12.5) H. EYRING, J. WALTER and G. E. KIMBALL, *Quantum Chemistry*, Wiley, New York (1963).

(12.6) C. A. COULSON, *Valence*, Oxford at the Clarendon Press (1952).

(12.7) J. C. SLATER, *Quantum Theory of Molecules and Solids,* Vol. 1 : *Electronic Structure of Molecules,* McGraw-Hill, New York (1963).

(12.8) *Handbuch der Physik, Vol. XXXVII, 1 and 2, Molecules*, S. FLÜGGE, ed., Springer Verlag, Berlin (1961).

(12.9) D. LANGBEIN, *Theory of Van der Waals Attraction*, Springer Tracts in Modern Physics, Vol. 72, Springer Verlag, Berlin (1974).

(12.10) C. H. TOWNES and A. L. SCHAWLOW, *Microwave Spectroscopy*, McGraw-Hill, New York (1955).

(12.11) P. ENCRENAZ, *Les Molécules Interstellaires*, Delachaux et Niestlé, Neuchâtel (1974).

See also (11.9), (11.11) and (11.14).

C - ARTICLES

(12.12) B. V. DERJAGUIN, "The Force Between Molecules", *Scientific American* **203**, 47 (July 1960).

(12.13) A. C. WAHL, "Chemistry by Computer", *Scientific American* **222**, 54 (April 1970).

(12.14) B. E. TURNER, "Interstellar Molecules", *Scientific American* **228**,51 (March 1973).

(12.15) P. M. SOLOMON, "Interstellar Molecules", *Physics Today* **26**, 32 (March 1973).

See also (16.25).

13. SOLID STATE PHYSICS

A - INTRODUCTORY LEVEL

(13.1) C. KITTEL, *Elementary Solid State Physics*, Wiley, New York (1962).

(13.2) C. KITTEL, *Introduction to Solid State Physics*, 3rd ed., Wiley, New York (1966).

(13.3) J. M. ZIMAN, *Principles of the Theory of Solids*, Cambridge University Press, London (1972).

(13.4) F. SEITZ, *Modern Theory of Solids*, McGraw-Hill, New York (1940).

B - MORE ADVANCED LEVEL

General texts:

(13.5) C. KITTEL, *Quantum Theory of Solids*, Wiley, New York (1963).

(13.6) R. E. PEIERLS, *Quantum Theory of Solids*, Oxford University Press, London (1964).

(13.7) N. F. MOTT and H. JONES, *The Theory of the Properties of Metals and Alloys*, Clarendon Press, Oxford (1936); Dover, New York (1958).

More specialized texts:

(13.8) M. BORN and K. HUANG, *Dynamical Theory of Crystal Lattices*, Oxford University Press, London (1954).

(13.9) J. M. ZIMAN, *Electrons and Phonons*, Oxford University Press, London (1960).

(13.10) H. JONES, *The Theory of Brillouin Zones and Electronic States in Crystals*, North Holland, Amsterdam (1962).

(13.11) J. CALLAWAY, *Energy Band Theory*, Academic Press, New York (1964).

(13.12) R. A. SMITH, *Wave Mechanics of Crystalline Solids*, Chapman and Hall, London (1967).

(13.13) D. PINES and P. NOZIÈRES, *The Theory of Quantum Liquids*, Benjamin, New York (1966).

(13.14) D. A. WRIGHT, *Semi-Conductors*, Associated Book Publishers, London (1966).

(13.15) R. A. SMITH, *Semi-Conductors*, Cambridge University Press, London (1964).

C - ARTICLES

(13.16) R. L. SPROULL, "The Conduction of Heat in Solids", *Scientific American* **207**, 92 (Dec. 1962).

(13.17) A. R. MACKINTOSH, "The Fermi Surface of Metals", *Scientific American* **209**, 110 (July 1963).

(13.18) D. N. LANGENBERG, D. J. SCALAPINO and B. N. TAYLOR, "The Josephson Effects", *Scientific American* **214**, 30 (May 1966).

(13.19) G. L. POLLACK, "Solid Noble Gases", *Scientific American* **215**, 64 (Oct. 1966).

(13.20) B. BERTMAN and R. A. GUYER, "Solid Helium", *Scientific American* **217**, 85 (Aug. 1967).

(13.21) N. MOTT, "The Solid State", *Scientific American* **217**, 80 (Sept. 1967).

(13.22) M. Ya. AZBEL', M. I. KAGANOV and I. M. LIFSHITZ, "Conduction Electrons in Metals", *Scientific American* **228**, 88 (Jan. 1973).

(13.23) W. A. HARRISON, "Electrons in Metals", *Physics Today* **22**, 23 (Oct. 1969).

14. MAGNETIC RESONANCE

(14.1) A. ABRAGAM, *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford (1961);

(14.2) C. P. SLICHTER, *Principles of Magnetic Resonance*, Harper and Row, New York (1963).

(14.3) G. E. PAKE, *Paramagnetic Resonance*, Benjamin, New York (1962).

See also Ramsey (11.11), Chaps. V, VI and VII.

ARTICLES

(14.4) G. E. PAKE, "Fundamentals of Nuclear Magnetic Resonance Absorption", I and II, *Am. J. Phys.* **18**, 438 and 473 (1950).

(14.5) E. M. PURCELL, "Nuclear Magnetism", *Am. J. Phys.* **22**, 1 (1954).

(14.6) G. E. PAKE, "Magnetic Resonance", *Scientific American* **199**, 58 (Aug. 1958).

(14.7) K. WÜTHRICH and R. C. SHULMAN, "Magnetic Resonance in Biology", *Physics Today* **23**, 43 (April 1970).

(14.8) F. BLOCH, "Nuclear Induction", *Phys. Rev.* **70**, 460 (1946).

Numerous other references, in particular to the original articles, can be found in:

(14.9) R. E. NORBERG, "Resource Letter NMR-EPR-1 on Nuclear Magnetic Resonance and Electron Paramagnetic Resonance", *Am. J. Phys.* **33**, 71 (1965).

15. QUANTUM OPTICS ; MASERS AND LASERS

A - OPTICAL PUMPING. MASERS AND LASERS

(15.1) R. A. BERNHEIM, *Optical Pumping: An Introduction*, Benjamin, New York (1965). Contient de nombreuses références. De plus, plusieurs articles originaux y sont reproduits.

(15.2) *Quantum Optics and Electronics, Les Houches Lectures 1964*, C. DE WITT, A. BLANDIN and C. COHEN-TANNOUDJI eds., Gordon and Breach, New York (1965).

(15.3) *Quantum Optics, Proceedings of the Scottish Universities Summer School 1969*, S. M. KAY and A. MAITLAND eds., Academic Press, London (1970).

The proceedings of these two summer schools contain several lectures related to optical pumping and quantum electronicss.

(15.4) W. E. LAMB JR., *Quantum Mechanical Amplifiers*, in *Lectures in Theoretical Physics*, Vol. II, W. BRITTIN and D. DOWNS eds., Interscience Publishers, New York (1960).

(15.5) M. SARGENT III, M. O. SCULLY and W. E. LAMB JR., *Laser Physics*, Addison-Wesley, New York (1974).

(15.6) A. E. SIEGMAN, *An Introduction to Lasers and Masers*, McGraw-Hill, New York (1971).

(15.7) L. ALLEN, *Essentials of Lasers*, Pergamon Press, Oxford (1969). This short book contains the reprints of several original articles related to lasers.

(15.8) L. ALLEN and J. H. EBERLY, *Optical Resonance and Two-Level Atoms*, Wiley Interscience, New York (1975).

(15.9) A. YARIV, *Quantum Electronics*, Wiley, New York (1967).

(15.10) H. M. NUSSENZVEIG, *Introduction to Quantum Optics*, Gordon and Breach, London (1973).

B - ARTICLES

Two "Resource Letters" give and comment a large number of references :

(15.11) H. W. MOOS, "Resource Letter MOP-1 on Masers (Microwave through Optical) and on Optical Pumping", *Am. J. Phys.* **32**, 589 (1964).

(15.12) P. CARRUTHERS, "Resource Letter QSL-1 on Quantum and Statistical Aspects of Light", *Am. J. Phys.*, **31**, 321 (1963).

Reprints of many important papers on lasers have been collected in:

(15.13) *Laser Theory*, F. S. BARNES ed., I.E.E.E. Press, New York (1972).

(15.14) H. LYONS, "Atomic Clocks", *Scientific American* **196,** 71 (Feb. 1957).

(15.15) J. P. GORDON, "The Maser", *Scientific American* **199,** 42 (Dec. 1958).

(15.16) A. L. BLOOM, "Optical Pumping", *Scientific American* **203**, 72 (Oct. 1960).

(15.17) A. L. SCHAWLOW, "Optical Masers", *Scientific American* **204,** 52 (June 1961).

"Advances in Optical Masers", *Scientific American* **209,** 34 (July 1963). "Laser Light", *Scientific. American.,* **219,** 120 (Sept. 1968).

(15.18) M. S. FELD and V. S. LETOKHOV, "Laser Spectroscopy", *Scientific American* **229**, 69 (Dec. 1973).

C - NON-LINEAR OPTICS

(15.19) G. C. BALDWIN, *An Introduction to Non-Linear Optics*, Plenum Press, New York (1969).

(15.20) F. ZERNIKE and J. E. MIDWINTER, *Applied Non-Linear Optics*, Wiley Interscience, New York (1973).

(15.21) N. BLOEMBERGEN, *Non-Linear Optics*, Benjamin, New York (1965).

See also the lectures of this author in references (15.2) and (15.3).

D - ARTICLES

(15.22) J. A. GIORDMAINE, "The Interaction of Light with Light", *Scientific American* **210**, 38 (Apr. 1964).

"Non-Linear Optics", *Physics Today* **22**, 39 (Jan. 1969).

16. NUCLEAR PHYSICS AND PARTICLE PHYSICS

A - INTRODUCTION TO NUCLEAR PHYSICS

(16.1) L. VALENTIN, *Physique Subatomique: Noyaux et Particules*, Hermann, Paris (1975).

(16.2) D. HALLIDAY, *Introductory Nuclear Physics*, Wiley, New York (1960).

(16.3) R. D. EVANS, *The Atomic Nucleus*, McGraw-Hill, New York (1955).

(16.4) M. A. PRESTON, *Physics of the Nucleus*, Addison-Wesley, Reading, Mass. (1962).

(16.5) E. SEGRE, *Nuclei and Particles*, Benjamin, New York (1965).

B - MORE ADVANCED NUCLEAR PHYSICS TEXTS

(16.6) A. DESHALIT and H. FESHBACH, *Theoretical Nuclear Physics, Vol. 1: Nuclear Structure*, Wiley, New York (1974).

(16.7) J. M. BLATT and V. F. WEISSKOPF, *Theoretical Nuclear Physics*, Wiley, New York (1963).

(16.8) E. FEENBERG, *Shell Theory of the Nucleus*, Princeton University Press (1955).

(16.9) A. BOHR and B. R. MOTTELSON, *Nuclear Structure*, Benjamin, New York (1969).

C - INTRODUCTION TO PARTICLE PHYSICS

(16.10) D. H. FRISCH and A. M. THORNDIKE, *Elementary Particles*, Van Nostrand, Princeton (1964).

(16.11) C. E. SWARTZ, *The Fundamental Particles*, Addison-Wesley, Reading, Mass. (1965).

(16.12) R. P. FEYNMAN, *Theory of Fundamental Processes*, Benjamin, New York (1962).

(16.13) R. OMNES, *Introduction à l'Etude des Particules Elémentaires*, Ediscience, Paris (1970).

(16.14) K. NISHIJIMA, *Fundamental Particles*, Benjamin, New York (1964).

D - MORE ADVANCED PARTICLE PHYSICS TEXTS

(16.15) B. DIU, *Qu'est-ce qu'une Particule Elémentaire?* Masson, Paris (1965).

(16.16) J. J. SAKURAI, *Invariance Principles and Elementary Particles*, Princeton University Press (1964).

(16.17) G. KÄLLEN, *Elementary Particle Physics*, Addison-Wesley, Reading, Mass. (1964).

(16.18) A. D. MARTIN and T. D. SPEARMAN, *Elementary Particle Theory*, North Holland, Amsterdam (1970).

(16.19) A. O. WEISSENBERG, *Muons*, North Holland, Amsterdam (1967).

E - ARTICLES

(16.20) M. G. MAYER, "The Structure of the Nucleus", *Scientific American* **184**, 22 (March 1951).

(16.21) R. E. PEIERLS, "The Atomic Nucleus", *Scientific American* **200**, 75 (Jan. 1959).

(16.22) E. U. BARANGER, "The Present Status of the Nuclear Shell Model", *Physics Today,* **26**, 34 (June 1973).

(16.23) S. DE BENEDETTI, "Mesonic Atoms", *Scientific American* **195**, 93 (Oct. 1956).

(16.24) S. DE BENEDETTI, "The Mössbauer Effect", *Scientific American* **202**, 72 (April 1960).

(16.25) R. H. HERBER, "Mössbauer Spectroscopy", *Scientific American* **225**, 86 (Oct. 1971).

(16.26) S. PENMAN, "The Muon", *Scientific American* **205**, 46 (July 1961).

(16.27) R. E. MARSHAK, "The Nuclear Force", *Scientific American* **202**, 98 (March 1960).

(16.28) M. GELL-MANN and E. P. ROSENBAUM, "Elementary Particles", *Scientific American* **197**, 72 (July 1957).

(16.29) G. F. CHEW, M. GELL-MANN and A. H. ROSENFELD, "Strongly Interacting Particles", *Scientific American* **210**, 74 (Feb. 1964).

(16.30) V. F. WEISSKOPF, "The Three Spectroscopies", *Scientific American* **218**, 15 (May 1968).

(16.31) U. AMALDI, "Proton Interactions at High Energies", *Scientific American* **229**, 36 (Nov. 1973).

(16.32) S. WEINBERG, "Unified Theories of Elementary-Particle Interaction", *Scientific American* **231**, 50 (July 1974).

(16.33) S. D. DRELL, "Electron-Positron Annihilation and the New Particles", *Scientific American* **232**,50 (June 1975).

(16.34) R. WILSON, "Form Factors of Elementary Particles", *Physics Today* **22**, 47 (Jan. 1969).

(16.35) E. S. ABERS and B. W. LEE, "Gauge Theories", *Physics Reports* (Amsterdam), **9C**, 1 (1973).

Absorption
and emission of photons, 2073
collision with, 971
of a quantum, a photon, 1311, 1353
of field, 2149
of several photons, 1368
rates, 1334
Acceptor (electron acceptor), 1495
Acetylene (molecule), 878
Action, 341, 1539, 1980
Addition
of angular momenta, 1015, 1043
of spherical harmonics, 1059
of two spins 1/2, 1019
Adiabatic
branching of the potential, 932
Adjoint
matrix, 123
operator, 112
Algebra (commutators), 165
Allowed energy band, 381, 1481, 1491
Ammonia (molecule), 469, 873
Amplitude
scattering amplitude, 929, 953
Angle (quantum), 2258
Angular momentum
addition of momenta, 1015, 1043
and rotations, 717
classical, 1529
commutation relations, 669, 725
conservation, 668, 736, 1016
coupling, 1016
electromagnetic field, 1968, 2043
half-integral, 987
of identical particles, 1497(ex.)
of photons, 1370
orbital, 667, 669, 685
quantization, 394
quantum, 667
spin, 987, 991
standard representation, 677, 691
two coupled momenta, 1091
Anharmonic oscillator, 502, 1135
Annihilation operator, 504, 513, 514, 1597
Annihilation-creation (pair), 1831, 1878
Anomalous
average value, 1828, 1852
dispersion, 2149
Zeeman effect, 987
Anti-normal correlation function, 1782, 1789
Anti-resonant term, 1312
Anti-Stokes (Raman line), 532, 752
Antibunching (photon), 2121
Anticommutation, 1599
field operator, 1754
Anticrossing of levels, 415, 482
Antisymmetric ket, state, 1428, 1431
Antisymmetrizer, 1428, 1431
Applications of the perturbation theory, 1231
Approximation
central field approximation, 1459
secular approximation, 1374
Argument (EPR), 2205
Atom(s), *see* helium, hydrogenoid
donor, 837
dressed, 2129, 2133
many-electron atoms, 1459, 1467
mirrors for atoms, 2153
muonic atom, 541
single atom fluorescence, 2121
Atomic
beam (deceleration), 2025
orbital, 869, 1496(ex.)
parameters, 41
Attractive bosons, 1747
Autler-Townes
doublet, 2144
effect, 1410
Autoionization, 1468
Average value (anomalous), 1828
Azimuthal
quantum number, 811

Band (energy), 381
Bardeen-Cooper-Schrieffer, 1889
Barrier (potential barrier), 68, 367, 373
Basis

Quantum Mechanics, Volume II, Second Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

change of bases, 174
characteristic relations, 101, 119
continuous basis in the space of states, 99
mixed basis in the space of states, 99
BCHSH inequalities, 2209, 2210
BCS, 1889
broken pairs and excited pairs, 1920
coherent length, 1909
distribution functions, 1899
elementary excitations, 1923
excited states, 1919
gap, 1894, 1896, 1923
pairs (wave function of), 1901
phase locking, 1893, 1914, 1916
physical mechanism, 1914
two-particle distribution, 1901
Bell's
inequality, 2208
theorem, 2204, 2208
Benzene (molecule), 417, 495
Bessel
Bessel-Parseval relation, 1507
spherical Bessel function, 944
spherical equation, 961
spherical function, 966
Biorthonormal decomposition, 2194
Bitter, 2059
Blackbody radiation, 651
Bloch
equations, 463, 1358, 1361
theorem, 659
Bogolubov
excitations, 1661
Hamiltonian, 1952
operator method, 1950
phonons, spectrum, 1660
transformation, 1950
Bogolubov-Valatin transformation, 1836, 1919
Bohr, 2207
electronic magneton, 856
frequencies, 249
magneton, *see* front cover pages
model, 40, 819
nuclear magneton, 1237
radius, 820
Boltzmann
constant, *see* front cover pages
distribution, 1630
Born
approximation, 938, 977, 1320
Born-Oppenheimer approximation, 528, 1177, 1190
Born-von Karman conditions, 1490
Bose-Einstein
condensation, 1446, 1638, 1940
condensation (repulsive bosons), 1933
condensation of pairs, 1857
distribution, 652, 1630
statistics, 1446
Bosons, 1434
at non-zero temperature, 1745
attractive, 1747
attractive instability, 1745
condensed, 1638
in a Fock state, 1775
paired, 1881
Boundary conditions (periodic), 1489
Bra, 103, 104, 119
Bragg reflection, 382
Brillouin
formula, 452
zone, 614
Broadband
detector, 2165
optical excitation, 1332
Broadening (radiative), 2138
Broken pairs and excited pairs (BCS), 1920
Brossel, 2059
Bunching of bosons, 1777

C.S.C.O., 133, 137, 153, 236
Canonical
commutation relations, 142, 223, 1984
ensemble, 2289
Hamilton-Jacobi canonical equations, 214
Hamilton-Jacobi equations, 1532
Cauchy principal part, 1517
Center of mass, 812, 1528
Center of mass frame, 814
Central

field approximation, 1459
potential, 1533
Central potential, 803, 841
scattering, 941
stationary states, 804
Centrifugal potential, 809, 888, 893
Chain (von Neumann), 2201
Chain of coupled harmonic oscillators, 611
Change
of bases, 124, 174, 1601
of representation, 124
Characteristic equation, 129
Characteristic relation of an orthonormal basis, 116
Charged harmonic oscillator in an electric field, 575
Charged particle
in an electromagnetic field, 1536
Charged particle in a magnetic field, 240, 321, 771
Chemical bond, 417, 869, 1189, 1210
Chemical potential, 1486, 2287
Circular quanta, 761, 783
Classical
electrodynamics, 1957
histories, 2272
Clebsch-Gordan coefficients, 1038, 1051
Closure relation, 93, 117
Coefficients
Clebsch-Gordan, 1038
Einstein, 1334, 2083
Coherences (of the density matrix), 307
Coherent length (BCS), 1909
Coherent state (field), 2008
Coherent superposition of states, 253, 301, 307
Collision, 923
between identical particles, 1454, 1497(ex.)
between identical particles in classical mechanics, 1420
between two identical particles, 1450
cross section, 926
scattering states, 928
total scattering cross section, 926
with absorption, 971
Combination
of atomic orbitals, 1172
Commutation, 1599
canonical relations, 142, 223
field operator, 1754
of pair field operators, 1861
relations, 1984
Commutation relations
angular momentum, 669, 725
field, 1989, 1996
Commutator algebra, 165
Commutator(s), 91, 167, 171, 187
of functions of operators, 168
Compatibility of observables, 232
Complementarity, 45
Complete set of commuting observables (C.S.C.O.), 133, 137, 236
Complex variables (Lagrangian), 1982
Compton wavelength of the electron, 825, 1235
Condensates
relative phase, 2237
with spins, 2254
Condensation
BCS condensation energy, 1917
Bose-Einstein, 1446, 1857, 1933
Condensed bosons, 1638
Conduction band, 1492
Conductivity (solid), 1492
Configurations, 1467
Conjugate momentum, 214, 323, 1531, 1983, 1987, 1995
Conjugation (Hermitian), 111
Conservation
local conservation of probability, 238
of angular momentum, 668, 736, 1016
of energy, 248
of probability, 237
Conservative systems, 245, 315
Constants of the motion, 248, 317
Contact term, 1273
Contact term (Fermi), 1238, 1247
Contextuality, 2231
Continuous
spectrum, 133, 219, 264, 1316
variables (in a Lagrangian), 1984
Continuum of final states, 1316, 1378, 1380
Contractions, 1802

Convolution product of two functions, 1510
Cooling
Doppler, 2026
down atoms, 2025
evaporative, 2034
Sisyphus, 2034
sub-Doppler, 2155
subrecoil, 2034
Cooper model, 1927
Cooper pairs, 1927
Cooperative effects (BCS), 1916
Correlation functions, 1781, 1804
anti-normal, 1782, 1789
dipole and field, 2113
for one-photon processes, 2084
normal, 1782, 1787
of the field, spatial, 1758
Correlations, 2231
between two dipoles, 1157
between two physical systems, 296
classical and quantum, 2221
introduced by a collision, 1104
Coulomb
field, 1962
gauge, 1965
Coulomb potential
cross section, 979
Coupling
between angular momenta, 1016
between two angular momenta, 1091
between two states, 412
effect on the eigenvalues, 438
spin-orbit coupling, 1234, 1241
Creation and annihilation operators, 504, 513, 514, 1596, 1990
Creation operator (pair of particles), 1813, 1846
Critical velocity, 1671
Cross section
and phase shifts, 951
scattering cross section, 926, 933, 953, 972
Current
metastable current in superfluid, 1667
of particles, 1758
of probability, 240
probability current in hydrogen atom, 851
Cylindrical symmetry, 899(ex.)

Darwin term, 1235, 1279
De Broglie
relation, 10
wavelength, *see* front cover pages, 11, 35
Decay of a discrete state, 1378
Deceleration of an atomic beam, 2025
Decoherence, 2199
Decomposition (Schmidt), 2193
Decoupling (fine or hyperfine structure), 1262, 1291
Degeneracy
essential, 811, 825, 845
exchange degeneracy, 1423
exchange degeneracy removal, 1435
lifted by a perturbation, 1125
rotation invariance, 1072
systematic and accidental, 203
Degenerate eigenvalue, 127, 203, 217, 260
Degereracy
lifted by a perturbation, 1117
parity, 199
Delta Dirac function, 1515
potential well and barriers, 83–85(ex.)
use in quantum mechanics, 97, 106, 280
Density
Lagrangian, 1986
of probability, 264
of states, 389, 1316, 1484, 1488
operator, 449, 1391
operator and matrix, 299
particle density operator, 1756
Density functions
one and two-particle, 1502(ex.)
Depletion (quantum), 1940
Derivative of an operator, 169
Detection probability amplitude (photon), 2166
Detectors (photon), 2165
Determinant
Slater determinant, 1438, 1679
Deuterium, 834, 1107(ex.)
Diagonalization

of a 2×2 matrix, 429
of an operator, 128
Diagram (dressed-atom), 2133
Diamagnetism, 855
Diatomic molecules
rotation, 739
Diffusion (momentum), 2030
Dipole
-dipole interaction, 1142, 1153
-dipole magnetic interaction, 1237
electric dipole transition, 863
electric moment, 1080
Hamiltonian, 2011
magnetic dipole moment, 1084
magnetic term, 1272
trap, 2151
Dirac, *see* Fermi
delta function, 97, 106, 280, 1515
equation, 1233
notation, 102
Direct
and exchange terms, 1613, 1632, 1634, 1646, 1650
term, 1447, 1453
Discrete
bases of the state space, 91
spectrum, 132, 217
Dispersion (anomalous), 2149
Dispersion and absorption (field), 2147
Distribution
Boltzmann, 1630
Bose-Einstein, 1630
Fermi-Dirac, 1630
function (bosons), 1629
function (fermions), 1629
functions, 1625, 1733
functions (BCS), 1899
Distribution law
Bose-Einstein, 652
Divergence (energy), 2007
Donor atom, 837, 1495
Doppler
cooling, 2026
effect, 2022
effect (relativistic), 2022
free spectroscopy, 2105
temperature, 2033
Double
condensate, 2237
resonance method, 2059
spin condensate, 2254
Doublet (Autler-Townes), 2144
Down-conversion (parametric), 2181
Dressed
states and energies, 2133
Dressed-atom, 2129, 2133
diagram, 2133
strong coupling, 2141
weak coupling, 2137

E.P.R., 1225(ex.)
Eckart (Wigner-Eckart theorem), *see* Wigner
Effect
Autler-Townes, 2144
Mössbauer, 2040
photoelectric, 2110
Effective Hamiltonian, 2141
Ehrenfest theorem, 242, 319, 522
Eigenresult, 9
Eigenstate, 217, 232
Eigenvalue, 11, 25, 176, 216
degenerate, 217, 260
equation, 126, 429
of an operator, 126
Eigenvector, 176
of an operator, 126
Einstein, 2110
coefficients, 1334, 1356, 2083
EPR argument, 297, 1104
model, 534, 653
Planck-Einstein relations, 3
temperature, 659
Einstein-Podolsky-Rosen, 2204, 2261
Elastic
scattering, 925
scattering (photon), 2086
scattering, form factor, 1411(ex.)
total cross section, 972
Elastically bound electron model, 1350
Electric
conductivity of a solid, 1492
Electric dipole
Hamiltonian, 2011
interaction, 1342

matrix elements, 1344
moment, 1080
selection rules, 1345
transition and selection rules, 863
transitions, 2056
Electric field (quantized), 2000, 2005
Electric polarisability
NH_3, 484
Electric polarizability
of the 1s state in Hydrogen, 1299
Electric quadrupole
Hamiltonian, 1347
moment, 1082
transitions, 1348
Electric susceptibility
bound electron, 577
of an atom, 1351
Electrical
susceptibility, 1223(ex.)
Electrodynamics
classical, 1957
quantum, 1997
Electromagnetic field
and harmonic oscillators, 1968
and potentials, 321
angular momentum, 1968, 2043
energy, 1966
Lagrangian, 1986, 1992
momentum, 1967, 2019
polarization, 1970
quantization, 631, 637
Electromagnetic interaction of an atom with a wave, 1340
Electromagnetism
fields and potentials, 1536
Electron spin, 393, 985
Electron(s)
configurations, 1463
gas in solids, 1491
in solids, 1177, 1481
mass and charge, *see* front cover pages
Electronic
configuration, 1459
paramagnetic resonance, 1225(ex.)
shell, 827
Elements of reality, 2205
Emergence of a relative phase, 2248, 2253
Emission
of a quantum, 1311
photon, 2080
spontaneous, 2081, 2135
stimulated (or induced), 2081
Energy, *see* Conservation, Uncertainty
and momentum of the transverse electromagnetic field, 1973
band, 381
bands in solids, 1177, 1481
conservation, 248
electromagnetic field, 1966
Fermi energy, 1772
fine structure energy levels, 986
free energy, 2290
levels, 359
levels of harmonic oscillator, 509
levels of hydrogen, 823
of a paired state, 1869
recoil energy, 2023
Ensemble
canonical, 2289
grand canonical, 2291
microcanonical, 2285
statistical ensembles, 2295
Entanglement
quantum, 2187, 2193, 2203, 2242
swapping, 2232
Entropy, 2286
EPR, 2204, 2261
elements of reality, 2205
EPRB, 2205
paradox/argument, 1104
Equation of state
ideal quantum gas, 1640
repulsive bosons, 1745
Equation(s)
Bloch, 1361
Hamilton-Jacobi, 1982, 1983, 1988
Lagrange, 1982, 1993
Lorentz, 1959
Maxwell, 1959
Schrödinger, 11, 12, 306
von Neumann, 306
Essential degeneracy, 811, 825
Ethane (molecule), 1223
Ethylene (molecule), 536, 881

Evanescent wave, 29, 67, 70, 78, 285
Evaporative cooling, 2034
Even operators, 196
Evolution
  field operator, 1765
  of quantum systems, 223
  of the mean value, 241
  operator, 313, 2069
  operator (expansion), 2070
  operator (integral equation), 2069
Exchange, 1611
  degeneracy, 1423
  degeneracy removal, 1435
  energy, 1469
  hole, 1774
  integral, 1474
  term, 1447, 1451, 1453
Excitations
  BCS, 1923
  Bogolubov, 1661
  vacuum, 1623
Excited states (BCS), 1919
Exciton, 838
Exclusion principle (Pauli), 1437, 1444, 1463, 1484
Extensive (or intensive) variables, 2292

Fermi
  contact term, 1238
  energy, 1445, 1481, 1486, 1772
  gas, 1481
  golden rule, 1318
  level, 1486, 1621
  radius, 1621
  surface (modified), 1914
  , *see* Fermi-Dirac
Fermi level
  and electric conductivity, 1492
Fermi-Dirac
  distribution, 1486, 1630, 1717
  statistics, 1446
Fermions, 1434
  in a Fock state, 1771
  paired, 1874
Ferromagnetism, 1477
Feynman
  path, 2267
  postulates, 341
Fictitious spin, 435, 1359
Field
  absorption, 2149
  commutation relations, 1989, 1996
  dispersion and absorption, 2147
  intense laser, 2126
  interaction energy, 1764
  kinetic energy, 1763
  normal variables, 1971
  operator, 1752
  operator (evolution), 1763, 1765
  pair field operator, 1861
  potential energy, 1764
  quantization, 1765, 1999
  quasi-classical state, 2008
  spatial correlation functions, 1758
Final states continuum, 1378, 1380
Fine and hyperfine structure, 1231
Fine structure
  constant, *see* front cover pages, 825
  energy levels, 1478
  Hamiltonian, 1233, 1276, 1478
  Helium atom, 1478
  Hydrogen, 1238
  of spectral lines, 986
  of the states $1s$, $2s$ et $2p$, 1276
Fletcher, 2111
Fluctuations
  boson occupation number, 1633
  intensity, 2125
  vacuum, 644, 2007
Fluorescence (single atom), 2121
Fluorescence triplet, 2144
Fock
  space, 1593, 2004
  state, 1593, 1614, 1769, 2103
Forbidden, *see* Band
  energy band, 381, 390, 1481
  transition, 1345
Forces
  van der Waals, 1151
Form factor
  elastic scattering, 1411(ex.)
Forward scattering (direct and exchange), 1874
Fourier

series and transforms, 1505
Fragmentation (condensate), 1654, 1776
Free
electrons in a box, 1481
energy, 2290
particle, 14
quantum field (Fock space), 2004
spherical wave, 941, 944, 961
spherical waves and plane waves, 967
Free particle
stationary states with well-defined angular momentum, 959
stationary states with well-defined momentum, 19
wave packet, 14, 57, 347
Frequency
Bohr, 249
components of the field (positive and negative), 2072
Rabi's frequency, 1325
Friction (coefficient), 2028
Function
of operators, 166
periodic functions, 1505
step functions, 1521
Fundamental state, 41

Gap (BCS), 1894, 1896, 1923
Gauge, 1343, 1536, 1960, 1963
Coulomb, 1965
invariance, 321
Lorenz, 1965
Gaussian
wave packet, 57, 292, 2305
Generalized velocities, 214, 1530
Geometric quantization, 2311
Gerlach, *see* Stern
GHZ state, 2222, 2227
Gibbs-Duhem relation, 2296
Golden rule (Fermi), 1318
Good quantum numbers, 248
Grand canonical, 1626, 2291
Grand potential, 1627, 1721, 2292
Green's function, 337, 936, 1781, 1786, 1789
evolution, 1785
Greenberger-Horne-Zeilinger, 2227
Groenewold's formula, 2315
Gross-Pitaevskii equation, 1643, 1657
Ground state, 363
harmonic oscillator, 509, 520
Hydrogen atom, 1228(ex.)
Group velocity, 55, 60, 614
Gyromagnetic ratio, 396, 455
orbital, 860
spin, 988

H_2^+ molecular ion, 85(ex.), 417, 1189
Hadronic atoms, 840
Hall effect, 1493
Hamilton
function, 1532
function and equations, 1531
Hamilton-Jacobi canonical equations, 214, 1532, 1982, 1983, 1988
Hamiltonian, 223, 245, 1527, 1983, 1988, 1995
classical, 1531
effective, 2141
electric dipole, 1342, 2011
electric quadrupole, 1347
fine structure, 1233, 1276
hyperfine, 1237, 1267
magnetic dipolar, 1347
of a charged particle in a vector potential, 1539
of a particle in a central potential, 806, 1533
of a particle in a scalar potential, 225
of a particle in a vector potential, 225, 323, 328
Hanbury Brown and Twiss, 2120
Hanle effect, 1372(ex.)
Hard sphere
scattering, 980, 981(ex.)
Harmonic oscillator, 497
in an electric field, 575
in one dimension, 527, 1131
in three dimensions, 569
in two dimensions, 755
infinite chain of coupled oscillators, 611
quasiclassical states, 583
thermodynamic equilibrium, 647

three-dimensional, 841, 899(ex.)
two coupled oscillators, 599
Hartree-Fock
approximation, 1677, 1701
density operator (one-particle), 1691
equations, 1686, 1731
for electrons, 1695
mean field, 1677, 1693
potential, 1706
thermal equilibrium, 1711, 1733
time-dependent, 1701, 1708
Healing length, 1652
Heaviside step function, 1521
Heisenberg
picture, 317, 1763
relations, 19, 39, 41, 45, 55, 232, 290
Helicity (photon), 2051
Helium
energy levels, 1467
ion, 838
isotopes, 1480
isotopes ^{3}He and ^{4}He, 1435, 1446
solidification, 535
Hermite polynomials, 516, 547, 561
Hermitian
conjugation, 111
matrix, 124
operator, 115, 124, 130
Histories (classical), 2272
Hole
creation and annihilation, 1622
exchange, 1774
Holes, 1621
Hybridization of atomic orbitals, 869
Hydrogen, 645
atom, 803
atom in a magnetic field, 853, 855, 862
atom, relativistic energies, 1245
Bohr model, 40, 819
energy levels, 823
fine and hyperfine stucture, 1231
ionisation energy, *see* front cover pages
ionization energy, 820
maser, 1251
molecular ion, 85(ex.), 417, 1189
quantum theory, 41
radial equation, 821
Stark effect, 1298
stationary states, 851
stationary wave functions, 830
Hydrogen-like systems in solid state physics, 837
Hydrogenoid systems, 833
Hyperfine
decoupling, 1262
Hamiltonian, 1237, 1267
Hyperfine structure, *see* Hydrogen, muonium, positronium, Zeeman effect, 1231
Muonium, 1281

Ideal gas, 1625, 1787, 1791, 1804
correlations, 1769
Identical particles, 1419, 1591
Induced
emission, 1334, 1366, 2081
emission of a quantum, 1311
emission of photons, 1355
Inequality (Bell's), 2208
Infinite one-dimensional well, 271
Infinite potential well, 74
in two dimensions, 201
Infinitesimal unitary operator, 178
Insulator, 1492
Integral
exchange integral, 1474
scattering equation, 935
Intense laser fields, 2126
Intensive (or extensive) variables, 2292
Interaction
between magnetic dipoles, 1141
dipole-dipole interaction, 1141, 1153
electromagnetic interaction of an atom with a wave, 1340
field and particles, 2009
field and atom, 2010
magnetic dipole-dipole interaction, 1237
picture, 353, 1393, 2070
tensor interaction, 1141
Interference
photons, 2167
two-photon, 2170, 2183
Ion H_2^+, 1189

Ionization
photo-ionization, 2109
tunnel ionization, 2126
Isotropic radiation, 2079

Jacobi, *see* Hamilton

Kastler, 2059, 2062
Ket, *see* state, 103, 119
for identical particles, 1436
Kuhn, *see* Thomas

Lagrange
equations, 1530, 1982, 1993
fonction and equations, 214
multipliers, 2281
Lagrangian, 1530, 1980
densities, 1986
electromagnetic field, 1986, 1992
formulation of quantum mechanics, 339
of a charged particle in an electromagnetic field, 1538
particle in an electromagnetic field, 323
Laguerre-Gaussian beams, 2065
Lamb shift, 645, 1245, 1388, 2008
Landau levels, 771
Landé factor, 1072, 1107(ex.), 1256, 1292
Laplacian, 1527
of $1/r$, 1524
of $Y_l^m(\theta,\varphi)/r^{l+1}$, 1526
Larmor
angular frequency, 857
precession, 394, 396, 410, 455, 857, 1071
Laser, 1359, 1365
Raman laser, 2093
saturation, 1370
trap, 2151
Lattices (optical), 2153
Least action
principle of, 1539
Legendre
associated function, 714
polynomial, 713
Length (healing), 1652
Level
anticrossing, 415, 482
Fermi level, 1621
Lifetime, 343, 485, 645
of a discrete state, 1386
radiative, 2081
Lifting of degeneracy by a perturbation, 1125
Light
quanta, 3
shifts, 1334, 2138, 2151, 2156
Linear, *see* operator
combination of atomic orbitals, 1172
operators, 90, 108, 163
response, 1350, 1357, 1364
superposition of states, 253
susceptibility, 1365
Local conservation of probability, 238
Local realism, 2209, 2230
Longitudinal
fields, 1961
relaxation, 1400
relaxation time, 1401
Lorentz equations, 1959
Lorenz (gauge), 1965

Magnetic
dipole term, 1272
dipole-dipole interaction, 1237
effect of a magnetic field on the levels of the Hydrogen atom, 1251
hyperfine Hamiltonian, 1267
interactions, 1232, 1237
quantum number, 811
resonance, 455
susceptibility, 1224, 1487
Magnetic dipole
Hamiltonian, 1347
transitions and selection rules, 1084, 1098, 1348
Magnetic dipoles
interactions between two dipoles, 1141
Magnetic field
and vector potential, 321
charged particle in a, 240, 771
effects on hydrogen atom, 853, 855
harmonic oscillator in a, 899(ex.)
Hydrogen atom in a magnetic field, 1263, 1289

multiplets, 1074
quantized, 2000, 2005
Magnetism (spontaneous), 1737
Many-electron atoms, 1459
Maser, 477, 1359, 1365
hydrogen, 1251
Mass correction (relativistic), 1234
Master equation, 1358
Matrice(s), 119, 121
diagonalization of a 2×2 matrix, 429
Pauli matrices, 425
unitary matrix, 176
Maxwell's equations, 1959
Mean field (Hartree-Fock), 1693, 1708, 1725
Mean value of an observable, 228
evolution, 241
Measurement
general postulates, 216, 226
ideal von Neumann measurement, 2196
of a spin 1/2, 394
of observables, 216
on a part of a physical system, 293
state after measurement, 221, 227
Mendeleev's table, 1463
Metastable superfluid flow, 1671
Methane (molecule), 883
Microcanonical ensemble, 2285
Millikan, 2111
Minimal wave packet, 290, 520, 591
Mirrors for atoms, 2153
Mixing of states, 1121, 1137
Model
Cooper model, 1927
Einstein model, 534
elastically bound electron, 1350
vector model of atom, 1071
Modes
vibrational modes, 599, 611
Modes (radiation), 1974, 1975
Molecular ion, 417
Molecule(s)
chemical bond, 417, 869, 873, 878, 883, 1189
rotation, 796
vibration, 527, 1137
vibration-rotation, 885
Mollow, 2144
Moment
quadrupole electric moment, 1225(ex.)
Momentum, 1539
conjugate, 214, 323, 1983, 1987, 1995
diffusion, 2030
electromagnetic field, 1967, 2019
mechanical momentum, 328
Monogamy (quantum), 2221
Mössbauer effect, 1415, 2040
Motional narrowing, 1323
condition, 1323, 1398, 1408
Multiphoton transition, 1368, 2040, 2097
Multiplets, 1072, 1074, 1467
Multipliers (Lagrange), 2281
Multipolar waves, 2052
Multipole
moments, 1077
Multipole operators
introduction, 1077, 1083
parity, 1082
Muon, 527, 541, 1281
Muonic atom, 541, 839
Muonium, 835
hyperfine structure, 1281
Zeeman effect, 1281

Narrowing (motional), 1323, 1408
condition, 1398
Natural width, 345, 1388
Need for a quantum treatment, 2118, 2120
Neumann
spherical function, 967
Neutron mass, *see* front cover pages
Non-destructive detection of a photon, 2159
Non-diagonal order (BCS), 1912
Non-locality, 2204
Non-resonant excitation, 1350
Non-separability, 2207
Nonlinear
response, 1357, 1368
susceptibility, 1369
Norm
conservation, 238
of a state vector, 104, 237
of a wave function, 13, 90, 99

Normal
correlation function, 1782, 1787
variables, 602, 616, 631, 633
variables (field), 1971
Nuclear
multipole moments, 1088
Bohr magneton, 1237
Nucleus
spin, 1088
volume effect, 1162, 1268
Number
occupation number, 1439, 1593
photon number, 2135
total number of particles in an ideal gas, 1635

Observable(s), 130
C.S.C.O., 133, 137
commutation, 232
compatibility, 232
for identical particles, 1429, 1441
mean value, 228
measurement of, 216, 226
quantization rules, 223
symmetric observables, 1441
transformation by permutation, 1434
whose commutator is $i\hbar$, 187, 289
Occupation number, 1439, 1593
operator, 1598
Odd operators, 196
One-particle
Hartree-Fock density operator, 1691
operators, 1603, 1605, 1628, 1756
Operator(s)
adjoint operator, 112
annihilation operator, 504, 513, 514, 1597
creation and annihilation, 1990
creation operator, 504, 513, 514, 1596
derivative of an operator, 169
diagonalization, 126, 128
even and odd operators, 196
evolution operator, 313, 2069
field, 1752
function of, 166
Hermitian operators, 115
linear operators, 90, 108, 163
occupation number, 1598
one-particle operator, 1603, 1605, 1628, 1756
parity operator, 193
particle density operator, 1756
permutation operators, 1425, 1430
potential, 168
product of, 90
reduced to a single particle, 1607
representation, 121
restriction, 165
restriction of, 1125
rotation operator, 1001
symmetric, 1628, 1755
translation operator, 190
two-particle operator, 1608, 1610, 1631, 1756
unitary operators, 173
Weyl operator, 2300
Oppenheimer, *see* Born, 1177, 1190
Optical
excitation (broadband), 1332
lattices, 2153
pumping, 2062, 2140
Orbital
angular momentum (of radiation), 2052
atomic orbital, 1496(ex.)
hybridization, 869
linear combination of atomic orbitals, 1172
quantum number, 1463
state space, 988
Order parameter for pairs, 1851
Orthonormal basis, 91, 99, 101, 133
characteristic relation, 116
Orthonormalization
and closure relations, 101, 140
relation, 116
Oscillation(s)
between two discrete states, 1374
between two quantum states, 418
Rabi, 2134
Oscillator
anharmonic, 502
harmonic, 497
strength, 1352

Pair(s)

annihilation-creation of pairs, 1831, 1874, 1887
BCS, wave function, 1909
Cooper, 1927
of particles (creation operator), 1813, 1846
pair field (commutation), 1861
pair field operator, 1845
pair wave function, 1851
Paired
bosons, 1881
fermions, 1874
state energy, 1869
states, 1811
states (building), 1818
Pairing term, 1878
Paramagnetism, 855
Parametric down-conversion, 2181
Parity, 2106
degeneracy, 199
of a permutation operator, 1431
of multipole operators, 1082
operator, 193
Parseval
Parseval-Plancherel equality, 20
Parseval-Plancherel formula, 1511, 1521
Partial
reflection, 79
trace of an operator, 309
waves in the potential, 948
waves method, 941
Particle (current), 1758
Particles and holes, 1621
Partition function, 1626, 1627, 1717
Path
integral, 2267
space-time path, 339
Pauli
exclusion principle, 1437, 1444, 1463, 1481
Hamiltonian, 1009(ex.)
matrices, 425, 991
spin theory, 986
spinor, 993
Penetrating orbit, 1463
Penrose-Onsager criterion, 1776, 1860, 1947
Peres, 2212
Periodic
boundary conditions, 1489
classification of elements, 1463
functions, 1505
potential (one-dimensional), 375
Permutation operators, 1425, 1430
Perturbation
applications of the perturbation theory, 1231
lifting of a degeneracy, 1125
one-dimensional harmonic oscillator, 1131
random perturbation, 1320, 1325, 1390
sinusoidal, 1311
stationary perturbation theory, 1115
Perturbation theory
time dependent, 1303
Phase
locking (BCS), 1893, 1916
locking (bosons), 1938, 1944
relative phase between condensates, 2237, 2248
velocity, 37
Phase shift (collision), 951, 1497(ex.)
with imaginary part, 971
Phase velocity, 21
Phonons, 611, 626
Bogolubov phonons, 1660
Photodetection
double, 2172, 2184
single, 2169, 2171
Photoelectric effect, 1412(ex.), 2110
Photoionization, 2109, 2165
rate, 2115, 2124
two-photon, 2123
Photon, 3, 631, 651, 2004, 2005, 2110
absorption and emission, 2067
angular momentum, 1370
antibunching, 2121
detectors, 2165
non-destructive detection, 2159
number, 2135
scattering (elastic), 2086
scattering by an atom, 2085
vacuum, 2007
, *see* Absorption, Emission
Picture

Heisenberg, 317, 1763
interaction, 1393, 2070
Pitaevskii (Gross-Pitaevskii equation), 1643, 1657
Plancherel, *see* Parseval
Planck
constant, *see* front cover pages, 3
law , 2083
Planck-Einstein relations, 3, 10
Plane wave, 14, 19, 95, 943
Podolsky (EPR argument), 297, 1104
Pointer states, 2199
Polarizability
of the 1s state in Hydrogen, 1299
Polarization
electromagnetic field, 1970
of Zeeman components, 1295
space-dependent, 2156
Polynomial method (harmonic oscillator), 555, 842
Polynomials
Hermite polynomials, 516, 547, 561
Position and momentum representations, 181
Positive and negative frequency components, 2072
Positron, 1281
Positronium, 836
hyperfine structure, 1281
Zeeman effect, 1281
Postulate (von Neumann projection), 2202
Postulates of quantum mechanics, 215
Potential
adiabatic branching, 932
barrier, 26, 68, 367, 373
centrifugal potential, 809, 888, 893
Coulomb potential, cross section, 979
cylindrically symmetric, 899(ex.)
Hartree-Fock, 1706
infinite one-dimensional well, 74
operator, 168
scalar and vector potentials, 1536, 1960, 1963
scattering by a, 923
self-consistent potential, 1461
square potential, 63
square well, 29
step, 28, 65, 75, 284
well, 71, 367
well (arbitrary shape), 359
well (infinite one-dimensional), 271
well (infinite two-dimensional, 201
Yukawa potential, 977
Precession
Larmor precession, 396, 1071
Thomas precession, 1235
Preparation of a state, 235
Pressure (ideal quantum gas), 1640
Principal part, 1517
Principal quantum number, 827
Principle
of least action, 1539, 1980
of spectral decomposition, 11, 216
of superposition, 237
Probability
amplitude, 11, 253, 259
conservation, 237
current, 240, 283, 333, 349, 932
current in hydrogen atom, 851
density, 11, 264
fluid, 932
of photon absorption, 2076
of the measurement results, 9, 11
transition probability, 439
Process (pair annihilation-creation), 1878, 1887
Product
convolution product of functions, 1510
of matrices, 122
of operators, 90
scalar product, 101, 141, 149, 161
state (tensor product), 311
tensor product, 147
tensor product, applications, 441
Projection theorem, 1070
Projector, 109, 133, 165, 218, 222, 1108(ex.)
Propagator
for the Schrödinger equation, 335
of a particle, 2267, 2272
Proper result, 9
Proton
mass, *see* front cover pages
spin and magnetic moment, 1237, 1274
Pumping, 1358

Pure (state or case), 301

Quadrupolar electric moment, 1082, 1225(ex.)
Quanta (circular), 761, 783
Quantization
  electrodynamics, 1997
  electromagnetic field, 631, 637, 1997
  of a field, 1765
  of angular momentum, 394, 677
  of energy, 3, 11, 71, 359
  of measurement results, 9, 216, 398
  of the measurement results, 405
  rules, 11, 223, 226, 2274
Quantum
  angle, 2258
  electrodynamics, 1245, 1282, 1997
  entanglement, 2187, 2193
  monogamy, 2221
  number
    orbital, 1463
    principal quantum number, 827
  numbers (good), 248
  resonance, 417
  treatment needed, 2118, 2120
Quasi-classical
  field states, 2008
  states, 765, 791, 801
  states of the harmonic oscillator, 583
Quasi-particles, 1736, 1840
  Bogolubov phonons, 1954
  Quasi-particle vacuum, 1836

Rabi
  formula, 440, 460, 1324, 1376
  formula), 419
  frequency, 1325
  oscillation, 2134
Radial
  equation, 842
  equation (Hydrogen), 821
  equation in a central potential, 808
  integral, 1277
  quantum number, 811
Radiation
  isotropic, 2079
  pressure, 2024
Radiative
  broadening, 2138
  cascade of the dressed atom, 2145
Raman
  effect, 532, 740, 1373(ex.)
  laser, 2093
  scattering, 2091
  scattering (stimulated), 2093
Random perturbation, 1320, 1325, 1390
Rank (Schmidt), 2196
Rate (photoionization), 2115, 2124
Rayleigh
  line, 752
  scattering, 532, 2089
Realism (local), 2205, 2209
Recoil
  blocking, 2036
  effect of the nucleus, 834
  energy, 1415, 2023
  free atom, 2020
  suppression, 2040
Reduced
  density operator, 1607
  mass, 813
Reduction of the wave packet, 221, 279
Reflection on a potential step, 285
Refractive index, 2149
Reiche, *see* Thomas
Relation (Gibbs-Duhem), 2296
Relative
  motion, 814
  particle, 814
  phase between condensates, 2248, 2258
  phase between spin condensates, 2253
Relativistic
  corrections, 1233, 1478
  Doppler effect, 2022
  mass correction, 1234
Relaxation, 465, 1358, 1390, 1413, 1414(ex.)
  general equations, 1397
  longitudinal, 1400
  longitudinal relaxation time, 1401
  transverse, 1403
  transverse relaxation time, 1406
Relay state, 2086, 2098, 2106
Renormalization, 2007
Representation(s)
  change of, 124
  in the state space, 116

of operators, 121
position and momentum, 139, 181
Schrödinger equation, 183–185
Repulsion between electrons, 1469
Resonance
magnetic resonance, 455
quantum resonance, 417, 1158
scattering resonance, 69, 954, 983(ex.)
two resonnaces with a sinusoidal excitation, 1365
width, 1312
with sinusoidal perturbation, 1311
Restriction of an operator, 165, 1125
Rigid rotator, 740, 1222(ex.)
Ritz theorem, 1170
Root mean square deviation
general definition, 230
Rosen (EPR argument), 297, 1104
Rotating frame, 459
Rotation(s)
and angular momentum, 717
invariance and degeneracy, 734
of diatomic molecules, 739
of molecules, 796, 885
operator(s), 720, 1001
rotation invariance, 1478
rotation invariance and degeneracy, 1072
Rotator
rigid rotator, 740, 1222(ex.)
Rules
quantization rules, 2274
selection rules, 197
Rutherford's formula, 979
Rydberg constant, *see* front cover pages

Saturation
of linear response, 1368
of the susceptibility, 1369
Scalar
and vector potentials, 321, 1536
interaction between two angular momenta, 1091
observable, operator, 732, 737
potential, 225
product, 89, 92, 101, 141, 149, 161
product of two coherent states, 593
Scattering
amplitude, 929, 953
by a central potential, 941
by a hard sphere, 980, 981(ex.)
by a potential, 923
cross section, 933, 953, 972
cross section and phase shifts, 951
inelastic, 2091
integral equation, 935
of particles with spin, 1102
of spin 1/2 particles, 1108(ex.)
photon, 2086
Raman, 2091
Rayleigh, 532, 2089
resonance, 954, 983(ex.)
resonant, 2089
stationary scattering states, 951
stationary states, 928
stimulated Raman, 2093
Schmidt
decomposition, 2193
rank, 2196
Schottky anomaly, 654
Schrödinger, 2190
equation, 11, 12, 223, 306
equation in momentum representation, 184
equation in position representation, 183
equation, physical implications, 237
equation, resolution for conservative systems, 245
picture, 317
Schwarz inequality, 161
Second
quantization, 1766
harmonic generation, 1368
Secular approximation, 1316, 1374
Selection rules, 197, 863, 2014, 2056
electric quadrupolar, 1348
magnetic dipolar, 1098, 1348
Self-consistent potential, 1461
Semiconductor, 837, 1493
Separability, 2207, 2223
Separable density operator, 2223
Shell (electronic), 827
Shift

light shift, 2138
of a discrete state, 1387
Singlet, 1024, 1474
Sinusoidal perturbation, 1311, 1374
Sisyphus
cooling, 2034
effect, 2155
Slater determinant, 1438, 1679
Slowing down atoms, 2025
Solids
electronic bands, 1177
energy bands of electrons, 1491
energy bands of electrons in solids, 381
hydrogen-like systems in solid state physics, 837
Space (Fock), 1593
Space-dependent polarization, 2156
Space-time path, 339, 1539
Spatial correlations (ideal gas), 1769
Specific heat
of an electron gas, 1484
of metals, 1487
of solids, 653
two level system, 654
Spectral
decomposition principle, 7, 11, 216
function, 1795
terms, 1469
Spectroscopy (Doppler free), 2105
Spectrum
BCS elementary excitation, 1923
continuous, 219, 264
discrete, 132, 217
of an observable, 126, 216
Spherical
Bessel equation, 961
Bessel function, 944, 966
free spherical waves, 961
free wave, 944
Neumann function, 967
wave, 941
waves and plane waves, 967
Spherical harmonics, 689, 705
addition of, 1059
expression for $l = 0, 1, 2$, 709
general expression, 707
Spin
and magnetic moment of the proton, 1237
angular momentum, 987
electron, 985, 1289
fictitious, 435
gyromagnetic ratio, 396, 455, 988
nuclear, 1088
of the electron, 393
Pauli theory, 986, 988
quantum description, 985, 991
rotation operator, 1001
scattering of particles with spin, 1102
spin 1 and radiation, 2044, 2049, 2050
system of two spins, 441
Spin 1/2
density operator, 449
ensemble of, 1358
fictitious, 1359
interaction between two spins, 1141
preparation and measurement, 401
scattering of spin 1/2 particles, 1108(ex.)
Spin-orbit coupling, 1018, 1234, 1241, 1279
Spin-statistics theorem, 1434
Spinor, 993
rotation, 1005
Spontaneous
emission, 343, 645, 1301, 2081, 2135
emission of photons, 1356
magnetism of fermions, 1737
Spreading of a wave packet, 59, 348
Square
barrier of potential, 26, 68
potential, 26, 63, 75, 283
potential well, 71, 271
spherical well, 982(ex.)
Standard representation (angular momentum), 677, 691
Stark effect in Hydrogen atom, 1298
State(s), *see* Density operator
density of, 389, 1316, 1484, 1488
Fock, 1593, 1614, 1769, 2103
ground state, 363
mixing of states by a perturbation, 1121
orbital state space, 988
paired, 1811

pointer states, 2199
quasi-classical states, 583, 765, 791, 801
relay state, 2086, 2098, 2106
stable and unstable states, 485
state after measurement, 221
state preparation, 235
stationary, 63, 359, 375
stationary state, 24, 246
stationary states in a central potential, 804
unstable, 343
vacuum state, 1595
vector, 102, 215
Stationary
perturbation theory, 1115
phase condition, 18, 54
scattering states, 928, 951
states, 24, 63, 246, 359
states in a periodic potential, 375
states with well-defined angular momentum, 944, 959
states with well-defined momentum, 943
Statistical
entropy, 2217
mechanics (review of), 2285
mixture of states, 253, 299, 304, 450
Statistics
Bose-Einstein, 1446
Fermi-Dirac, 1446
Step
function, 1521
potential, 28, 65, 75, 284
Stern-Gerlach experiment, 394
Stimulated
(or induced) emission, 1334, 1366, 2081
Raman scattering, 2093
Stokes Raman line, 532, 752
Stoner (spontaneous magnetism), 1737
Strong coupling (dressed-atom), 2141
Subrecoil cooling, 2034
Sum rule (Thomas-Reiche-Kuhn), 1352
Superfluidity, 1667, 1674
Superposition
of states, 253
principle, 7, 237
principle and physical predictions, 253
Surface (modified Fermi surface), 1914
Susceptibility, *see* Linear, nonlinear, tensor
electric susceptibility of an atom, 1351
electrical susceptibility, 577, e1223
electrical susceptibility of NH_3, 484
magnetic susceptibility, 1224
tensor, 1224, 1410(ex.)
Swapping (entanglement), 2232
Symmetric
ket, state, 1428, 1431
observables, 1429, 1441
operators, 1603, 1605, 1608, 1610, 1628, 1631, 1755
Symmetrization
of observables, 224
postulate, 1434
Symmetrizer, 1428, 1431
System
time evolution of a quantum system, 223
two-level system, 435
Systematic
and accidental degeneracies, 203
degeneracy, 845

Temperature (Doppler), 2033
Tensor
interaction, 1141
product, 147, 441
product of operators, 149
product state, 295, 311
product, applications, 201
susceptibility tensor, 1224
Term
direct and exchange terms, 1613, 1632, 1634, 1646, 1650
pairing, 1878
spectral terms, 1467, 1469
Theorem
Bell, 2204, 2208
Bloch, 659
projection, 1070
Ritz, 1170
Wick, 1799, 1804

Wigner-Eckart, 1065, 1085, 1254
Thermal wavelength, 1635
Thermodynamic equilibrium, 308
harmonic oscillator, 647
ideal quantum gas, 1625
spin 1/2, 452
Thermodynamic potential (minimization), 1715
Thomas precession, 1235
Thomas-Reiche-Kuhn sum rule, 1352
Three-dimensional harmonic oscillator, 569, 841, 899(ex.)
Three-level system, 1409(ex.)
Three-photon transition, 1370
Time evolution of quantum systems, 223
Time-correlations (fluorescent photons), 2145
Time-dependent
Gross-Pitaevskii equation, 1657
perturbation theory, 1303
Time-energy uncertainty relation, 250, 279, 345, 1312, 1389
Torsional oscillations, 536
Torus (flow in a), 1667
Total
elastic scattering cross section, 972
reflection, 67, 75
scattering cross section (collision), 926
Townes
Autler-Townes effect, 1410
Trace
of an operator, 163
partial trace of an operator, 309
Transform (Wigner), 2297
Transformation
Bogolubov, 1950
Bogolubov-Valatin, 1836, 1919
Gauge, 1960
of observables by permutation, 1434
Transition, *see* Probability, Forbidden, Electric dipole, Magnetic dipole, Quadrupole electric dipole, 2056
magnetic dipole transition, 1098
probability, 439, 1308, 1321, 1355
probability per unit time, 1319
probability, spin 1/2, 460
three-photon transition, 1370
two-photon, 2097
virtual, 2100
Translation operator, 190, 579, 791
Transpositions, 1431
Transverse
fields, 1961
relaxation, 1403
relaxation time, 1406
Trap
dipolar, 2151
laser, 2151
Triplet, 1024, 1474
fluorescence triplet, 2144
Tunnel
effect, 29, 70, 365, 476, 540, 1177
ionization, 2126
Two coupled harmonic oscillators, 599
Two-dimensional
harmonic oscillator, 755
infinite potential well, 201
wave packets, 49
Two-level system, 393, 411, 435, 1357
Two-particle operators, 1608, 1610, 1631, 1756
Two-photon
absorption, 1373(ex.)
interference, 2170, 2183
transition, 1409(ex.), 2097

Uncertainty
relation, 19, 39, 41, 45, 232, 290
time-energy uncertainty relation, 1312
Uniqueness of the measurement result, 2201
Unitary
matrix, 125, 176
operator, 173, 314
transformation of operators, 177
Unstable states, 343

Vacuum
electromagnetism, 644, 2007
excitations, 1623
fluctuations, 2007
photon vacuum, 2007
quasi-particule vacuum, 1836
state, 1595
Valence band, 1493

Van der Waals forces, 1151
Variables
intensive or extensive, 2292
normal variables, 602, 616, 631, 633
Variational method, 1169, 1190, 1228(ex.)
Vector
model, 1091
model of the atom, 1071, 1256
observable, operator, 732
operator, 1065
potential, 225
potential of a magnetic dipole, 1268
Velocity
critical, 1671
generalized velocities, 214, 1530
group velocity, 23, 614
phase velocity, 21, 37
Vibration(s)
modes, 599, 611
modes of a continuous system, 631
of molecules, 885, 1137
of nuclei in a crystal, 534, 611, 653
of the nuclei in a molecule, 527
Violations of Bell's inequalities, 2210, 2265
Virial theorem, 350, 1210
Virtual transition, 2100
Volume effect, 544, 840, 1162, 1268
Von Neumann
chain, 2201
equation, 306
ideal measurement, 2196
reduction postulate, 2202
statistical entropy, 2217
Vortex in a superfluid, 1667

Water (molecule), 873, 874
Wave (evanescent), 67
Wave function, 88, 140, 226
BCS pairs, 1901, 1909
Hydrogen, 830
norm, 90
pair wave functions, 1851
particle, 11
Wave packet(s)
Gaussian, 57, 2305
in a potential step, 75
in three dimensions, 53
minimal, 290, 520, 591
motion in a harmonic potential, 596
one-photon, 2168
particle, 13
photon, 2163
propagation, 20, 57, 242, 398
reduction, 221, 227, 265, 279
spreading, 57, 59, 347, 348(ex.)
two-dimension, 49
two-photons, 2181
Wave(s)
de Broglie wavelength, 10, 35
evanescent, 29
free spherical waves, 961
multipolar, 2052
partial waves, 948
plane, 14, 19, 943
wave function, 11, 88, 140, 226
Wave-particle duality, 3, 45
Wavelength
Compton wavelength, 1235
de Broglie, 10
Weak coupling (dressed-atom), 2137
Well
potential square well, 29
potential well, 367
Weyl
operator, 2300
quantization, 2311
Which path type of experiments, 2202
Wick's theorem, 1799, 1804
Wigner transform, 2297
Wigner-Eckart theorem, 1065, 1085, 1254

Young (double slit experiment), 4
Yukawa potential, 977

Zeeman
components, polarizations, 865
effect, 855, 862, 987, 1251, 1253, 1257, 1261, 1281
polarization of the components, 1295
slower, 2025
Zeeman effect
Hydrogen, 1289
in muonium, 1281
in positronium, 1281
Muonium, 1284

Zone (Brillouin zone), 614

WILEY-VCH

This new edition of the unrivalled textbook introduces concepts such as the quantum theory of scattering by a potential, special and general cases of adding angular momenta, time-independent and time-dependent perturbation theory, and systems of identical particles. The entire book has been revised to take into account new developments in quantum mechanics curricula.

The textbook retains its original style in is new edition: fundamental concepts are explained in *chapters,* each followed by *complements* that provide more detailed discussions, examples and applications.

- The quantum mechanics classic in a new edition: written by 1997 Nobel laureate Claude Cohen-Tannoudji and his colleagues Bernard Diu and Franck Laloë
- As easily comprehensible as possible: all steps of the physical background and its mathematical representation are spelled out explicitly
- Comprehensive: in addition to the fundamentals themselves, the book contains more than 170 worked examples plus exercises

Claude Cohen-Tannoudji *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris where he also studied and received his PhD in 1962. In 1973 he became Professor of atomic and molecular physics at the Collège des France. His main research interests were optical pumping, quantum optics and atom-photon interactions. In 1997, Claude Cohen-Tannoudji, together with Steven Chu and William D. Phillips, was awarded the Nobel Prize in Physics for his research on laser cooling and trapping of neutral atoms.*

Bernard Diu *was Professor at the Denis Diderot University (Paris VII). He was engaged in research at the Laboratory of Theoretical Physics and High Energy where his focus was on strong interactions physics and statistical mechanics.*

Franck Laloë *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris. His first assignment was with the University of Paris VI before he was appointed to the CNRS, the French National Research Center. His research was focused on optical pumping, statistical mechanics of quantum gases, musical acoustics and the foundations of quantum mechanics.*

Cover Design: Wiley
Cover Image: © antishock/Getty Images

ISBN 978-3-527-34554-0

www.wiley-vch.de

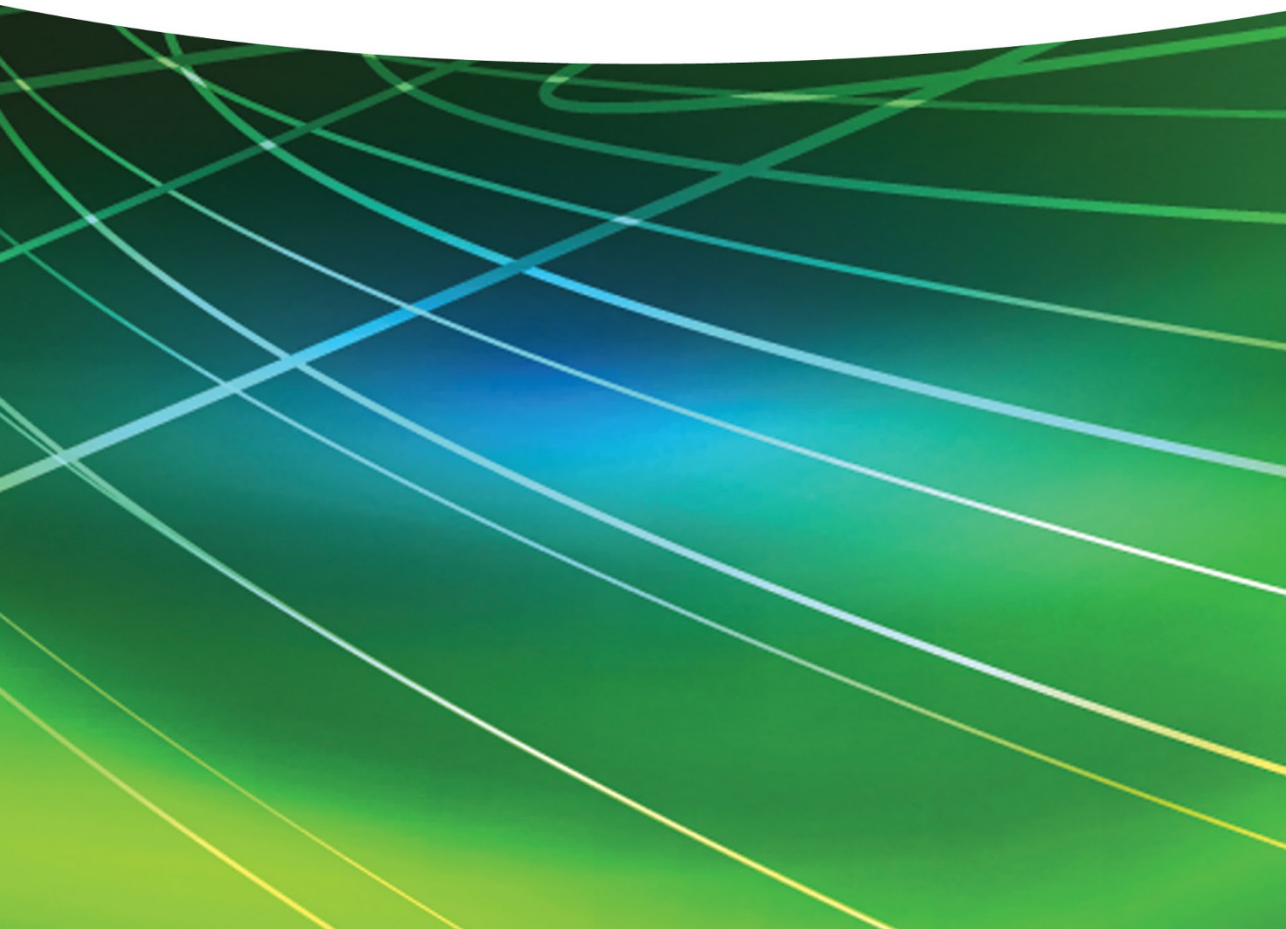
Also available as an e-book

WILEY-VCH

Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë

Quantum Mechanics

Volume III: Fermions, Bosons, Photons, Correlations, and Entanglement



QUANTUM MECHANICS

Volume III

Fermions, Bosons, Photons, Correlations, and Entanglement

Claude Cohen-Tannoudji, Bernard Diu,
and Franck Laloë

Translated from the French by **Nicole Ostrowsky and Dan Ostrowsky**

Authors

Prof. Dr. Claude Cohen-Tannoudji
Laboratoire Kastler Brossel (ENS)
24 rue Lhomond
75231 Paris Cedex 05
France

Prof. Dr. Bernard Diu
4 rue du Docteur Roux
91440 Boures-sur-Yvette
France

Prof. Dr. Frank Laloë
Laboratoire Kastler Brossel (ENS)
24 rue Lhomond
75231 Paris Cedex 05
France

Cover Image
© antishock/Getty Images

First Edition

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data:
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de.

© 2020 WILEY-VCH Verlag GmbH & Co. KGaA, Boschstr. 12, 69469 Weinheim, Germany

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Print ISBN 978-3-527-34555-7
ePDF ISBN 978-3-527-82274-4
ePub ISBN 978-3-527-82275-1

Cover Design Tata Consulting Services
Printing and Binding CPI Ebner & Spiegel

Printed on acid-free paper.

Directions for Use

This book is composed of chapters and their complements:

– *The chapters* contain the fundamental concepts. Except for a few additions and variations, they correspond to a course given in the last year of a typical undergraduate physics program (Volume I) or of a graduate program (Volumes II and III). The 21 chapters are *complete in themselves* and can be studied independently of the complements.

– *The complements* follow the corresponding chapter. Each is labelled by a letter followed by a subscript, which gives the number of the chapter (for example, the complements of Chapter V are, in order, A_V, B_V, C_V, etc.). They can be recognized immediately by the symbol • that appears at the top of each of their pages.

The complements vary in character. Some are intended to expand the treatment of the corresponding chapter or to provide a more detailed discussion of certain points. Others describe concrete examples or introduce various physical concepts. One of the complements (usually the last one) is a collection of exercises.

The *difficulty* of the complements varies. Some are very simple examples or extensions of the chapter. Others are more difficult and at the graduate level or close to current research. In any case, the reader should have studied the material in the chapter before using the complements.

The complements are generally independent of one another. *The student should not try to study all the complements of a chapter at once.* In accordance with his/her aims and interests, he/she should choose a small number of them (two or three, for example), plus a few exercises. The other complements can be left for later study. To help with the choise, the complements are listed at the end of each chapter in a "*reader's guide*", which discusses the difficulty and importance of each.

Some passages within the book have been set in small type, and these can be omitted on a first reading.

Foreword

Quantum mechanics is a branch of physics whose importance has continually increased over the last decades. It is essential for understanding the structure and dynamics of microscopic objects such as atoms, molecules and their interactions with electromagnetic radiation. It is also the basis for understanding the functioning of numerous new systems with countless practical applications. This includes lasers (in communications, medicine, milling, etc.), atomic clocks (essential in particular for the GPS), transistors (communications, computers), magnetic resonance imaging, energy production (solar panels, nuclear reactors), etc. Quantum mechanics also permits understanding surprising physical properties such as superfluidity or supraconductivity. There is currently a great interest in entangled quantum states whose non-intuitive properties of nonlocality and nonseparability permit conceiving remarkable applications in the emerging field of quantum information. Our civilization is increasingly impacted by technological applications based on quantum concepts. This why a particular effort should be made in the teaching of quantum mechanics, which is the object of these three volumes.

The first contact with quantum mechanics can be disconcerting. Our work grew out of the authors' experiences while teaching quantum mechanics for many years. It was conceived with the objective of easing a first approach, and then aiding the reader to progress to a more advance level of quantum mechanics. The first two volumes, first published more than forty years ago, have been used throughout the world. They remain however at an intermediate level. They have now been completed with a third volume treating more advanced subjects. Throughout we have used a progressive approach to problems, where no difficulty goes untreated and each aspect of the diverse questions is discussed in detail (often starting with a classical review).

This willingness to go further "without cheating or taking shortcuts" is built into the book structure, using two distinct linked texts: *chapters* and *complements*. As we just outlined in the "Directions for use", the chapters present the general ideas and basic concepts, whereas the complements illustrate both the methods and concepts just exposed.

Volume I presents a general introduction of the subject, followed by a second chapter describing the basic mathematical tools used in quantum mechanics. While this chapter can appear long and dense, the teaching experience of the authors has shown that such a presentation is the most efficient. In the third chapter the postulates are announced and illustrated in many of the complements. We then go on to certain important applications of quantum mechanics, such as the harmonic oscillator, which lead to numerous applications (molecular vibrations, phonons, etc.). Many of these are the object of specific complements.

Volume II pursues this development, while expanding its scope at a slightly higher level. It treats collision theory, spin, addition of angular momenta, and both time-dependent and time-independent perturbation theory. It also presents a first approach to the study of identical particles. In this volume as in the previous one, each theoretical concept is immediately illustrated by diverse applications presented in the complements. Both volumes I and II have benefited from several recent corrections, but there have also been additions. Chapter XIII now contains two sections §§ D and E that treat random perturbations, and a complement concerning relaxation has been added.

Volume III extends the two volumes at a slightly higher level. It is based on the use of the creation and annihilation operator formalism (second quantization), which is commonly used in quantum field theory. We start with a study of systems of identical particles, fermions or bosons. The properties of ideal gases in thermal equilibrium are presented. For fermions, the Hartree-Fock method is developed in detail. It is the base of many studies in chemistry, atomic physics and solid state physics, etc. For bosons, the Gross-Pitaevskii equation and the Bogolubov theory are discussed. An original presentation that treats the pairing effect of both fermions and bosons permits obtaining the BCS (Bardeen-Cooper-Schrieffer) and Bogolubov theories in a unified framework. The second part of volume III treats quantum electrodynamics, its general introduction, the study of interactions between atoms and photons, and various applications (spontaneous emission, multiphoton transitions, optical pumping, etc.). The dressed atom method is presented and illustrated for concrete cases. A final chapter discusses the notion of quantum entanglement and certain fundamental aspects of quantum mechanics, in particular the Bell inequalities and their violations.

Finally note that we have not treated either the philosophical implications of quantum mechanics, or the diverse interpretations of this theory, despite the great interest of these subjects. We have in fact limited ourselves to presenting what is commonly called the "orthodox point of view". It is only in Chapter XXI that we touch on certain questions concerning the foundations of quantum mechanics (nonlocality, etc.). We have made this choice because we feel that one can address such questions more efficiently after mastering the manipulation of the quantum mechanical formalism as well as its numerous applications. These subjects are addressed in the book *Do we really understand quantum mechanics?* (F. Laloë, Cambridge University Press, 2019); see also section 5 of the bibliography of volumes I and II.

Acknowledgments:

Volumes I and II:

The teaching experience out of which this text grew were group efforts, pursued over several years. We wish to thank all the members of the various groups and particularly Jacques Dupont-Roc and Serge Haroche, for their friendly collaboration, for the fruitful discussions we have had in our weekly meetings and for the ideas for problems and exercises that they have suggested. Without their enthusiasm and valuable help, we would never have been able to undertake and carry out the writing of this book.

Nor can we forget what we owe to the physicists who introduced us to research, Alfred Kastler and Jean Brossel for two of us and Maurice Levy for the third. It was in the context of their laboratories that we discovered the beauty and power of quantum mechanics. Neither have we forgotten the importance to us of the modern physics taught at the C.E.A. by Albert Messiah, Claude Bloch and Anatole Abragam, at a time when graduate studies were not yet incorporated into French university programs.

We wish to express our gratitude to Ms. Aucher, Baudrit, Boy, Brodschi, Emo, Heywaerts, Lemirre, Touzeau for preparation of the mansucript.

Volume III:

We are very grateful to Nicole and Daniel Ostrowsky, who, as they translated this Volume from French into English, proposed numerous improvements and clarifications. More recently, Carsten Henkel also made many useful suggestions during his translation of the text into German; we are very grateful for the improvements of the text that resulted from this exchange. There are actually many colleagues and friends who greatly contributed, each in his own way, to finalizing this book. All their complementary remarks and suggestions have been very helpful and we are in particular thankful to:

Pierre-François Cohadon
Jean Dalibard
Sébastien Gleyzes
Markus Holzmann
Thibaut Jacqmin
Philippe Jacquier
Amaury Mouchet
Jean-Michel Raimond
Félix Werner

Some delicate aspects of Latex typography have been resolved thanks to Marco Picco, Pierre Cladé and Jean Hare. Roger Balian, Edouard Brézin and William Mullin have offered useful advice and suggestions. Finally, our sincere thanks go to Geneviève Tastevin, Pierre-François Cohadon and Samuel Deléglise for their help with a number of figures.

Volume I

Table of contents vii

I WAVES AND PARTICLES. INTRODUCTION TO THE BASIC IDEAS OF QUANTUM MECHANICS 1

READER'S GUIDE FOR COMPLEMENTS 33

A_I Order of magnitude of the wavelengths associated with material particles 35

B_I Constraints imposed by the uncertainty relations 39

C_I Heisenberg relation and atomic parameters 41

D_I An experiment illustrating the Heisenberg relations 45

E_I A simple treatment of a two-dimensional wave packet 49

F_I The relationship between one- and three-dimensional problems 53

G_I One-dimensional Gaussian wave packet: spreading of the wave packet 57

H_I Stationary states of a particle in one-dimensional square potentials 63

J_I Behavior of a wave packet at a potential step 75

K_I Exercises 83

****

II THE MATHEMATICAL TOOLS OF QUANTUM MECHANICS 87

READER'S GUIDE FOR COMPLEMENTS 159

A_{II} The Schwarz inequality 161

B_{II} Review of some useful properties of linear operators 163

C_{II} Unitary operators 173

D_{II} A more detailed study of the r and p representations 181

E_{II} Some general properties of two observables, Q and P, whose commutator is equal to $i\hbar$ 187

F_{II} The parity operator 193

G_{II} An application of the properties of the tensor product: the two-dimensional infinite well **201**

H_{II} Exercises **205**

III THE POSTULATES OF QUANTUM MECHANICS **213**

READER'S GUIDE FOR COMPLEMENTS **267**

A_{III} Particle in an infinite one-dimensional potential well **271**

B_{III} Study of the probability current in some special cases **283**

C_{III} Root mean square deviations of two conjugate observables **289**

D_{III} Measurements bearing on only one part of a physical system **293**

E_{III} The density operator **299**

F_{III} The evolution operator **313**

G_{III} The Schrödinger and Heisenberg pictures **317**

H_{III} Gauge invariance **321**

J_{III} Propagator for the Schrödinger equation **335**

K_{III} Unstable states. Lifetime **343**

L_{III} Exercises **347**

M_{III} Bound states in a "potential well" of arbitrary shape **359**

N_{III} Unbound states of a particle in the presence of a potential well or barrier **367**

O_{III} Quantum properties of a particle in a one-dimensional periodic structure **375**

****

IV APPLICATIONS OF THE POSTULATES TO SIMPLE CASES: SPIN 1/2 AND TWO-LEVEL SYSTEMS **393**

READER'S GUIDE FOR COMPLEMENTS **423**

A_{IV} The Pauli matrices **425**

B_{IV} Diagonalization of a 2×2 Hermitian matrix **429**

C_{IV} Fictitious spin 1/2 associated with a two-level system **435**

D_{IV} System of two spin 1/2 particles **441**

E_{IV} Spin $1/2$ density matrix **449**

F_{IV} Spin 1/2 particle in a static and a rotating magnetic fields: magnetic resonance **455**

G_{IV} A simple model of the ammonia molecule **469**

H_{IV} Effects of a coupling between a stable state and an unstable state **485**

J_{IV} Exercises **491**

****

V THE ONE-DIMENSIONAL HARMONIC OSCILLATOR **497**

READER'S GUIDE FOR COMPLEMENTS **525**

A_V Some examples of harmonic oscillators **527**

B_V Study of the stationary states in the x representation. Hermite polynomials **547**

C_V Solving the eigenvalue equation of the harmonic oscillator by the polynomial method **555**

D_V Study of the stationary states in the momentum representation **563**

E_V The isotropic three-dimensional harmonic oscillator **569**

F_V A charged harmonic oscillator in a uniform electric field **575**

G_V Coherent "quasi-classical" states of the harmonic oscillator **583**

H_V Normal vibrational modes of two coupled harmonic oscillators **599**

J_V Vibrational modes of an infinite linear chain of coupled harmonic oscillators; phonons **611**

K_V Vibrational modes of a continuous physical system. Photons **631**

L_V One-dimensional harmonic oscillator in thermodynamic equilibrium at a temperature T **647**

M_V Exercises **661**

****

VI GENERAL PROPERTIES OF ANGULAR MOMENTUM IN QUANTUM MECHANICS **667**

READER'S GUIDE FOR COMPLEMENTS 703

A_{VI} Spherical harmonics 705

B_{VI} Angular momentum and rotations 717

C_{VI} Rotation of diatomic molecules 739

D_{VI} Angular momentum of stationary states of a two-dimensional harmonic oscillator 755

E_{VI} A charged particle in a magnetic field: Landau levels 771

F_{VI} Exercises 795

****

VII PARTICLE IN A CENTRAL POTENTIAL, HYDROGEN ATOM 803

READER'S GUIDE FOR COMPLEMENTS 831

A_{VII} Hydrogen-like systems 833

B_{VII} A soluble example of a central potential: The isotropic three-dimensional harmonic oscillator 841

C_{VII} Probability currents associated with the stationary states of the hydrogen atom 851

D_{VII} The hydrogen atom placed in a uniform magnetic field. Paramagnetism and diamagnetism. The Zeeman effect 855

E_{VII} Some atomic orbitals. Hybrid orbitals 869

F_{VII} Vibrational-rotational levels of diatomic molecules 885

G_{VII} Exercises 899

INDEX 901

****

Volume II

VOLUME II **923**

Table of contents **v**

VIII AN ELEMENTARY APPROACH TO THE QUANTUM THEORY OF SCATTERING BY A POTENTIAL **923**

READER'S GUIDE FOR COMPLEMENTS **957**

A_{VIII} The free particle: stationary states with well-defined angular momentum **959**

B_{VIII} Phenomenological description of collisions with absorption **971**

C_{VIII} Some simple applications of scattering theory **977**

****

IX ELECTRON SPIN **985**

READER'S GUIDE FOR COMPLEMENTS **999**

A_{IX} Rotation operators for a spin 1/2 particle **1001**

B_{IX} Exercises **1009**

****

X ADDITION OF ANGULAR MOMENTA **1015**

READER'S GUIDE FOR COMPLEMENTS **1041**

A_X Examples of addition of angular momenta **1043**

B_X Clebsch-Gordan coefficients **1051**

C_X Addition of spherical harmonics **1059**

D_X Vector operators: the Wigner-Eckart theorem **1065**

E_X Electric multipole moments **1077**

F_X Two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ coupled by an interaction $a\mathbf{J}_1 \cdot \mathbf{J}_2$ **1091**

G_X Exercises **1107**

****

XI STATIONARY PERTURBATION THEORY 1115

READER'S GUIDE FOR COMPLEMENTS 1129

A_{XI} A one-dimensional harmonic oscillator subjected to a perturbing potential in x, x^2, x^3 1131

B_{XI} Interaction between the magnetic dipoles of two spin 1/2 particles 1141

C_{XI} Van der Waals forces 1151

D_{XI} The volume effect: the influence of the spatial extension of the nucleus on the atomic levels 1162

E_{XI} The variational method 1169

F_{XI} Energy bands of electrons in solids: a simple model 1177

G_{XI} A simple example of the chemical bond: the H_2^+ ion 1189

H_{XI} Exercises 1221

****

XII AN APPLICATION OF PERTURBATION THEORY: THE FINE AND HYPERFINE STRUCTURE OF THE HYDROGEN ATOM 1231

READER'S GUIDE FOR COMPLEMENTS 1265

A_{XII} The magnetic hyperfine Hamiltonian 1267

B_{XII} Calculation of the average values of the fine-structure Hamiltonian in the $1s$, $2s$ and $2p$ states 1276

C_{XII} The hyperfine structure and the Zeeman effect for muonium and positronium 1281

D_{XII} The influence of the electronic spin on the Zeeman effect of the hydrogen resonance line 1289

E_{XII} The Stark effect for the hydrogen atom 1298

****

XIII APPROXIMATION METHODS FOR TIME-DEPENDENT PROBLEMS 1303

READER'S GUIDE FOR COMPLEMENTS 1337

A_{XIII} Interaction of an atom with an electromagnetic wave 1339

B_{XIII} Linear and non-linear responses of a two-level system subject to a sinusoidal perturbation 1357

C_{XIII} Oscillations of a system between two discrete states under the effect of a sinusoidal resonant perturbation 1374

D_{XIII} Decay of a discrete state resonantly coupled to a continuum of final states 1378

E_{XIII} Time-dependent random perturbation, relaxation 1390

F_{XIII} Exercises 1409

****

XIV SYSTEMS OF IDENTICAL PARTICLES 1419

READER'S GUIDE FOR COMPLEMENTS 1457

A_{XIV} Many-electron atoms. Electronic configurations 1459

B_{XIV} Energy levels of the helium atom. Configurations, terms, multiplets 1467

C_{XIV} Physical properties of an electron gas. Application to solids 1481

D_{XIV} Exercises 1496

****

APPENDICES 1505

I Fourier series and Fourier transforms 1505

II The Dirac δ-"function" 1515

III Lagrangian and Hamiltonian in classical mechanics 1527

BIBLIOGRAPHY OF VOLUMES I AND II 1545

INDEX 1569

****

Volume III

VOLUME III **1591**

Table of contents **v**

XV CREATION AND ANNIHILATION OPERATORS FOR IDENTICAL PARTICLES **1591**
A General formalism . . . 1592
B One-particle symmetric operators . . . 1603
C Two-particle operators . . . 1608

READER'S GUIDE FOR COMPLEMENTS **1617**

A_{XV} Particles and holes **1621**
1 Ground state of a non-interacting fermion gas . . . 1621
2 New definition for the creation and annihilation operators . . . 1622
3 Vacuum excitations . . . 1623

B_{XV} Ideal gas in thermal equilibrium; quantum distribution functions **1625**
1 Grand canonical description of a system without interactions . . . 1626
2 Average values of symmetric one-particle operators . . . 1628
3 Two-particle operators . . . 1631
4 Total number of particles . . . 1635
5 Equation of state, pressure . . . 1640

C_{XV} Condensed boson system, Gross-Pitaevskii equation **1643**
1 Notation, variational ket . . . 1643
2 First approach . . . 1645
3 Generalization, Dirac notation . . . 1648
4 Physical discussion . . . 1651

D_{XV} Time-dependent Gross-Pitaevskii equation **1657**
1 Time evolution . . . 1657
2 Hydrodynamic analogy . . . 1664
3 Metastable currents, superfluidity . . . 1667

E_{XV} Fermion system, Hartree-Fock approximation **1677**
1 Foundation of the method . . . 1678
2 Generalization: operator method . . . 1688

F_{XV} Fermions, time-dependent Hartree-Fock approximation **1701**
1 Variational ket and notation . . . 1701
2 Variational method . . . 1702
3 Computing the optimizer . . . 1705
4 Equations of motion . . . 1707

G_{XV} Fermions or Bosons: Mean field thermal equilibrium **1711**
1 Variational principle . . . 1712
2 Approximation for the equilibrium density operator . . . 1716
3 Temperature dependent mean field equations . . . 1725

H_{XV} Applications of the mean field method for non-zero temperature **1733**
1 Hartree-Fock for non-zero temperature, a brief review . . . 1733
2 Homogeneous system . . . 1734
3 Spontaneous magnetism of repulsive fermions . . . 1737
4 Bosons: equation of state, attractive instability . . . 1745

****

XVI FIELD OPERATOR **1751**
A Definition of the field operator . . . 1752
B Symmetric operators . . . 1755
C Time evolution of the field operator (Heisenberg picture) . . . 1763
D Relation to field quantization . . . 1765

READER'S GUIDE FOR COMPLEMENTS **1767**

A_{XVI} Spatial correlations in an ideal gas of bosons or fermions **1769**
1 System in a Fock state . . . 1769
2 Fermions in the ground state . . . 1771
3 Bosons in a Fock state . . . 1775

B_{XVI} Spatio-temporal correlation functions, Green's functions **1781**
1 Green's functions in ordinary space . . . 1781
2 Fourier transforms . . . 1790
3 Spectral function, sum rule . . . 1795

C_{XVI} Wick's theorem **1799**
1 Demonstration of the theorem . . . 1799
2 Applications: correlation functions for an ideal gas . . . 1804

****

XVII PAIRED STATES OF IDENTICAL PARTICLES **1811**
A Creation and annihilation operators of a pair of particles . . . 1813
B Building paired states . . . 1818
C Properties of the kets characterizing the paired states . . . 1822
D Correlations between particles, pair wave function . . . 1830
E Paired states as a quasi-particle vacuum; Bogolubov-Valatin transformations 1836

READER'S GUIDE FOR COMPLEMENTS 1843

A_{XVII} Pair field operator for identical particles 1845
1 Pair creation and annihilation operators . . . 1846
2 Average values in a paired state . . . 1851
3 Commutation relations of field operators . . . 1861

B_{XVII} Average energy in a paired state 1869
1 Using states that are not eigenstates of the total particle number . . . 1869
2 Hamiltonian . . . 1871
3 Spin 1/2 fermions in a singlet state . . . 1874
4 Spinless bosons . . . 1881

C_{XVII} Fermion pairing, BCS theory 1889
1 Optimization of the energy . . . 1890
2 Distribution functions, correlations . . . 1899
3 Physical discussion . . . 1914
4 Excited states . . . 1919

D_{XVII} Cooper pairs 1927
1 Cooper model . . . 1927
2 State vector and Hamiltonian . . . 1927
3 Solution of the eigenvalue equation . . . 1929
4 Calculation of the binding energy for a simple case . . . 1929

E_{XVII} Condensed repulsive bosons 1933
1 Variational state, energy . . . 1935
2 Optimization . . . 1937
3 Properties of the ground state . . . 1940
4 Bogolubov operator method . . . 1950

****

XVIII REVIEW OF CLASSICAL ELECTRODYNAMICS 1957
A Classical electrodynamics . . . 1959
B Describing the transverse field as an ensemble of harmonic oscillators . . . 1968

READER'S GUIDE FOR COMPLEMENTS 1977

A_{XVIII} Lagrangian formulation of electrodynamics 1979
1 Lagrangian with several types of variables . . . 1980
2 Application to the free radiation field . . . 1986
3 Lagrangian of the global system field + interacting particles . . . 1992

****

XIX QUANTIZATION OF ELECTROMAGNETIC RADIATION 1997
A Quantization of the radiation in the Coulomb gauge 1999
B Photons, elementary excitations of the free quantum field 2004
C Description of the interactions 2009

READER'S GUIDE FOR COMPLEMENTS 2017

A_{XIX} Momentum exchange between atoms and photons 2019
1 Recoil of a free atom absorbing or emitting a photon 2020
2 Applications of the radiation pressure force: slowing and cooling atoms . . . 2025
3 Blocking recoil through spatial confinement 2036
4 Recoil suppression in certain multi-photon processes 2040

B_{XIX} Angular momentum of radiation 2043
1 Quantum average value of angular momentum for a spin 1 particle 2044
2 Angular momentum of free classical radiation as a function of normal variables 2047
3 Discussion 2050

C_{XIX} Angular momentum exchange between atoms and photons 2055
1 Transferring spin angular momentum to internal atomic variables 2056
2 Optical methods 2058
3 Transferring orbital angular momentum to external atomic variables 2065

****

XX ABSORPTION, EMISSION AND SCATTERING OF PHOTONS BY ATOMS 2067
A A basic tool: the evolution operator 2068
B Photon absorption between two discrete atomic levels 2073
C Stimulated and spontaneous emissions 2080
D Role of correlation functions in one-photon processes 2084
E Photon scattering by an atom 2085

READER'S GUIDE FOR COMPLEMENTS 2095

A_{XX} A multiphoton process: two-photon absorption 2097
1 Monochromatic radiation 2097
2 Non-monochromatic radiation 2101
3 Discussion 2105

B_{XX} Photoionization 2109
1 Brief review of the photoelectric effect 2110
2 Computation of photoionization rates 2112
3 Is a quantum treatment of radiation necessary to describe photoionization? . 2118
4 Two-photon photoionization 2123
5 Tunnel ionization by intense laser fields 2126

C_{XX} Two-level atom in a monochromatic field. Dressed-atom method 2129
1 Brief description of the dressed-atom method . . . 2130
2 Weak coupling domain . . . 2137
3 Strong coupling domain . . . 2141
4 Modifications of the field. Dispersion and absorption . . . 2147

D_{XX} Light shifts: a tool for manipulating atoms and fields 2151
1 Dipole forces and laser trapping . . . 2151
2 Mirrors for atoms . . . 2153
3 Optical lattices . . . 2153
4 Sub-Doppler cooling. Sisyphus effect . . . 2155
5 Non-destructive detection of a photon . . . 2159

E_{XX} Detection of one- or two-photon wave packets, interference 2163
1 One-photon wave packet, photodetection probability . . . 2165
2 One- or two-photon interference signals . . . 2167
3 Absorption amplitude of a photon by an atom . . . 2174
4 Scattering of a wave packet . . . 2176
5 Example of wave packets with two entangled photons . . . 2181

****

XXI QUANTUM ENTANGLEMENT, MEASUREMENTS, BELL'S INEQUALITIES 2187
A Introducing entanglement, goals of this chapter . . . 2188
B Entangled states of two spin-1/2 systems . . . 2190
C Entanglement between more general systems . . . 2193
D Ideal measurement and entangled states . . . 2196
E "Which path" experiment: can one determine the path followed by the photon in Young's double slit experiment? . . . 2202
F Entanglement, non-locality, Bell's theorem . . . 2204

READER'S GUIDE FOR COMPLEMENTS 2215

A_{XXI} Density operator and correlations; separability 2217
1 Von Neumann statistical entropy . . . 2217
2 Differences between classical and quantum correlations . . . 2221
3 Separability . . . 2223

B_{XXI} GHZ states, entanglement swapping 2227
1 Sign contradiction in a GHZ state . . . 2227
2 Entanglement swapping . . . 2232

C_{XXI} Measurement induced relative phase between two condensates 2237
1 Probabilities of single, double, etc. position measurements . . . 2239
2 Measurement induced enhancement of entanglement . . . 2242
3 Detection of a large number Q of particles . . . 2245

D_{XXI} Emergence of a relative phase with spin condensates; macroscopic non-locality and the EPR argument **2253**
1 Two condensates with spins 2254
2 Probabilities of the different measurement results 2255
3 Discussion 2259

****

APPENDICES **2267**

IV Feynman path integral **2267**
1 Quantum propagator of a particle 2267
2 Interpretation in terms of classical histories 2272
3 Discussion; a new quantization rule 2274
4 Operators 2276

V Lagrange multipliers **2281**
1 Function of two variables 2281
2 Function of N variables 2283

VI Brief review of Quantum Statistical Mechanics **2285**
1 Statistical ensembles 2285
2 Intensive or extensive physical quantities 2292

VII Wigner transform **2297**
1 Delta function of an operator 2299
2 Wigner distribution of the density operator (spinless particle) 2299
3 Wigner transform of an operator 2310
4 Generalizations 2318
5 Discussion: Wigner distribution and quantum effects 2319

BIBLIOGRAPHY OF VOLUME III **2325**

INDEX **2333**

Chapter XV

Creation and annihilation operators for identical particles

A General formalism 1592
A-1 Fock states and Fock space 1593
A-2 Creation operators $a^\dagger$ 1596
A-3 Annihilation operators a 1597
A-4 Occupation number operators (bosons and fermions) 1598
A-5 Commutation and anticommutation relations 1599
A-6 Change of basis 1601
B One-particle symmetric operators 1603
B-1 Definition 1603
B-2 Expression in terms of the operators a and $a^\dagger$ 1604
B-3 Examples 1606
B-4 Single particle density operator 1607
C Two-particle operators 1608
C-1 Definition 1608
C-2 A simple case: factorization 1609
C-3 General case 1610
C-4 Two-particle reduced density operator 1610
C-5 Physical discussion; consequences of the exchange 1611

Introduction

For a system composed of identical particles, the particle numbering used in Chapter XIV, the last chapter of Volume II [2], does not really have much physical significance. Furthermore, when the particle number gets larger than a few units, applying the symmetrization postulate to numbered particles often leads to complex calculations. For

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

example, computing the average value of a symmetric operator requires the symmetrization of the bra, the ket, and finally the operator, which introduces a large number of terms[1]. They seem different, a priori, but at the end of the computation many are found to be equal, or sometimes cancel each other. Fortunately, these lengthy calculations may be avoided using an equivalent method based on creation and annihilation operators in a "Fock space". The simple commutation (or anticommutation) rules satisfied by these operators are the expression of the symmetrization (or antisymmetrization) postulate. The non-physical particle numbering is replaced by assigning "occupation numbers" to individual states, which is more natural for treating identical particles.

The method described in this chapter and the following is sometimes called "second quantization"[2]. It deals with operators that no longer conserve the particle number, hence acting in a state space larger than those we have previously considered; this new space is called the "Fock space" (§ A). These operators which change the particle number appear mainly in the course of calculations, and often regroup at the end, keeping the total particle number constant. Examples will be given (§ B) for one-particle symmetric operators, such as the total linear momentum or angular momentum of a system of identical particles. We shall then study two-particle symmetric operators (§ C), such as the energy of a system of interacting identical particles, their spatial correlation function, etc. In quantum statistical mechanics, the Fock space is well adapted to computations performed in the "grand canonical" ensemble, where the total number of particles may fluctuate since the system is in contact with an external reservoir. Furthermore, as we shall see in the following chapters, the Fock space is very useful for describing physical processes where the particle number changes, as in photon absorption or emission.

A. General formalism

We denote $\mathcal{E}_N$ the state space of a system of N distinguishable particles, which is the tensor product of N individual state spaces $\mathcal{E}_1$:

$$\mathcal{E}_N = \mathcal{E}_1(1) \otimes \mathcal{E}_1(2) \otimes .. \otimes \mathcal{E}_1(N) \tag{A-1}$$

Two sub-spaces of $\mathcal{E}_N$ are particularly important for identical particles, as they contain all their accessible physical states: the space $\mathcal{E}_S(N)$ of the completely symmetric states for bosons, and the space $\mathcal{E}_A(N)$ of the completely antisymmetric states for fermions. The projectors onto these two sub-spaces are given by relations (B-49) and (B-50) of Chapter XIV:

$$S_N = \frac{1}{N!} \sum_\alpha P_\alpha \tag{A-2}$$

and:

$$A_N = \frac{1}{N!} \sum_\alpha \varepsilon_\alpha P_\alpha \tag{A-3}$$

[1] For a one-particle symmetric operator, which includes the sum of N terms, both the ket and bra contain $N!$ terms. The matrix element will therefore involve $N\,(N!)^2$ terms, a very large number once N exceeds a few units.

[2] A commonly accepted but a somewhat illogical expression, since no new quantification comes in addition to that of the usual postulates of Quantum Mechanics; its essential ingredient is the symmetrization of identical particles.

where the P_α are the $N!$ permutation operators for the N particles, and ε_α the parity of P_α (in this chapter we have added for clarity the index N to the projectors S and A defined in Chapter XIV).

A-1. Fock states and Fock space

Starting from an arbitrary orthonormal basis $\{|u_k\rangle\}$ of the state space for one particle, we constructed in § C-3-d of Chapter XIV a basis of the state space for N identical particles. Its vectors are characterized by the occupation numbers n_i, with:

$$n_1 + n_2 + .. + n_k + .. = N \tag{A-4}$$

where n_1 is the occupation number of the first basis vector $|u_1\rangle$ (i.e. the number of particles in $|u_1\rangle$), n_2 that of $|u_2\rangle$, ..,n_k that of $|u_k\rangle$. In this series of numbers, some (even many) may be zero: a given state has no particular reason to always be occupied. It is therefore often easier to specify only the non-zero occupation numbers, which will be noted $n_i, n_j, .., n_l, ..$. This series indicates that the first basis state that has at least one particle is $|u_i\rangle$ and it contains n_i particles; the second occupied state is $|u_j\rangle$ with a population n_j, etc. As in (A-4), these occupation numbers add up to N.

Comment:

In this chapter we constantly use subscripts of different types, which should not be confused. The subscripts i, j, k, l, ..denote different basis vectors $\{|u_i\rangle\}$ of the state space $\mathcal{E}_1$ of a single particle; they span values given by the dimension of this state space, which often goes to infinity. They should not be confused with the subscripts used to number the particles, which can take N different values, and are labeled q, q', etc. Finally the subscript α distinguishes the different permutations of the N particles, and can therefore take $N!$ different values.

A-1-a. Fock states for identical bosons

For bosons, the basis vectors can be written as in (C-15) of Chapter XIV:

$$\begin{aligned} &|n_i, n_j, .., n_l, ..\rangle \\ &\quad = c\ S_N\, |1 : u_i;\ 2 : u_i;\ ..;\ n_i : u_i;\ \ n_i + 1 : u_j;\ ..\ \ n_i + n_j : u_j;\ ..\rangle \end{aligned} \tag{A-5}$$

where c is a normalization constant; on the right-hand side, n_i particles occupy the state $|u_i\rangle$, n_j the state $|u_j\rangle$, etc... (because of symmetrization, their order does not matter).

Let us calculate the norm of the right-hand term. It is composed of $N!$ terms, coming from each of the $N!$ permutations included in S_N, but only some of them are orthogonal to each other: all the permutations P_α leading to redistributions of the n_i first particles among themselves, of the next n_j particles among themselves, etc. yield the same initial ket. On the other hand, if a permutation changes the individual state of one (or more than one) particle, it yields a different ket, actually orthogonal to the initial ket. This means that the different permutations contained in S_N can be grouped into families of $n_i!n_j!..n_l!..$ equivalent permutations, all yielding the same ket; taking into account the factor $N!$ appearing in the definition of S_N, the coefficient in front of this ket becomes $n_i!n_j!..n_l!../N!$ and its contribution to the norm of the ket is equal to the square of this number. On the other hand, the number of orthogonal kets is

$N!/n_i!n_j!..n_l!..$ Consequently if c was equal to 1 in formula (A-5), the ket thus defined would have a norm equal to:

$$\frac{N!}{n_i!n_j!..n_l!..}\left[\frac{1}{N!}n_i!n_j!..n_l!..\right]^2 = \frac{n_i!n_j!..n_l!..}{N!} \tag{A-6}$$

We shall therefore choose for c the inverse of the square root of that number, leading to the normalized ket:

$$|n_i,\ n_j,..,n_l,..\rangle = \sqrt{\frac{N!}{n_i!n_j!..n_l!..}}\ S_N\,|1:u_i;\ 2:u_i;\ ..;\ n_i:u_i;\ \ n_i+1:u_j;..;\ n_i+n_j:u_j;\ ..\rangle \tag{A-7}$$

These states are called the "*Fock states*", for which the occupation numbers are well defined.

For the Fock states, it is sometimes handy to use a slightly different but equivalent notation. In (A-7), these states are defined by specifying the occupation numbers of all the states that are actually occupied ($n_i \geq 1$). Another option would be to indicate all the occupation numbers including those which are zero [3] – this is what we have explicitly done in (A-4). We then write the same kets as:

$$|n_1, n_2, .., n_i, .., n_l, ..\rangle \tag{A-8}$$

Another possibility is to specify a list of N occupied states, where u_i is repeated n_i times, u_j repeated n_j times, etc. :

$$|\underbrace{u_i, u_i,\ \ldots}_{n_i\text{-times}}, \underbrace{u_j, u_j,\ \ldots}_{n_j\text{-times}},\ \ldots\ , u_l,\ \ldots\rangle \tag{A-9}$$

As we shall see later, this latter notation is sometimes useful in computations involving both bosons and fermions.

A-1-b. Fock states for identical fermions

In the case of fermions, the operator A_N acting on a ket where two (or more) numbered particles are in the same individual state yields a zero result: there are no such states in the physical space $\mathcal{E}_A(N)$. Hence we concentrate on the case where all the occupation numbers are either 1 or 0. We denote $|u_i\rangle$, $|u_j\rangle$,..,$|u_l\rangle$, .. all the states having an occupation number equal to 1. The equivalent for fermions of formula (A-7) is written:

$$|u_i, u_j, .., u_l, ..\rangle = \begin{cases} \sqrt{N!}\ A_N\,|1:u_i;\ 2:u_j;\ ..;\ q:u_l;\ ..\rangle & \text{if all the } u_i \text{ are different} \\ 0 & \text{if two } u_i \text{ are identical} \end{cases} \tag{A-10}$$

[3] Remember that, by convention, $0! = 1$.

Taking into account the $1/N!$ factor appearing in definition (A-3) of A_N, the right-hand side of this equation is a linear superposition, with coefficients $1/\sqrt{N!}$, of $N!$ kets which are all orthogonal to each other (as we have chosen an orthonormal basis for the individual states $\{|u_k\rangle\}$); hence its norm is equal to 1. Consequently, Fock states for fermions are defined by (A-10). Contrary to bosons, the main concern is no longer how many particles occupy a state, but whether a state is occupied or not. Another difference with the boson case is that, for fermions, the order of the states matters. If for instance the first two states u_i and u_j are exchanged, we get the opposite ket:

$$|u_j, u_i, .., u_l, ..\rangle = -\,|u_i, u_j, .., u_l, ..\rangle \tag{A-11}$$

but it obviously does not change the physical meaning of the ket.

A-1-c. Fock space

The Fock states are the building blocks used to construct this whole chapter. We have until now considered separately the spaces $\mathcal{E}_{S,A}(N)$ associated with different values of the particle number N. We shall now regroup them into a single space, called the "*Fock space* ", using the direct sum[4] formalism. For bosons:

$$\mathcal{E}^S_{\rm Fock} = \mathcal{E}_S(0) \oplus \mathcal{E}_S(1) \oplus \mathcal{E}_S(2) \oplus .. \oplus \mathcal{E}_S(N) \oplus .. \tag{A-12}$$

and, for fermions:

$$\mathcal{E}^A_{\rm Fock} = \mathcal{E}_A(0) \oplus \mathcal{E}_A(1) \oplus \mathcal{E}_A(2) \oplus .. \oplus \mathcal{E}_A(N) \oplus .. \tag{A-13}$$

(the sums go to infinity). In both cases, we have included on the right-hand side a first term associated with a total number of particles equal to zero. The corresponding space, $\mathcal{E}_{S,A}(0)$, is defined as a one-dimensional space, containing a single state called "vacuum" and denoted $|0\rangle$ or $|\text{vac}\rangle$. For bosons as well as fermions, an orthonormal basis for the Fock space can be built with the Fock states $|n_1,\ n_2, ..., n_i, ., n_l..\rangle$, relaxing the constraint (A-4): the occupation numbers may then take on any (integer) values, including zeros for all, which corresponds to the vacuum ket $|0\rangle$. Linear combinations of all these basis vectors yield all the vectors of the Fock space, including linear superpositions of kets containing different particle numbers. It is not essential to attribute a physical interpretation to such superpositions since they can be considered as intermediate states of the calculation. Obviously, the Fock space contains many kets with well defined particle numbers: all those belonging to a single sub-space $\mathcal{E}_S(N)$ for bosons, or $\mathcal{E}_A(N)$ for fermions. Two kets having different particle numbers N are necessarily orthogonal; for example, all the kets having a non-zero total population are orthogonal to the vacuum state.

Comments:

(i) Contrary to the distinguishable particle case, the Fock space is not the tensor product of the spaces of states associated with particles numbered 1, 2,..., q, etc. First of all, for a

[4]The direct sum of two spaces $\mathcal{E}_P$ (with dimension P) and $\mathcal{E}_Q$ (with dimension Q) is a space $\mathcal{E}_{P+Q}$ with dimension $P+Q$, spanned by all the linear combinations of a vector from the first space with a vector from the second. A basis for $\mathcal{E}_{P+Q}$ may be simply obtained by grouping together a basis for $\mathcal{E}_P$ and one for $\mathcal{E}_Q$. For example, vectors of a two-dimensional plane belong to a space that is the direct sum of the one-dimensional spaces for the vectors of two axes of that plane.

fixed N, it only includes the totally symmetric (or antisymmetric) subspace of this tensor product; furthermore, the Fock space is the direct sum of such subspaces associated with each value of the particle number N.

The Fock space is, however, the tensor product of Fock spaces $\mathcal{E}_{\rm Fock}^{u_i}$ associated with the individual orthogonal states $|u_i\rangle$, each $\mathcal{E}_{\rm Fock}^{u_i}$ being spanned by the kets $|n_i\rangle$ where n_i takes on all integer values (from zero to infinity for bosons, from zero to one for fermions):

$$\mathcal{E}_{\rm Fock}^{S,A} = \mathcal{E}_{\rm Fock}^{u_1} \otimes \mathcal{E}_{\rm Fock}^{u_2} \otimes .. \otimes \mathcal{E}_{\rm Fock}^{ui} \otimes .. \tag{A-14}$$

This is because the Fock states, which are a basis for $\mathcal{E}_{\rm Fock}^{S,A}$, may be written as the tensor product:

$$|n_1,\ n_2, .., n_i, ..\rangle = |n_1\rangle \otimes |n_2\rangle \otimes .. \otimes |n_i\rangle \otimes .. \tag{A-15}$$

It is often said that each individual state defines a "mode" of the system of identical particles. Decomposing the Fock state into a tensor product allows considering the modes as describing different and distinguishable variables. This will be useful on numerous occasions (see for example Complements $\rm B_{XV}$, $\rm D_{XV}$ and $\rm E_{XV}$).

(ii) One should not confuse a Fock state with an arbitrary state of the Fock space. The occupation numbers of individual states are all well defined in a Fock state (also called "number state"), whereas an arbitrary state of the Fock space is a linear superposition of these eigenstates, with several non-zero coefficients.

A-2. Creation operators $a^\dagger$

Choosing a basis of individual states $\{|u_i\rangle\}$, we now define the action in the Fock space of the creation operator[5] $a_{u_i}^\dagger$ on a particle in the state $|u_i\rangle$.

A-2-a. Bosons

For bosons, we introduce the linear operator $a_{u_i}^\dagger$ defined by:

$$a_{u_i}^\dagger\ |n_1, n_2, .., n_i, ..\rangle = \sqrt{n_i + 1}\ |n_1, n_2, .., n_i + 1, ..\rangle \tag{A-16}$$

As all the states of the Fock space may be obtained by a linear superposition of $|n_1,\ n_2, .., n_i, ..\rangle$, the action of $a_{u_i}^\dagger$ is defined in the entire space. It adds a particle to the system, which goes from a state of $\mathcal{E}_S(N)$ to a state of $\mathcal{E}_S(N+1)$, and in particular from the vacuum to a state having one single occupied state.

Creation operators acting on the vacuum allow building occupied states. Recurrent application of (A-16), leads to:

$$|n_1, n_2, .., n_i, ..\rangle = \frac{1}{\sqrt{n_1! n_2! .. n_i! ..}} \left[a_{u_1}^\dagger\right]^{n_1} \left[a_{u_2}^\dagger\right]^{n_2} .. \left[a_{u_i}^\dagger\right]^{n_i} .. |0\rangle \tag{A-17}$$

Comment:

Why was the factor $\sqrt{n_i + 1}$ introduced in (A-16)? We shall see later (§ B) that, together with the factors of (A-7), it simplifies the computations.

[5] A similar notation was used for the harmonic oscillator.

A-2-b. Fermions

For fermions, we define the operator $a^\dagger_{u_i}$ by:

$$a^\dagger_{u_i} \ |u_j, .., u_k, .., u_l, ..\rangle = |u_i, u_j, .., u_k, .., u_l, ..\rangle \tag{A-18}$$

where the newly created state $|u_i\rangle$ appears first in the list of states in the ket on the right-hand side. If we start from a ket where the individual state $|u_i\rangle$ is already occupied ($n_i = 1$), the action of $a^\dagger_{u_i}$ leads to zero, as in this case (A-10) gives:

$$a^\dagger_{u_i} \ |u_j, .., u_i, .., u_l, ..\rangle = |u_i, u_j, .., u_i, .., u_l, ..\rangle = 0 \tag{A-19}$$

Formulas (A-16) and (A-17) are also valid for fermions, with all the occupation numbers equal to 0 or 1(or else both members are zero).

Comment:

Definition (A-18) must not depend on the specific order of the individual states $u_j, .., u_k, .., u_l, ..$ in the ket on which the operator $a^\dagger_{u_i}$ acts. It can be easily verified that any permutation of the states simply multiply by its parity both members of the equality. It therefore remains valid independently of the order chosen for the individual states in the initial ket.

A-3. Annihilation operators a

We now study the Hermitian conjugate operator of $a^\dagger_{u_i}$, that we shall simply call a_{u_i} since taking twice in a row the Hermitian conjugate of an operator brings you back to the initial operator.

A-3-a. Bosons

For bosons, we deduce from (A-16) that the only non-zero matrix elements of $a^\dagger_{u_i}$ in the Fock states orthonormal basis are:

$$\langle n_1, n_2, .., n_i + 1, ..| \, a^\dagger_{u_i} \, |n_1, n_2, .., n_i, ..\rangle = \sqrt{n_i + 1} \tag{A-20}$$

They link two vectors having equal occupation numbers except for n_i, which increases by one going from the ket to the bra.

The matrix elements of the Hermitian conjugate of $a^\dagger_{u_i}$ are obtained from relation (A-20), using the general definition (B-49) of Chapter II. The only non-zero matrix elements of a_{u_i} are thus:

$$\langle n_1, n_2, .., n_i, ..| \, a_{u_i} \, |n_1, n_2, .., n_i + 1, ..\rangle = \sqrt{n_i + 1} \tag{A-21}$$

Since the basis we use is complete, we can deduce the action of the operators a_i on kets having given occupation numbers:

$$a_{u_i} \ |n_1, n_2, .., n_i, ..\rangle = \sqrt{n_i} \, |n_1, n_2, .., n_i - 1, ..\rangle \tag{A-22}$$

(note that we have replaced n_i by $n_i - 1$). As opposed to $a^\dagger_{u_i}$, which adds a particle in the state $|u_i\rangle$, the operator a_{u_i} takes one away; it yields zero when applied on a ket where the state $|u_i\rangle$ is empty to begin with, such as the vacuum state:

$$a_{u_i} \, |0\rangle = 0 \tag{A-23}$$

We call a_{u_i} "*the annihilation operator*" for the state $|u_i\rangle$.

A-3-b. Fermions

For fermions, relation (A-18) allows writing the matrix elements:

$$\langle u_i, u_j, .., u_k, ..u_l, ..|\, a^\dagger_{u_i}\, |u_j, .., u_k, .., u_l, ..\rangle = 1 \tag{A-24}$$

The only non-zero elements are those where all the individual occupied states are left unchanged in the bra and the ket, except for the state u_i only present in the bra, but not in the ket. As for the occupation numbers, none change, except for n_i which goes from 0 (in the ket) to 1 (in the bra).

The Hermitian conjugation operation then yields the action of the corresponding annihilation operator:

$$a_{u_i}\ \ |u_i, u_j, u_k, .., u_l, ..\rangle = |u_j, u_k, .., u_l, ..\rangle \tag{A-25}$$

or, if initially the state $|u_i\rangle$ is not occupied:

$$a_{u_i}\ \ |u_j, u_k, .., u_l, ..\rangle = 0 \tag{A-26}$$

Relations (A-22) and (A-23) are also valid for fermions, with the usual condition that all occupation numbers should be equal to 0 or 1; otherwise, the relations amount to $0 = 0$.

Comment:

To use relation (A-25) when the state $|u_i\rangle$ is already occupied but not listed in the first position, we first have to bring it there; if it requires an odd permutation, a change of sign will occur. For example:

$$a_{u_2}\, |u_1, u_2\rangle = -\, |u_1\rangle \tag{A-27}$$

For fermions, the operators a and $a^\dagger$ therefore act on the individual state that is listed in the first position in the N-particle ket; a destroys the first state in the list, and $a^\dagger$ creates a new state placed at the beginning of the list. Forgetting this could lead to errors in sign.

A-4. Occupation number operators (bosons and fermions)

Consider the operator $\widehat{n}_{u_i}$ defined by:

$$\widehat{n}_{u_i} = a^\dagger_{u_i} a_{u_i} \tag{A-28}$$

and its action on a Fock state. For bosons, if we apply successively formulas (A-22) and (A-16), we see that this operator yields the same Fock state, but multiplied by its occupation number n_i. For fermions, if $|u_i\rangle$ is empty in the Fock state, relation (A-26) shows that the action of the operator $\widehat{n}_{u_i}$ yields zero. If the state $|u_i\rangle$ is already occupied, we must first permute the states to bring $|u_i\rangle$ to the first position, which may eventually change the sign in front of the Fock space ket. The successive application on this ket of (A-25) and (A-19) shows that the action of the operator $\widehat{n}_{u_i}$ leaves this ket unchanged; we then move the state $|u_i\rangle$ back to its initial position, which may introduce a second change in sign, canceling the first one. We finally obtain for fermions the same result as for bosons, except that the n_i can only take the values 1 and 0. In both cases the Fock states are the eigenvectors of the operator $\widehat{n}_{u_i}$ with the occupation numbers as

eigenvalues; consequently, this operator is named the "*occupation number operator of the state* $|u_i\rangle$". The operator $\widehat{N}$ associated with the total number of particles is simply the sum:

$$\widehat{N} = \sum_i \widehat{n}_{u_i} = \sum_i a^\dagger_{u_i} a_{u_i} \tag{A-29}$$

A-5. Commutation and anticommutation relations

Creation and annihilation operators have very simple commutation (for the bosons) and anticommutation (for the fermions) properties, which make them easy tools for taking into account the symmetrization or antisymmetrization of the state vectors.

To simplify the notation, each time the equations refer to a single basis of individual states $|u_i\rangle$, we shall write a_i instead of a_{u_i}. If, however, it can lead to ambiguity, we will return to the full notation.

A-5-a. Bosons: commutation relations

Consider, for bosons, the two operators $a^\dagger_i$ and $a^\dagger_j$. If both subscripts i and j are different, they correspond to orthogonal states $|u_i\rangle$ and $|u_j\rangle$. Using twice (A-16) then yields:

$$a^\dagger_i a^\dagger_j \,|n_1, n_2, .., n_i, .., n_j, ..\rangle = \sqrt{n_i + 1}\sqrt{n_j + 1}\;\; |n_1, n_2, .., n_i + 1, .., n_j + 1, ..\rangle \tag{A-30}$$

Changing the order of the operators yields the same result. As the Fock states form a basis, we can deduce that the commutator of $a^\dagger_i$ and $a^\dagger_j$ is zero if $i \neq j$. In the same way, it is easy to show that both operator products $a_i a_j$ and $a_j a_i$ acting on the same ket yield the same result (a ket having two occupation numbers lowered by 1); a_i and a_j thus commute if $i \neq j$. Finally the same procedure allows showing that a_i and $a^\dagger_j$ commute if $i \neq j$. Now, if $i = j$, we must evaluate the commutator of a_i and $a^\dagger_i$. Let us apply (A-16) and (A-22) successively, first in that order, and then in the reverse order:

$$\begin{aligned} a_i a^\dagger_i \,|n_1, n_2, .., n_i, ..\rangle &= (n_i + 1)\,|n_1, n_2, .., n_i, ..\rangle \\ a^\dagger_i a_i \,|n_1, n_2, .., n_i, ..\rangle &= (n_i)\,|n_1, n_2, .., n_i, ..\rangle \end{aligned} \tag{A-31}$$

The commutator of a_i and $a^\dagger_i$ is therefore equal to 1 for all the values of the subscript i. All the previous results are summarized in three equalities valid for bosons:

$$[a_i, a_j] = 0 \qquad \left[a^\dagger_i, a^\dagger_j\right] = 0 \qquad \left[a_i, a^\dagger_j\right] = \delta_{ij} \tag{A-32}$$

A-5-b. Fermions: anticommutation relations

For fermions, let us first assume that the subscripts i and j are different. The successive action of $a^\dagger_i$ and $a^\dagger_j$ on an occupation number ket only yields a non-zero ket if $n_i = n_j = 0$; using twice (A-18) leads to:

$$a^\dagger_i a^\dagger_j \;\; |u_k, u_l .., u_m, ..\rangle = |u_i, u_j, u_k, u_l, .., u_m, ..\rangle \tag{A-33}$$

but, if we change the order:

$$a^\dagger_j a^\dagger_i \;\; |u_k, u_l, .., u_m, ..\rangle = |u_j, u_i, u_k, u_l, .., u_m, ..\rangle = -\,|u_i, u_j, u_k, u_l, .., u_m, ..\rangle \tag{A-34}$$

Consequently the sign change that goes with the permutation of the two individual states leads to:

$$a_i^\dagger a_j^\dagger = -a_j^\dagger a_i^\dagger \qquad \text{if } i \neq j \tag{A-35}$$

If we define the anticommutator $[A, B]_+$ of two operators A and B by:

$$[A, B]_+ = AB + BA \tag{A-36}$$

(A-35) may be written as:

$$\left[a_i^\dagger, a_j^\dagger\right]_+ = 0 \qquad \text{if } i \neq j \tag{A-37}$$

Taking the Hermitian conjugate of (A-35), we get:

$$a_j a_i = -a_i a_j \qquad \text{if } i \neq j \tag{A-38}$$

which can be written as:

$$[a_i, a_j]_+ = 0 \qquad \text{if } i \neq j \tag{A-39}$$

Finally, we show by the same method that the anticommutator of a_i and $a_j^\dagger$ is zero except when it acts on a ket where $n_i = 1$ and $n_j = 0$; those two occupation numbers are then interchanged. The computation goes as follows:

$$a_j^\dagger a_i \;\; |u_i, u_l, ..u_m, ..\rangle = a_j^\dagger \, |u_l, ..u_m, ..\rangle = |u_j, u_l, ..u_m, ..\rangle \tag{A-40}$$

and:

$$\begin{aligned} a_i a_j^\dagger \, |u_i, u_l, .., u_m, ..\rangle &= a_i \, |u_j, u_i, u_l, .., u_m, ..\rangle \\ &= -a_i \, |u_i, u_j, u_l, ...u_m, ..\rangle = - \, |u_j, u_l, .., u_m, ..\rangle \end{aligned} \tag{A-41}$$

Adding those two equations yields zero, hence proving that the anticommutator is zero:

$$\left[a_i, a_j^\dagger\right]_+ = 0 \qquad \text{si } i \neq j \tag{A-42}$$

In the case where $i = j$, the limitation on the occupation numbers (0 or 1) leads to:

$$\left[a_i\right]^2 = 0 \qquad \text{and} \qquad \left[a_i^\dagger\right]^2 = 0 \tag{A-43}$$

Equalities (A-37) and (A-39) are still valid if i and j are equal. We are now left with the computation of the anticommutator of a_i and $a_i^\dagger$. Let us first examine the product $a_i a_i^\dagger$; it yields zero if applied to a ket having an occupation number $n_i = 1$, but leaves unchanged any ket with $n_i = 0$, since the particle created by $a_i^\dagger$ is then annihilated by a_i. We get the inverse result for the product $a_i^\dagger a_i$ where the order has been inverted: it yields zero if $n_i = 0$, and leaves the ket unchanged if $n_i = 1$. Finally, whatever the occupation number ket is, one of the terms of the anticommutator yields zero, the other 1, and the net result is always 1. Therefore:

$$\left[a_i, a_i^\dagger\right]_+ = 1 \tag{A-44}$$

All the previous results valid for fermions are summarized in the following three relations, which are for fermions the equivalent of relations (A-32) for bosons:

$$[a_i, a_j]_+ = 0 \qquad \left[a_i^\dagger, a_j^\dagger\right]_+ = 0 \qquad \left[a_i, a_j^\dagger\right]_+ = \delta_{ij} \tag{A-45}$$

A-5-c. Common relations for bosons and fermions

To regroup the results valid for bosons and fermions in common relations, we introduce the notation:

$$[A,B]_{-\eta} = AB - \eta\ BA \tag{A-46}$$

with:

$$\begin{array}{ll} \eta = & 1 \quad \text{for bosons} \\ \eta = & -1 \quad \text{for fermions} \end{array} \tag{A-47}$$

so that (A-46) is the commutator of A and B for bosons, and their anticommutator for fermions. We then have:

$$\begin{array}{l} \left[a_i, a_j\right]_{-\eta} = 0 \quad \text{for all } i \text{ and } j \text{ (equal or different)} \\ \left[a_i^\dagger, a_j^\dagger\right]_{-\eta} = 0 \quad \text{for all } i \text{ and } j \text{ (equal or different)} \end{array} \tag{A-48}$$

and the only non-zero combinations are:

$$\left[a_i, a_j^\dagger\right]_{-\eta} = \delta_{ij} \qquad\qquad \left[a_i^\dagger, a_j\right]_{-\eta} = -\eta\delta_{ij} \tag{A-49}$$

A-6. Change of basis

What are the effects on the creation and annihilation operators of a change of basis for the individual states? The operators $a_{u_i}^\dagger$ and a_{u_i} have been introduced by their action on the Fock states, defined by relations (A-7) and (A-10) for which a given basis of individual states $\{|u_i\rangle\}$ was chosen. One could also choose any another orthonormal basis $\{|v_s\rangle\}$ and define in the same way bases for the Fock state and creation $a_{v_s}^\dagger$ and annihilation a_{v_s} operators. What is the relation between these new operators and the ones we defined earlier with the initial basis?

For creation operators acting on the vacuum state $|0\rangle$, the answer is quite straightforward: the action of $a_{v_s}^\dagger$ on $|0\rangle$ yields a one-particle ket, which can be written as:

$$a_{v_s}^\dagger |0\rangle = |1 : v_s\rangle = \sum_i \langle u_i|\ v_s\rangle\ |1 : u_i\rangle = \sum_i \langle u_i|\ v_s\rangle\ a_{u_i}^\dagger |0\rangle \tag{A-50}$$

This result leads us to expect a simple linear relation of the type:

$$a_{v_s}^\dagger = \sum_i \langle u_i|\ v_s\rangle\ \ a_{u_i}^\dagger \tag{A-51}$$

with its Hermitian conjugate:

$$a_{v_s} = \sum_i \langle v_s|\ u_i\rangle\ \ a_{u_i} \tag{A-52}$$

Equation (A-51) implies that creation operators are transformed by the same unitary relation as the individual states. Commutation or anticommutation relations are then conserved, since:

$$\left[a_{v_s}, a_{v_t}^\dagger\right]_{-\eta} = \sum_{i,j} \langle v_s|\ u_i\rangle\ \langle u_j|\ v_t\rangle\ \left[a_{u_i}, a_{u_j}^\dagger\right]_{-\eta} = \sum_{i,j} \langle v_s|\ u_i\rangle\ \langle u_j|\ v_t\rangle\ \delta_{ij} \tag{A-53}$$

which amounts to (as expected):

$$\left[a_{v_s}, a^\dagger_{v_t}\right]_{-\eta} = \sum_i \langle v_s|\ u_i\rangle \langle u_i|\ v_t\rangle = \delta_{st} \tag{A-54}$$

Furthermore, it is straightforward to show that the creation operators commute (or anticommute), as do the annihilation operators.

Equivalence of the two bases

We have not yet shown the complete equivalence of the two bases, which can be done following two different approaches. In the first one, we use (A-51) and (A-52) to define the creation and annihilation operators in the new basis. The associated Fock states are defined by replacing the $a^\dagger_{u_i}$ by the $a^\dagger_{v_s}$ in relations (A-17) for the bosons, and (A-18) for the fermions. We then have to show that these new Fock states are still related to the states with numbered particles as in (A-18) for bosons, and (A-10) for fermions. This will establish the complete equivalence of the two bases.

We shall follow a second approach where the two bases are treated completely symmetrically. Replacing in relations (A-7) and (A-10) the u_i by the v_s, we construct the new Fock basis. We next define the operators $a^\dagger_{v_s}$ by transposing relations (A-17) and (A-18) to the new basis. We then must verify that these operators obey relation (A-51), without limiting ourselves, as in (A-50), to their action on the vacuum state.

(i) Bosons

Relations (A-7) and (A-17) lead to:

$$\begin{aligned}&\left[a^\dagger_{u_i}\right]^{n_i}\left[a^\dagger_{u_j}\right]^{n_j} .. |0\rangle \\ &= \sqrt{N!}\ S_N\ \ |1:u_i;\ 2:u_i;\ ..;\ n_i:u_i;\ \ n_i+1:u_j;..;\ n_i+n_j:u_j;\ ..\rangle\end{aligned} \tag{A-55}$$

where, on the right-hand side, the n_i first particles occupy the same individual state u_i, the following n_j particles, numbered from n_i+1 to n_i+n_j, the individual state u_j, etc. The equivalent relation in the second basis can be written:

$$\begin{aligned}&\left[a^\dagger_{v_s}\right]^{p_s}\left[a^\dagger_{v_t}\right]^{p_t} .. |0\rangle \\ &= \sqrt{N!}\ S_N\ \ |1:v_s;\ 2:v_s;\ ..;\ p_s:v_s;\ \ p_s+1:v_t;..;\ p_s+p_t:v_t;\ ..\rangle\end{aligned} \tag{A-56}$$

with:

$$n_i + n_j + .. = p_s + p_t + .. = N \tag{A-57}$$

Replacing on the right-hand side of (A-56), the first ket $|v_s\rangle$ by:

$$|v_s\rangle = \sum_i \langle u_i|\ v_s\rangle\, |u_i\rangle \tag{A-58}$$

we obtain:

$$\sum_i \langle u_i|\ v_s\rangle\, \sqrt{N!}\ S_N\, |1:u_i;\ 2:v_s;\ ..;\ p_s:v_s;\ \ p_s+1:v_t;..;\ p_s+p_t:v_t;\ ..\rangle \tag{A-59}$$

Following the same procedure for all the basis vectors of the right-hand side, we can replace it by:

$$\sum_i \langle u_i|\ v_s\rangle \sum_{i'} \langle u_{i'}|\ v_s\rangle \,..\sum_j \langle u_j|\ v_t\rangle \sum_{j'} \langle u_{j'}|\ v_t\rangle \,..$$
$$\times\sqrt{N!}\ S_N\,|1:u_i;\ 2:u_{i'};\ ..;\ p_s+1:u_j;\ p_s+2:u_{j'};\ ..\rangle \tag{A-60}$$

or else[6], taking into account (A-55):

$$\left[\sum_i \langle u_i|\ v_s\rangle\, a^\dagger_{u_i}\right]\left[\sum_{i'} \langle u_{i'}|\ v_s\rangle\, a^\dagger_{u_{i'}}\right] ..\left[\sum_j \langle u_j|\ v_t\rangle\, a^\dagger_{u_j}\right]\left[\sum_{j'} \langle u_{j'}|\ v_t\rangle\, a^\dagger_{u_{j'}}\right] ..|0\rangle \tag{A-61}$$

We have thus shown that the operators $\left[a^\dagger_{v_s}\right]^{p_s}\left[a^\dagger_{v_t}\right]^{p_t}$ .. act on the vacuum state in the same way as the operators defined by (A-51), raised to the powers p_s, p_t , ..

When the occupation numbers p_s, p_t, .. can take on any values, the kets (A-56) span the entire Fock space. Writing the previous equality for $p_s$and p_s+1, we see that the action on all the basis kets of $a^\dagger_{v_s}$ and of $\sum_i \langle u_i|\ v_s\rangle\, a^\dagger_{u_i}$ yields the same result, establishing the equality between these two operators. Relation (A-52) can be readily obtained by Hermitian conjugation.

(ii) Fermions

The demonstration is identical, with the constraint that the occupation numbers are 0 or 1. As this requires no changes in the operator or state order, it involves no sign changes.

B. One-particle symmetric operators

Using creation and annihilation operators makes it much easier to deal, in the Fock space, with physical operators that are thus symmetric (§ C-4-a-β of Chapter XIV). We first study the simplest of such operators, those which act on a single particle and are called "one-particle operators".

B-1. Definition

Consider an operator $\widehat{f}$ defined in the space of individual states; $\widehat{f}(q)$ acts in the state space of particle q. It could be for example the momentum of the q-th particle, or its angular momentum with respect to the origin. We now build the operator associated with the total momentum of the N-particle system, or its total angular momentum, which is the sum over q of all the $\widehat{f}(q)$ associated with the individual particles.

A one-particle symmetric operator acting in the space $\mathcal{E}_S(N)$ for bosons - or $\mathcal{E}_A(N)$ for fermions - is therefore defined by:

$$\widehat{F}^{(N)} = \sum_{q=1}^{N} \widehat{f}(q) \tag{B-1}$$

[6]In this relation, the first p_s sums are identical, as are the next p_t sums, etc.

(contrary to states, which are symmetric for bosons and antisymmetric for fermions, the physical operators are always symmetric). The operator $\widehat{F}$ acting in the Fock space is defined as the operator $\widehat{F}^{(N)}$ acting either in $\mathcal{E}_S(N)$ or in $\mathcal{E}_A(N)$, depending on the specific case. Since the basis for the entire Fock space is the union of the bases of these spaces for all values of N, the operator $\widehat{F}$ is thus well defined in the direct sum of all these subspaces. To summarize:

$$\widehat{F}^{(N)} \quad ; \quad N = 1, 2, 3, ... \quad \Longrightarrow \widehat{F} \tag{B-2}$$

Using (B-1) directly to compute the matrix elements of $\widehat{F}$ often leads to tedious manipulations. Starting with an operator involving numbered particles, we place it between states with numbered particles; we then symmetrize the bra, the ket, and take into account the symmetry of the operator (*cf.* footnote 1). This introduces several summations (on the particles and on the permutations) that have to be properly regrouped to be simplified. We will now show that expressing $\widehat{F}$ in terms of creation and annihilation operators avoids all these intermediate calculations, taking nevertheless into account all the symmetry properties.

B-2. Expression in terms of the operators a and $a^\dagger$

We choose a basis $\{|u_i\rangle\}$ for the individual states. The matrix elements f_{kl} of the one-particle operator $\widehat{f}$ are given by:

$$f_{kl} = \langle u_k | \, \widehat{f} \, | u_l \rangle \tag{B-3}$$

They can be used to expand the operator itself as follows:

$$\widehat{f}(q) = \sum_{k,l} |q : u_k\rangle \langle q : u_k| \, \widehat{f}(q) \, |q : u_l\rangle \langle q : u_l| = \sum_{k,l} f_{kl} \; |q : u_k\rangle \langle q : u_l| \tag{B-4}$$

B-2-a. Action of $F^{(N)}$ on a ket with N particles

Using in (B-1) the expression (B-4) for $\widehat{f}(q)$ leads to:

$$\widehat{F}^{(N)} = \sum_{k,l} f_{kl} \sum_{q=1}^{N} \; |q : u_k\rangle \langle q : u_l| \tag{B-5}$$

The action of $\widehat{F}^{(N)}$ on a symmetrized ket written as (A-9) therefore includes a sum over k and l of terms:

$$\left[\sum_{q=1}^{N} \; |q : u_k\rangle \langle q : u_l| \right] \quad |u_i, u_i, .., u_j, u_j, ..\rangle \tag{B-6}$$

with coefficients f_{kl}. Let us use (A-7) or (A-10) to compute this ket for given values of k and l. As the operator contained in the bracket is symmetric with respect to the exchange of particles, it commutes with the two operators S_N and A_N (§ C-4-a-β of Chapter XIV)), and the ket can be written as:

$$\sqrt{\frac{N!}{n_i! n_j! .. n_m! ..}} \, S_N \, A_N \left[\sum_{q=1}^{N} \; |q : u_k\rangle \langle q : u_l| \right] \\ |1 : u_i; \; 2 : u_i; ..; \; n_i : u_i; \; n_i + 1 : u_j; \; ..; \; q : u_m; \; ..\rangle \tag{B-7}$$

In the summation over q, the only non-zero terms are those for which the individual state $|u_l\rangle$ coincides with the individual state $|u_m\rangle$ occupied in the ket on the right by the particle labeled q; there are n_l different values of q that obey this condition (i.e. none or one for fermions). For these n_l terms, the operator $|q:u_k\rangle\langle q:u_l|$ transforms the state $|u_m\rangle$ into $|u_k\rangle$, then S_N (or A_N) reconstructs a symmetrized (but not normalized) ket:

$$\sqrt{\frac{N!}{n_i!n_j!..n_m!..}}\begin{matrix}S_N\\A_N\end{matrix}\ |1:u_i;\ ..;\ n_i+1:u_j;\ ..;\ q:u_k;\ ..\rangle \tag{B-8}$$

This ket is always the same for all the numbers q among the n_l selected ones (for fermions, this term might be zero, if the state $|u_k\rangle$ was already occupied in the initial ket). We shall then distinguish two cases:

(i) For $k \neq l$, and for bosons, the ket written in (B-8) equals:

$$\sqrt{\frac{n_k+1}{n_l}}\,|u_i,u_i,..,u_j,u_j,..,u_k,..\rangle \tag{B-9}$$

where the square root factor comes from the variation in the occupation numbers n_k and n_l, which thus change the numerical coefficients in the definition (A-7) of the Fock states. As this ket is obtained n_l times, this factor becomes $\sqrt{(n_k+1)n_l}$. This is exactly the factor obtained by the action on the same symmetrized ket of the operator $a_k^\dagger a_l$, which also removes a particle from the state $|u_l\rangle$ and creates a new one in the state $|u_k\rangle$. Consequently, the operator $a_k^\dagger a_l$ reproduces exactly the same effect as the sum over q.

For fermions, the result is zero except when, in the initial ket, the state $|u_l\rangle$ was occupied by a particle, and the state $|u_k\rangle$ empty, in which case no numerical factor appears; as before, this is exactly what the action of the operator $a_k^\dagger a_l$ would do.

(ii) if $k = l$, for bosons the only numerical factor involved is n_l, coming from the number of terms in the sum over q that yields the same symmetrized ket. For fermions, the only condition that yields a non-zero result is for the state $|u_l\rangle$ to be occupied, which also leads to the factor n_l. In both cases, the sum over q amounts to the action of the operator $a_k^\dagger a_l$.

We have shown that:

$$\sum_{q=1}^{N}\ |q:u_k\rangle\langle q:u_l| = a_k^\dagger a_l \tag{B-10}$$

The summation over k and l in (B-5) then yields:

$$\widehat{F}^{(N)} = \sum_{k,l}\ f_{kl}\ \ a_k^\dagger a_l = \sum_{k,l}\langle u_k|\,\widehat{f}\,|u_l\rangle\ \ a_k^\dagger a_l \tag{B-11}$$

B-2-b. Expression valid in the entire Fock space

The right-hand side of (B-11) contains an expression completely independent of the space $\mathcal{E}_S(N)$ or $\mathcal{E}_A(N)$ in which we defined the action of the operator $\widehat{F}^{(N)}$. Since we defined operator $\widehat{F}$ as acting as $\widehat{F}^{(N)}$ in each of these subspaces having fixed N, we can simply write:

$$\boxed{\widehat{F} = \sum_{k,l}\langle u_k|\,\widehat{f}\,|u_l\rangle\ \ a_k^\dagger a_l} \tag{B-12}$$

This is the expression of one-particle symmetric operators we were looking for. Its form is valid for any value of N and the particles are no longer numbered; it contains equal numbers of creation and annihilation operators, which only act on occupation numbers.

Comment:

Choosing the proper basis $\{|u_i\rangle\}$, it is always possible to diagonalize the Hermitian operator $\widehat{f}$ and write:

$$f_{kl} = f_k \times \delta_{kl} \tag{B-13}$$

Equality (B-11) is then simply written as:

$$\widehat{F} = \sum_k f_k \ a_k^\dagger a_k = \sum_k f_k \ \widehat{n}_k \tag{B-14}$$

where $\widehat{n}_k = \widehat{n}_{u_k}$ is the occupation number operator in the state $|u_k\rangle$ defined in (A-28).

B-3. Examples

A first very simple example is the operator $\widehat{N}$, already described in (A-29), and corresponding to the total number of particles:

$$\widehat{N} = \sum_i \widehat{n}_i = \sum_i a_i^\dagger a_i \tag{B-15}$$

As expected, this operator does not depend on the basis $\{|u_i\rangle\}$ chosen to count the particles, as we now show. Using the unitary transformations of operators (A-51) and (A-52), and with the full notation for the creation and annihilation operators to avoid any ambiguity, we get:

$$\sum_i a_{u_i}^\dagger a_{u_i} = \sum_i \sum_{s,t} \langle v_s|\ u_i\rangle \ \langle u_i|\ v_t\rangle \ a_{v_s}^\dagger a_{v_t} = \sum_{s,t} \delta_{st} \ a_{v_s}^\dagger a_{v_t} \tag{B-16}$$

which shows that:

$$\widehat{N} = \sum_i a_{u_i}^\dagger a_{u_i} = \sum_s a_{v_s}^\dagger a_{v_s} \tag{B-17}$$

For a spinless particle one can also define the operator corresponding to the probability density at point $\mathbf{r}_0$:

$$\widehat{f} = |\mathbf{r}_0\rangle \ \langle \mathbf{r}_0| \tag{B-18}$$

Relation (B-12) then leads to the "particle local density" (or "single density") operator:

$$\widehat{D}(\mathbf{r}_0) = \sum_{k,l} u_k^*(\mathbf{r}_0) u_l(\mathbf{r}_0) \ a_k^\dagger a_l \tag{B-19}$$

The same procedure as above shows that this operator is independent of the basis $\{|u_i\rangle\}$ chosen in the individual states space.

Let us assume now that the chosen basis is formed by the eigenvectors $|\mathbf{k}_i\rangle$ of a particle's momentum $\hbar\mathbf{k}_i$, and that the corresponding annihilation operators are noted $a_{\mathbf{k}_i}$. The operator associated with the total momentum of the system can be written as:

$$\widehat{\mathbf{P}} = \sum_{\mathbf{k}_i} \hbar\mathbf{k}_i \; a^\dagger_{\mathbf{k}_i} a_{\mathbf{k}_i} = \sum_{\mathbf{k}_i} \hbar\mathbf{k}_i \; \widehat{n}_{\mathbf{k}_i} \tag{B-20}$$

As for the kinetic energy of the particles, its associated operator is expressed as:

$$\widehat{H}_0 = \sum_{\mathbf{k}_i} \frac{\hbar^2\mathbf{k}_i^2}{2m} \; a^\dagger_{\mathbf{k}_i} a_{\mathbf{k}_i} = \sum_{\mathbf{k}_i} \frac{\hbar^2\mathbf{k}_i^2}{2m} \; \widehat{n}_{\mathbf{k}_i} \tag{B-21}$$

B-4. Single particle density operator

Consider the average value $\left\langle \widehat{F} \right\rangle$ of a one-particle operator $\widehat{F}$ in an arbitrary N-particle quantum state. It can be expressed, using relation (B-12), as a function of the average values of operator products $a^\dagger_k a_l$:

$$\left\langle \widehat{F} \right\rangle = \sum_{k,l} \langle u_k | \, \widehat{f} \, | u_l \rangle \;\; \left\langle a^\dagger_k a_l \right\rangle \tag{B-22}$$

This expression is close to that of the average value of an operator for a physical system composed of a single particle. Remember (Complement $\mathrm{E_{III}}$, § 4-b) that if a system is described by a single particle density operator $\widehat{\rho}_1\,(1)$, the average value of any operator $\widehat{f}\,(1)$ is written as:

$$\left\langle \widehat{f}\,(1) \right\rangle = \mathrm{Tr}\left\{ \widehat{f}\,(1)\,\widehat{\rho}_1\,(1) \right\} = \sum_{k,l} \langle u_k | \, \widehat{f} \, | u_l \rangle \;\; \langle u_l | \, \widehat{\rho}_1 \, | u_k \rangle \tag{B-23}$$

The above two expressions can be made to coincide if, for the system of identical particles, we introduce a "density operator reduced to a single particle" $\widehat{\rho}_1$ whose matrix elements are defined by:

$$\langle u_l | \, \widehat{\rho}_1 \, | u_k \rangle = \left\langle a^\dagger_k a_l \right\rangle \tag{B-24}$$

This reduced operator allows computing average values of all the single particle operators as if the system consisted only of a single particle:

$$\left\langle \widehat{F} \right\rangle = \mathrm{Tr}\left\{ \widehat{f} \; \widehat{\rho}_1 \right\} \tag{B-25}$$

where the trace is taken in the state space of a single particle.

The trace of the reduced density operator thus defined is not equal to unity, but to the average particle number as can be shown using (B-24) and (B-15):

$$\mathrm{Tr}\,\{\widehat{\rho}_1\} = \sum_k \left\langle a^\dagger_k a_k \right\rangle = \left\langle \widehat{N} \right\rangle \tag{B-26}$$

This normalization convention can be useful. For example, the diagonal matrix element of $\widehat{\rho}_1$ in the position representation is simply the average of the particle local density defined in (B-19):

$$\langle \mathbf{r}_0 | \, \widehat{\rho}_1 \, | \mathbf{r}_0 \rangle = \left\langle \widehat{D}(\mathbf{r}_0) \right\rangle \tag{B-27}$$

It is however easy to choose a different normalization for the reduced density operator: its trace can be made equal to 1 by dividing the right-hand side of definition (B-24) by the factor $\left\langle \widehat{N} \right\rangle$.

C. Two-particle operators

We now extend the previous results to the case of two-particle operators.

C-1. Definition

Consider a physical quantity involving two particles, labeled q and q'. It is associated with an operator $\widehat{g}(q, q')$ acting in the state space of these two particles (the tensor product of the two individual state's spaces). Starting from this binary operator, the easiest way to obtain a symmetric N-particle operator is to sum all the $\widehat{g}(q, q')$ over all the particles q and q', where the two subscripts q and q' range from 1 to N. Note, however, that in this sum all the terms where $q = q'$ add up to form a one-particle operator of exactly the same type as those studied in § B-1. Consequently, to obtain a real two-particle operator we shall exclude the terms where $q = q'$ and define:

$$\widehat{G}^{(N)} = \frac{1}{2} \sum_{q,q'=1;\ q \neq q'}^{N} \widehat{g}(q, q') \tag{C-1}$$

The factor $1/2$ present in this expression is arbitrary but often handy. If for example the operator describes an interaction energy that is the sum of the contributions of all the distinct pairs of particles, $\widehat{g}(q, q')$ and $\widehat{g}(q', q)$ corresponding to the same pair are equal and appear twice in the sum over q and q': the factor $1/2$ avoids counting them twice. Whenever $\widehat{g}(q, q') = \widehat{g}(q', q)$, it is equivalent to write $\widehat{G}^{(N)}$ in the form:

$$\widehat{G}^{(N)} = \sum_{q<q'}^{N} \widehat{g}(q, q') \tag{C-2}$$

As with the one-particle operators, expression (C-1) defines symmetric operators separately in each physical state's space having a given particle number N. This definition may be extended to the entire Fock space, which is their direct sum over all N. This results in a more general operator $\widehat{G}$, following the same scheme as for (B-2):

$$\widehat{G}^{(N)} \quad ; \quad N = 1, 2, 3, ... \quad \Longrightarrow \widehat{G} \tag{C-3}$$

C-2. A simple case: factorization

Let us first assume the operator $\widehat{g}(q,q')$ can be factored as:

$$\widehat{g}(q,q') = \widehat{f}(q)\widehat{h}(q') \tag{C-4}$$

The operator written in (C-1) then becomes:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{q,q'=1;\ q\neq q'}^{N} \widehat{f}(q)\widehat{h}(q') = \frac{1}{2}\left[\sum_{q=1}^{N}\widehat{f}(q)\sum_{q'=1}^{N}\widehat{h}(q') - \sum_{q=1}^{N}\widehat{f}(q)\widehat{h}(q)\right] \tag{C-5}$$

The right-hand side of this expression starts with a product of one-particle operators, each of which can be replaced, following (B-11), by its expression as a function of the creation and annihilation operators:

$$\sum_{q=1}^{N}\widehat{f}(q) = \sum_{i,k}\ \langle u_i|\,\widehat{f}\,|u_k\rangle\ a_i^\dagger a_k \qquad \text{and} \qquad \sum_{q'=1}^{N}\widehat{h}(q') = \sum_{j,l}\ \langle u_j|\,\widehat{h}\,|u_l\rangle\ a_j^\dagger a_l \tag{C-6}$$

As for the last term on the right-hand side of (C-5), it is already a single particle operator:

$$\sum_{q=1}^{N}\widehat{f}(q)\widehat{h}(q) = \sum_{i,l}\ \langle u_i|\,\widehat{f}\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l \tag{C-7}$$

This leads to:

$$\widehat{G}^{(N)} = \frac{1}{2}\left[\sum_{i,j,k,l}\langle u_i|\,\widehat{f}\,|u_k\rangle\,\langle u_j|\,\widehat{h}\,|u_l\rangle\left(a_i^\dagger a_k\right)\left(a_j^\dagger a_l\right) - \sum_{i,l}\langle u_i|\,\widehat{f}\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l\right] \tag{C-8}$$

We can then use general relations (A-49) to transform the operator product:

$$\left(a_i^\dagger a_k\right)\left(a_j^\dagger a_l\right) = \eta\ a_i^\dagger a_j^\dagger a_k a_l + \delta_{jk}\ a_i^\dagger a_l = a_i^\dagger a_j^\dagger a_l a_k + \delta_{jk}\ a_i^\dagger a_l \tag{C-9}$$

Including this form in the first term on the right-hand side of (C-8) yields, for the δ_{jk} contribution:

$$\sum_{i,l}\sum_{k}\langle u_i|\,\widehat{f}\,|u_k\rangle\,\langle u_k|\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l = \sum_{i,l}\ \langle u_i|\,\widehat{f}\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l \tag{C-10}$$

which exactly cancels the second term of (C-8). Consequently, we are left with:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{i,j,k,l}\langle u_i|\,\widehat{f}\,|u_k\rangle\,\langle u_j|\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_j^\dagger a_l a_k \tag{C-11}$$

As the right-hand side of this expression has the same form in all spaces having a fixed N, it is also valid for the operator $\widehat{G}$ acting in the entire Fock space.

C-3. General case

Any two-particle operator $\widehat{g}(q,q')$ may be decomposed as a sum of products of single particle operators:

$$\widehat{g}(q,q') = \sum_{\alpha,\beta} c_{\alpha,\beta}\ \widehat{f}_\alpha(q)\widehat{h}_\beta(q') \tag{C-12}$$

where the coefficients $c_{\alpha,\beta}$ are numbers[7]. Hence expression (C-1) can be written as:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{q,q'=1;\ q\neq q'}^{N} \widehat{g}(q,q') = \frac{1}{2}\sum_{\alpha,\beta} c_{\alpha,\beta}\sum_{q=1}^{N}\sum_{q'=1;\ q'\neq q}^{N} \widehat{f}_\alpha(q)\widehat{h}_\beta(q') \tag{C-13}$$

In this linear combination with coefficients $c_{\alpha,\beta}$, each term (corresponding to a given α and β) is of the form (C-5) and can therefore be replaced by expression (C-11). This leads to:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{\alpha,\beta} c_{\alpha,\beta}\sum_{i,j,k,l} \langle u_i|\,\widehat{f_\alpha}\,|u_k\rangle\,\langle u_j|\,\widehat{h_\beta}\,|u_l\rangle\ a_i^\dagger a_j^\dagger a_l a_k \tag{C-14}$$

The right-hand side of this equation has the same form in all the spaces of fixed N; hence it is valid in the entire Fock space. Furthermore, we recognize in the summation over α and β the matrix element of $\widehat{g}$ as defined by (C-12):

$$g_{i,j,k,l} = \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle = \sum_{\alpha,\beta} c_{\alpha,\beta}\ \langle u_i|\,\widehat{f_\alpha}\,|u_k\rangle\,\langle u_j|\,\widehat{h_\beta}\,|u_l\rangle \tag{C-15}$$

The final result is then:

$$\boxed{\widehat{G} = \frac{1}{2}\sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle\ a_i^\dagger a_j^\dagger a_l a_k} \tag{C-16}$$

which is the general expression for a two-particle symmetric operator.

As for the one-particle operators, each term of expression (C-16) for the two-particle operators contains equal numbers of creation and annihilation operators. Consequently, these symmetric operators do not change the total number of particles, as was obvious from their initial definition.

C-4. Two-particle reduced density operator

Relation (C-16) implies that the average value of any two-particle operator may be written as:

$$\left\langle \widehat{G} \right\rangle = \frac{1}{2}\sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle\ \left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle \tag{C-17}$$

[7] The two-particle state space is the tensor product of the two spaces of individual states (see § F-4-b of Chapter II). In the same way, the space of operators acting on two particles is the tensor product of the spaces of operators acting separately on these particles. For example, the operator for the interaction potential between two particles can be decomposed as a sum of products of two operators: the first one is a function of the position of the first particle, and the second one of the position of the second particle.

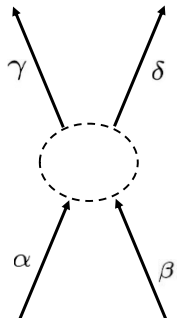


Figure 1: Physical interaction between two identical particles: initially in the states $|u_{k_\alpha}\rangle$ and $|u_{k_\beta}\rangle$ (schematized by the letters α and β), the particles are transferred to the states $|u_{k_\gamma}\rangle$ and $|u_{k_\delta}\rangle$ (schematized by the letters γ and δ)

This expression is similar to the average value of an operator $\widehat{g}\,(1,2)$ for a two-particle system having a density operator $\widehat{\rho}_2\,(1,2)$:

$$\langle \widehat{g}\,(1,2)\rangle = \sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle \times \langle 1:u_k;\ 2:u_l|\,\widehat{\rho}_2\,(1,2)\,|1:u_i;\ 2:u_j\rangle \tag{C-18}$$

which leads us to define a two-particle reduced density operator $\widehat{\rho}_2$:

$$\langle 1:u_k;\ 2:u_l|\,\widehat{\rho}_2\,|1:u_i;\ 2:u_j\rangle = \left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle \tag{C-19}$$

In this definition we have left out the factor 1/2 of (C-17) since this will lead to a normalization of $\widehat{\rho}_2$ often more handy: the matrix element of $\widehat{\rho}_2$ in the position representation yields directly the double density (as well as the field correlation function that we shall study in § B-3-b of Chapter XVI). The trace of $\widehat{\rho}_2$ is then written:

$$\begin{aligned}\mathrm{Tr}\,\{\widehat{\rho}_2\} &= \sum_{i,j} \left\langle a_i^\dagger a_j^\dagger a_j a_i \right\rangle = \sum_{i,j} \left\langle a_i^\dagger a_i a_j^\dagger a_j - \delta_{ij} a_i^\dagger a_i \right\rangle \\ &= \left\langle \widehat{N}\left(\widehat{N}-1\right)\right\rangle \end{aligned} \tag{C-20}$$

It is obviously possible to divide the right-hand side of the definition of $\widehat{\rho}_2$ either by the factor 2, or else by the factor $\left\langle \widehat{N}\left(\widehat{N}-1\right)\right\rangle$ if we wish its trace to be equal to 1.

C-5. Physical discussion; consequences of the exchange

As mentioned in the introduction of this chapter, the equations no longer contain labeled particles, permutations, symmetrizers and antisymmetrizers; the total number of particles N has also disappeared. We may now continue the discussion begun in § D-2 of Chapter XIV concerning the exchange terms, but in a more general way since we no longer specify the total particle number N.

C-5-a. Two terms in the matrix elements

Consider a physical process (schematized in Figure 1) where, in a system of N identical particles, an interaction produces a transfer from the two states $\left|u_{k_\alpha}\right\rangle$ and $\left|u_{k_\beta}\right\rangle$ towards the two states $\left|u_{k_\gamma}\right\rangle$ and $\left|u_{k_\delta}\right\rangle$; we assume that the four states we are dealing with are different. In the summation over i, j, k, l of (C-16), the only terms involved in this process are those where the bra contains either $i = k_\gamma$ and $j = k_\delta$, or the opposite $i = k_\delta$ and $j = k_\gamma$; as for the ket, it must contain either $k = k_\alpha$ and $l = k_\beta$, or the opposite $k = k_\beta$ and $l = k_\alpha$. We are then left with four terms:

$$\begin{aligned}
&\tfrac{1}{2}\left\langle 1 : u_{k_\gamma};\ 2 : u_{k_\delta}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\ a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}\\
&\tfrac{1}{2}\left\langle 1 : u_{k_\delta};\ 2 : u_{k_\gamma}\right|\widehat{g}\left|1 : u_{k_\beta};\ 2 : u_{k_\alpha}\right\rangle\ a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\alpha} a_{k_\beta}\\
&\tfrac{1}{2}\left\langle 1 : u_{k_\gamma};\ 2 : u_{k_\delta}\right|\widehat{g}\left|1 : u_{k_\beta};\ 2 : u_{k_\alpha}\right\rangle\ a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\alpha} a_{k_\beta}\\
&\tfrac{1}{2}\left\langle 1 : u_{k_\delta};\ 2 : u_{k_\gamma}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\ a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\beta} a_{k_\alpha}
\end{aligned} \tag{C-21}$$

However, since the numbers used to label the particles are dummy variables, the first two matrix elements shown in (C-21) are equal and so are the last two. In addition, the product of creation and annihilation operators obey the following relations, for bosons ($\eta = 1$) as well as for fermions ($\eta = -1$):

$$\begin{aligned}
a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\alpha} a_{k_\beta} &= a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}\\
a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\alpha} a_{k_\beta} &= a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\beta} a_{k_\alpha} = \eta\ a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}
\end{aligned} \tag{C-22}$$

These relations are obvious for bosons since we only commute either creation operators or annihilation operators. For fermions, as we assumed all the states were different, the anticommutation of operators a or of operators $a^\dagger$, leads to sign changes; these may cancel out depending on whether the number of anticommutations is even or odd. If we now double the sum of the first and last term of (C-21), we obtain the final contribution to (C-16):

$$\begin{aligned}
&a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}\\
&\qquad\times\ \Big[\left\langle 1 : u_{k_\gamma};\ 2 : u_{k_\delta}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\\
&\qquad\qquad+\ \eta\left\langle 1 : u_{k_\delta};\ 2 : u_{k_\gamma}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\Big]
\end{aligned} \tag{C-23}$$

Hence we are left with two terms whose relative sign depends on the nature (bosons or fermions) of the identical particles. They correspond to a different "switching point" for the incoming and outgoing individual states (Fig. 2).

For bosons, the product of the 4 operators in (C-23) acting on an occupation number ket introduces the square root:

$$\sqrt{n_{k_\alpha} n_{k_\beta}\left(n_{k_\gamma}+1\right)\left(n_{k_\delta}+1\right)} \tag{C-24}$$

For large occupation numbers, this square root may considerably increase the value of the matrix element. For fermions, however, this amplification effect does not occur. Furthermore, if the direct and exchange matrix elements of $\widehat{g}$ are equal, they will cancel each other in (C-23) and the corresponding transition amplitude of this process will be zero.

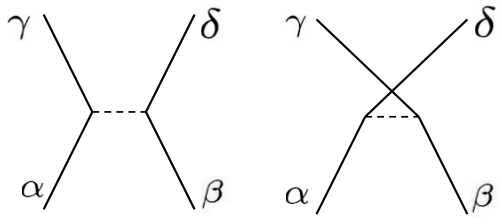


Figure 2: Two diagrams representing schematically the two terms appearing in equation (C-23); they differ by an exchange of the individual states of the exit particles. They correspond, in a manner of speaking, to a different "switching point" for the incoming and outgoing states. The solid lines represent the particles' free propagation, and the dashed lines their binary interaction.

C-5-b. Particle interaction energy, the direct and exchange terms

Many physics problems involve computing the average particle interaction energy. For the sake of simplicity, we shall only study here spinless particles (or, equivalently, particles being in the same internal spin state so that the corresponding quantum number does not come into play) and assume their interactions to be binary. These interactions are then described by an operator $\widehat{W}_{\rm int}$, diagonal in the $\{|\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N\rangle\}$ basis (eigenstates of all the particles' positions), which multiplies each of these states by the function:

$$W_{\rm int}(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = \sum_{q<q'} W_2(\mathbf{r}_q, \mathbf{r}_{q'}) \tag{C-25}$$

In this expression, the function $W_2(\mathbf{r}_q, \mathbf{r}_{q'})$ yields the diagonal matrix elements of the operator $\widehat{W}_2(\mathbf{R}_q, \mathbf{R}_{q'})$ associated with the two-particle (q, q') interaction, where $\mathbf{R}_q$ is the quantum operator associated with the classical position $\mathbf{r}_q$. The matrix elements of this operator in the $|u_k; u_l\rangle$ basis is simply obtained by inserting a closure relation for each of the two positions. This leads to:

$$\begin{aligned}&\langle 1:u_i;\ 2:u_j|\,\widehat{W}_2(\mathbf{R}_1,\mathbf{R}_2)\,|1:u_k;\ 2:u_l\rangle \\ &\qquad = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2\ W_2(\mathbf{r}_1,\mathbf{r}_2)\ u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_k(\mathbf{r}_1) u_l(\mathbf{r}_2)\end{aligned} \tag{C-26}$$

α. General expression:

Replacing in (C-16) operator $\widehat{g}(1,2)$ by $\widehat{W}_{\rm int}(\mathbf{R}_1, \mathbf{R}_2)$ and taking (C-26) into account, we get:

$$\widehat{W}_{\rm int} = \frac{1}{2} \sum_{i,j,k,l} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2\ W_2(\mathbf{r}_1,\mathbf{r}_2)\ u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_k(\mathbf{r}_1) u_l(\mathbf{r}_2)\ a_i^\dagger a_j^\dagger a_l a_k \tag{C-27}$$

We can thus write the average value of the interaction energy in any normalized state $|\Phi\rangle$ as:

$$\left\langle \widehat{W}_{\rm int} \right\rangle = \langle\Phi|\, \widehat{W}_2\, |\Phi\rangle = \frac{1}{2} \int {\rm d}^3 r_1 \int {\rm d}^3 r_2 \; W_2(\mathbf{r}_1, \mathbf{r}_2)\; G_2(\mathbf{r}_1, \mathbf{r}_2) \tag{C-28}$$

where $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is the spatial correlation function defined by:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i,j,k,l} u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_k(\mathbf{r}_1) u_l(\mathbf{r}_2) \;\; \langle\Phi|\, a_i^\dagger a_j^\dagger a_l a_k\, |\Phi\rangle \tag{C-29}$$

Consequently, knowing the correlation function $G_2(\mathbf{r}_1, \mathbf{r}_2)$ associated with the quantum state $|\Phi\rangle$ allows computing directly, by a double spatial integration, the average interaction energy in that state.

Actually, as we shall see in more detail in § B-3 of Chapter XVI, $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is simply the *double density*, equal to the probability density of finding any particle in $\mathbf{r}_1$ and another one in $\mathbf{r}_2$. The physical interpretation of (C-28) is simple: the average interaction energy is equal to the sum over all the particles' pairs of the interaction energy $W_{\rm int}(\mathbf{r}_1, \mathbf{r}_2)$ of a pair, multiplied by the probability of finding such a pair at points $\mathbf{r}_1$ and $\mathbf{r}_2$ (the factor 1/2 avoids the double counting of each pair).

β. Specific case: the Fock states

Let us assume the state $|\Phi\rangle$ is a Fock state, with specified occupation numbers n_i:

$$|\Phi\rangle = |n_1 : u_1;\; n_2 : u_2;\; ..;\; n_i : u_i;\; ..\rangle \tag{C-30}$$

We can compute explicitly, as a function of the n_i, the average values:

$$\langle\Phi|\, a_i^\dagger a_j^\dagger a_l a_k\, |\Phi\rangle \tag{C-31}$$

contained in (C-29). We first notice that to get a non-zero result, the two operators $a^\dagger$ must create particles in the same states from which they were removed by the two annihilation operators a. Otherwise the action of the four operators on the ket $|\Phi\rangle$ will yield a new Fock state orthogonal to the initial one, and hence a zero result. We must therefore impose either $i = k$ and $j = l$, or the opposite $i = l$ and $j = k$, or eventually the special case where all the subscripts are equal. The first case leads to what we call the "direct term", and the second, the "exchange term". We now compute their values.

(i) Direct term, $i = k$ and $j = l$, shown on the left diagram of Figure 3. If $i = j = k = l$, the four operators acting on $|\Phi\rangle$ reconstruct the same ket, multiplied by the factor $n_i\,(n_i - 1)$; this yields a zero result for fermions. If $i \neq j$, we can move the operator $a_k = a_i$ just to the right of the first operator $a_i^\dagger$ to form the particle number operator $\hat{n}_i$. This permutation in the operators' order does not change anything: for bosons, we are moving commuting operators, and for fermions, two anticommutations introduce two minus signs, which cancel each other. The same goes for the operators with subscript j, leading to the particle number $\hat{n}_j$. Finally, the direct term is equal to:

$$G_2^{\rm dir}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i \neq j} |u_i(\mathbf{r}_1)|^2\, |u_j(\mathbf{r}_2)|^2\, n_i n_j + \sum_i n_i\,(n_i - 1)\, |u_i(\mathbf{r}_1)|^2\, |u_i(\mathbf{r}_2)|^2 \tag{C-32}$$

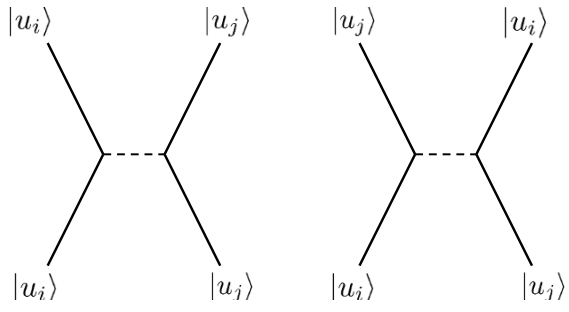


Figure 3: Schematic representation of a direct term (left diagram where each particle remains in the same individual state) and an exchange term (right diagram where the particles exchange their individual states). As in Figure 2, the solid lines represent the particles free propagation, and the dashed lines their binary interaction.

where the second sum is zero for fermions (n_i is equal to 0 or 1).

(ii) Exchange term, $i = l$ and $j = k$, shown on the right diagram of Figure 3. The case where all four subscripts are equal is already included in the direct term. To get the operators' product $\hat{n}_i\hat{n}_j$ starting from the product $a_i^\dagger a_j^\dagger a_i a_j$, we just have to permute the two central operators $a_j^\dagger a_i$; when $i \neq j$ this operation is of no consequence for bosons, but introduces a change of sign for fermions (anticommutation). The exchange term can therefore be written as $\eta G_2^{\rm ex}(\mathbf{r}_1, \mathbf{r}_2)$, with:

$$G_2^{\rm ex}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i \neq j} n_i n_j \left[u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_i(\mathbf{r}_2) u_j(\mathbf{r}_1) \right] \tag{C-33}$$

Finally, the spatial correlation function (or double density) $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is the sum of the direct and exchange terms:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = G_2^{\rm dir}(\mathbf{r}_1, \mathbf{r}_2) + \eta G_2^{\rm ex}(\mathbf{r}_1, \mathbf{r}_2) \tag{C-34}$$

where the factor η in front of the exchange term is 1 for bosons and -1 for fermions. The direct term only contains the product $|u_i(\mathbf{r}_1)|^2 |u_j(\mathbf{r}_2)|^2$ of the probability densities associated with the individual wave functions $u_i(\mathbf{r}_1)$ and $u_j(\mathbf{r}_2)$; it corresponds to non-correlated particles. We must add to it the exchange term, which has a more complex mathematical form and reveals correlations between the particles, even when they do not interact with each other. These correlations come from explicitly taking into account the fact that the particles are identical (symmetrization or antisymmetrization of the state vector). They are sometimes called "statistical correlations " and their spatial dependence will be studied in more detail in Complement $\mathrm{A_{XVI}}$.

Conclusion

The creation and annihilation operators introduced in this chapter lead to compact and general expressions for operators acting on any particle number N. These expressions involve the occupation numbers of the individual states but the particles are no longer

numbered. This considerably simplifies the computations performed on "N-body systems", like N interacting bosons or fermions. The introduction of approximations such as the mean field approximation used in the Hartree-Fock method (Complement $\mathrm{D_{XV}}$) will also be facilitated.

We have shown the complete equivalence between this approach and the one where we explicitly take into account the effect of permutations between numbered particles. It is important to establish this link for the study of certain physical problems. In spite of the overwhelming efficiency of the creation and annihilation operator formalism, the labeling of particles is sometimes useful or cannot be avoided. This is often the case for numerical computations, dealing with numbers or simple functions that require numbered particles and which, if needed, will be symmetrized (or antisymmetrized) afterwards.

In this chapter, we have only considered creation and annihilation operators with discrete subscripts. This comes from the fact that we have only used discrete bases $\{|u_i\rangle\}$ or $\{|v_j\rangle\}$ for the individual states. Other bases could be used, such as the position eigenstates $\{|\mathbf{r}\rangle\}$ of a spinless particle. The creation and annihilation operators will then be labeled by a continuous subscript $\mathbf{r}$. Fields of operators are thus introduced at each space point: they are called "field operators" and will be studied in the next chapter.

COMPLEMENTS OF CHAPTER XV, READER'S GUIDE

A_{XV} : PARTICLES AND HOLES	In an ideal gas of fermions, one can define creation and annihilation operators of holes (absence of a particle). Acting on the ground state, these operators allow building excited states. This is an important concept in condensed matter physics. *Easy to grasp, this complement can be considered to be a preliminary to Complement E_{XV}.*
B_{XV} : IDEAL GAS IN THERMAL EQUILIBRIUM; QUANTUM DISTRIBUTION FUNCTIONS	Studying the thermal equilibrium of an ideal gas of fermions or bosons, we introduce the distribution functions characterizing the physical properties of a particle or of a pair of particles. These distribution functions will be used in several other complements, in particular G_{XV} and H_{XV}. Bose-Einstein condensation is introduced in the case of bosons. The equation of state is discussed for both types of particles.

The list of complements continues on the next page

Series of four complements, discussing the behavior of particles interacting through a mean field created by all the others. Important, since the mean field concept is largely used throughout many domains of physics and chemistry.

C_{XV} : CONDENSED BOSON SYSTEM, GROSS-PITAEVSKII EQUATION

C_{XV} : This complement shows how to use a variational method for studying the ground state of a system of interacting bosons. The system is described by a one-particle wave function in which all the particles of the system accumulate. This wave function obeys the Gross-Pitaevskii equation.

D_{XV} : TIME-DEPENDENT GROSS-PITAEVSKII EQUATION

D_{XV} : This complement generalizes the previous one to the case where the Gross-Pitaevskii wave function is time-dependent. This allows us to obtain the excitation spectrum (Bogolubov spectrum), and to discuss metastable flows (superfluidity).

E_{XV} : FERMION SYSTEM, HARTREE-FOCK APPROXIMATION

E_{XV} : An ensemble of interacting fermions can be treated by a variational method, the Hartree-Fock approximation, which plays an essential role in atomic, molecular and solid state physics. In this approximation, the interaction of each particle with all the others is replaced by a mean field created by the other particles. The correlations introduced by the interactions are thus ignored, but the fermions' indistinguishability is accurately treated. This allows computing the energy levels of the system to an approximation that is satisfactory in many situations.

F_{XV} : FERMIONS, TIME-DEPENDENT HARTREE-FOCK APPROXIMATION

F_{XV} : We often have to study an ensemble of fermions in a time-dependent situation, as for example electrons in a molecule or a solid subjected to an oscillating electric field. The Hartree-Fock mean field method also applies to time-dependent problems. It leads to a set of coupled equations of motion involving a Hartree-Fock mean field potential, very similar to the one encountered for time-independent problems.

The list of complements continues on the next page

The mean field approximation can also be used to study the properties, at thermal equilibrium, of systems of interacting fermions or bosons. The variational method amounts to optimizing the one-particle reduced density operator. It permits generalizing to interacting particles a number of results obtained for an ideal gas (Complement B_{XV}).

G_{XV} : FERMIONS OR BOSONS: MEAN FIELD THERMAL EQUILIBRIUM	G_{XV}: The trial density operator at non-zero temperature can be optimized using a variational method. This leads to self-consistent Hartree-Fock equations, of the same type as those derived in Complement E_{XV}. We thus obtain an approximate value for the thermodynamic potential.
H_{XV} : APPLICATIONS OF THE MEAN FIELD METHOD FOR NON-ZERO TEMPERATURES (FERMIONS AND BOSONS)	H_{XV}: This complement discusses various applications of the method described in the previous complement: spontaneous magnetism of an ensemble of repulsive fermions, equation of state for bosons and instability in the presence of attractive interactions.

Complement A_{XV}

Particles and holes

1 Ground state of a non-interacting fermion gas 1621
2 New definition for the creation and annihilation operators 1622
3 Vacuum excitations 1623

Creation and annihilation operators are frequently used in solid state physics where the notion of particle and hole plays an important role. A good example is the study of metals or semiconductors, where we talk about an electron-hole pair created by photon absorption. A hole means an absence of a particle, but it has properties similar to a particle, like a mass, a momentum, an energy; the holes obey the same fermion statistics as the electrons they replace. Using creation or annihilation operators allows a better understanding of the hole concept. We will remain in the simple framework of a free particle gas, but the concepts can be generalized to the case of particles placed in an external potential or a Hartree-Fock mean potential (Complement E_{XV}).

1. Ground state of a non-interacting fermion gas

Consider a system of non-interacting fermions in their ground state. We assume for simplicity that they are all in the same spin state, and thus introduce no spin index (generalization to several spin states is fairly simple). As we showed in Complement C_{XIV}, this system in its ground state is described by a state where all the occupation numbers of the individual states having an energy lower than the Fermi energy E_F are equal to 1, and all the other individual states are empty. In momentum space, the only occupied states are all the individual states whose wave vector $\mathbf{k}$ is included in a sphere (called the "Fermi sphere") of radius k_F (the "Fermi radius") given by[1]:

$$\frac{\hbar^2 (k_F)^2}{2m} = E_F = \frac{\hbar^2}{2m} \left[\frac{6\pi^2 N}{L^3}\right]^{2/3} \tag{1}$$

where we have used the notation of formula (7) in Complement C_{XIV}: E_F is the Fermi energy (proportional to the particle density to the power 2/3), and L the edge length of the cube containing the N particles. When the system is in its ground state, all the individual states inside the Fermi sphere are occupied, whereas all the other individual states are empty. Choosing for the individual states basis $\{|u_i\rangle\}$ the plane wave basis, noted $\{|u_{\mathbf{k}_i}\rangle\}$ to explicit the wave vector $\mathbf{k}_i$, the occupation numbers are:

$$\begin{array}{ll} n_{\mathbf{k}_i} = 1 & \text{if } |\mathbf{k}_i| \leq k_F \\ n_{\mathbf{k}_i} = 0 & \text{if } |\mathbf{k}_i| > k_F \end{array} \tag{2}$$

[1] In Complement C_{XIV} we had assumed that both spin states of the electron gas were occupied, whereas this is not the case here. This explains why the bracket in formula (1) contains the coefficient $6\pi^2 N$ instead of $3\pi^2 N$.

In a macroscopic system, the number of occupied states is very large, of the order of the Avogadro number ($\simeq 10^{23}$). The ground state energy is given by:

$$E_0 = \sum_{|\mathbf{k}_i| \leq k_F} e_{\mathbf{k}_i} \tag{3}$$

with:

$$e_{\mathbf{k}_i} = \frac{\hbar^2 (k_i)^2}{2m} \tag{4}$$

The sum over $\mathbf{k}_i$ in (3) must be interpreted as a sum over all the $\mathbf{k}_i$ values that obey the boundary conditions in the box of volume L^3, as well as the restriction on the length of the vector $\mathbf{k}_i$ which must be smaller or equal to k_F.

2. New definition for the creation and annihilation operators

We now consider this ground state as a new "vacuum" $\widetilde{|0\rangle}$ and introduce creation operators that, acting on this vacuum, create excited states for this system. We define:

$$\begin{array}{l} \text{if } |\mathbf{k}_i| \leq k_F \begin{cases} b^\dagger_{\mathbf{k}_i} = a_{\mathbf{k}_i} \\ b_{\mathbf{k}_i} = a^\dagger_{\mathbf{k}_i} \end{cases} \\ \text{if } |\mathbf{k}_i| > k_F \begin{cases} c_{\mathbf{k}_i} = a_{\mathbf{k}_i} \\ c^\dagger_{\mathbf{k}_i} = a^\dagger_{\mathbf{k}_i} \end{cases} \end{array} \tag{5}$$

Outside the Fermi sphere, the new operators $c^\dagger_{\mathbf{k}_i}$ and $c_{\mathbf{k}_i}$ are therefore simple operators of creation (or annihilation) of a particle in a momentum state that is not occupied in the ground state. Inside the Fermi sphere, the results are just the opposite: operator $b^\dagger_{\mathbf{k}_i}$ creates a missing particle, that we shall call a "hole"; the adjoint operator $b_{\mathbf{k}_i}$ repopulates that level, hence destroying the hole. It is easy to show that the anticommutation relations for the new operators are:

$$\begin{array}{l} \left[b_{\mathbf{k}_i}, b_{\mathbf{k}_j}\right]_+ = \left[b^\dagger_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = 0 \\ \left[b_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = \delta_{\mathbf{k}_i \mathbf{k}_j} \end{array} \tag{6}$$

as well as:

$$\begin{array}{l} \left[c_{\mathbf{k}_i}, c_{\mathbf{k}_j}\right]_+ = \left[c^\dagger_{\mathbf{k}_i}, c^\dagger_{\mathbf{k}_j}\right]_+ = 0 \\ \left[c_{\mathbf{k}_i}, c^\dagger_{\mathbf{k}_j}\right]_+ = \delta_{\mathbf{k}_i \mathbf{k}_j} \end{array} \tag{7}$$

which are the same as for ordinary fermions. Finally, the cross anticommutation relations are:

$$\left[c_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = \left[c^\dagger_{\mathbf{k}_i}, b_{\mathbf{k}_j}\right]_+ = \left[c_{\mathbf{k}_i}, b_{\mathbf{k}_j}\right]_+ = \left[c^\dagger_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = 0 \tag{8}$$

3. Vacuum excitations

Imagine, for example, that with this new point of view we apply an annihilation operator $b_{\mathbf{k}_i}$, with $|\mathbf{k}_i| \leq k_F$, to the "new vacuum" $\widetilde{|0\rangle}$. The result must be zero since it is impossible to annihilate a non-existent hole. From the old point of view and according to (5), this amounts to applying the creation operator $a^\dagger_{\mathbf{k}_i}$ to a system where the individual state $|\mathbf{k}_i\rangle$ is already occupied, and the result is indeed zero, as expected. On the other hand, if we apply the creation operator $b^\dagger_{\mathbf{k}_i}$, with $|\mathbf{k}_i| \leq k_F$, to the new vacuum, the result is not zero: from the old point of view, it removes a particle from an occupied state, and in the new point of view it creates a hole that did not exist before. The two points of view are consistent.

Instead of talking about particles and holes, one can also use a general term, excitations (or "quasi-particles"). The creation operator of an excitation of $|\mathbf{k}_i| \leq k_F$ is the creation operator $b^\dagger_{\mathbf{k}_i} = a_{\mathbf{k}_i}$ of a hole ; the creation operator of an excitation of $|\mathbf{k}_i| > k_F$ is the creation operator $c^\dagger_{\mathbf{k}_i} = a^\dagger_{\mathbf{k}_i}$ of a particle. The vacuum state defined initially is a common eigenvector of all the particle annihilation operators, with eigenvalues zero; in a similar way, the new vacuum state $\widetilde{|0\rangle}$ is a common eigenvector of all the excitation annihilation operators. We therefore call it the "quasi-particle vacuum".

As we have neglected all particle interactions, the system Hamiltonian is written as:

$$\widehat{H} = \sum_{\mathbf{k}_i} e_{\mathbf{k}_i} n_{\mathbf{k}_i} = \sum_{\mathbf{k}_i} e_{\mathbf{k}_i} a^\dagger_{\mathbf{k}_i} a_{\mathbf{k}_i} = \sum_{|\mathbf{k}_i| \leq k_F} e_{\mathbf{k}_i} b_{\mathbf{k}_i} b^\dagger_{\mathbf{k}_i} + \sum_{|\mathbf{k}_i| > k_F} e_{\mathbf{k}_i} c^\dagger_{\mathbf{k}_i} c_{\mathbf{k}_i} \tag{9}$$

Taking into account the anticommutation relations between the operators $b_{\mathbf{k}_i}$ and $b^\dagger_{\mathbf{k}_i}$ we can rewrite this expression as:

$$\widehat{H} - E_0 = - \sum_{|\mathbf{k}_i| \leq k_F} e_{\mathbf{k}_i} b^\dagger_{\mathbf{k}_i} b_{\mathbf{k}_i} + \sum_{|\mathbf{k}_i| > k_F} e_{\mathbf{k}_i} c^\dagger_{\mathbf{k}_i} c_{\mathbf{k}_i} \tag{10}$$

where E_0 has been defined in (3) and simply shifts the origin of all the system energies. Relation (10) shows that holes (excitations with $|\mathbf{k}_i| \leq k_F$) have a negative energy, as expected since they correspond to missing particles. Starting from its ground state, to increase the system energy keeping the particle number constant, we must apply the operator $b^\dagger_{\mathbf{k}_i} c^\dagger_{\mathbf{k}_j}$ that creates both a particle and a hole: the system energy is then increased by the quantity $e_j - e_i$; inversely, to decrease the system energy, the adjoint operator $c_{\mathbf{k}_j} b_{\mathbf{k}_i}$ must be applied.

Comments:

(i) We have discussed the notion of hole in the context of free particles, but nothing in the previous discussion requires the one-particle energy spectrum to be simply quadratic as in (4). In semi-conductor physics for example, particles often move in a periodic potential, and occupy states in the "valence band" when their energy is lower than the Fermi level E_F, whereas the others occupy the "conduction band", separated from the previous band by an "energy gap". Sending a photon with an energy larger than this gap allows the creation of an electron-hole pair, easily studied in the formalism we just introduced.

A somewhat similar case occurs when studying the relativistic Dirac wave equation, where two energy continuums appear: one with energies greater than the electron rest energy mc^2 (where m is the electron mass, and c the speed of light), and one for negative energies less than $-mc^2$ associated with the positron (the antiparticle of the electron, having the opposite charge). The energy spectrum is relativistic, and thus different from formula (4), even inside each of those two continuums. However, the general formalism remains valid, the operators $b^\dagger_{\mathbf{k}_i}$ and $b_{\mathbf{k}_i}$ describing now, respectively, the creation and annihilation of a positron. The Dirac equation however leads to difficulties by introducing for example an infinity of negative energy states, assumed to be all occupied to avoid problems. A proper treatment of this type of relativistic problems must be done in the framework of quantum field theory.

(ii) An arbitrary N-particle Fock state $|\Phi\rangle$ does not have to be the ground state to be formally considered as a "quasi-particle vacuum". We just have to consider any annihilation operator on an already occupied individual state as a creation operator of a hole (i.e. of an excitation); we then define the corresponding hole (or excitation) annihilation operators, which all have in common the eigenvector $|\Phi\rangle$ with eigenvalue zero. This comment will be useful when studying the Wick theorem (Complement C_{XVI}). In § E of Chapter XVII, we shall see another example of a quasi-particle vacuum, but where, this time, the new annihilation operators are no longer acting on individual states but on states of pairs of particles.

Complement B_{XV}

Ideal gas in thermal equilibrium; quantum distribution functions

1 Grand canonical description of a system without interactions 1626
1-a Density operator . . . 1626
1-b Grand canonical partition function, grand potential . . . 1627
2 Average values of symmetric one-particle operators . . . 1628
2-a Fermion distribution function . . . 1629
2-b Boson distribution function . . . 1629
2-c Common expression . . . 1630
2-d Characteristics of Fermi-Dirac and Bose-Einstein distributions 1630
3 Two-particle operators . . . 1631
3-a Fermions . . . 1632
3-b Bosons . . . 1632
3-c Common expression . . . 1634
4 Total number of particles . . . 1635
4-a Fermions . . . 1635
4-b Bosons . . . 1637
5 Equation of state, pressure . . . 1640
5-a Fermions . . . 1640
5-b Bosons . . . 1641

This complement studies the average values of one- or two-particle operators for an ideal gas, in thermal equilibrium. It includes a discussion of several useful properties of the Fermi-Dirac and Bose-Einstein distribution functions, already introduced in Chapter XIV.

To describe thermal equilibrium, statistical mechanics often uses the grand canonical ensemble, where the particle number may fluctuate, with an average value fixed by the chemical potential μ (cf. Appendix VI, where you will find a number of useful concepts for reading this complement). This potential plays, with respect to the particle number, a role similar to the role the inverse of the temperature term $\beta = 1/k_B T$ plays with respect to the energy (k_B is the Boltzmann constant). In quantum statistical mechanics, Fock space is a good choice for the grand canonical ensemble as it easily allows changing the total number of particles. As a direct application of the results of §§ B and C of Chapter XV, we shall compute the average values of symmetric one- or two-particle operators for a system of identical particles in thermal equilibrium.

We begin in § 1 with the density operator for non-interacting particles, and then show in §§ 2 and 3 that the average values of the symmetric operators may be expressed in terms of the Fermi-Dirac and Bose-Einstein distribution functions, increasing their application range and hence their importance. In § 5, we shall study the equation of state for an ideal gas of fermions or bosons at temperature T and contained in a volume $\mathcal{V}$.

1. Grand canonical description of a system without interactions

We first recall how a system of non-interacting particles is described, in quantum statistical mechanics, by the grand canonical ensemble; more details on this subject can be found in Appendix VI, § 1-c.

1-a. Density operator

Using relations (42) and (43) of Appendix VI, we can write the grand canonical density operator ρ_{eq} (whose trace has been normalized to 1) as:

$$\rho_{eq} = \frac{1}{Z} e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \tag{1}$$

where Z is the grand canonical partition function:

$$Z = \mathrm{Tr}\left\{ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \right\} \tag{2}$$

In these relations, $\beta = 1/(k_B T)$ is the inverse of the absolute temperature T multiplied by the Boltzmann constant k_B, and μ, the chemical potential (which may be fixed by a large reservoir of particles). Operators $\widehat{H}$ and $\widehat{N}$ are, respectively, the system Hamiltonian and the particle number operator defined by (B-15) in Chapter XV.

Assuming the particles do not interact, equation (B-1) of Chapter XV allows writing the system Hamiltonian $\widehat{H}$ as a sum of one-particle operators, in each subspace having a total number of particles equal to N:

$$\widehat{H} = \sum_{q=1}^{N} \widehat{h}(q) \tag{3}$$

Let us call $\{|u_k\rangle\}$ the basis of the individual states that are the eigenstates of the operator $\widehat{h}$. Noting $a_k^\dagger$ and a_k the creation and annihilation operators of a particle in these states, $\widehat{H}$ may be written as in (B-14):

$$\widehat{H} = \sum_k e_k \; a_k^\dagger a_k = \sum_k e_k \; \widehat{n}_k \tag{4}$$

where the e_k are the eigenvalues of $\widehat{h}$. Operator (1) can also be written as:

$$\rho_{eq} = \frac{1}{Z} e^{-\beta \sum_k (e_k-\mu) \; a_k^\dagger a_k} = Z^{-1} \prod\nolimits_k e^{-\beta(e_k-\mu) \; a_k^\dagger a_k} \tag{5}$$

We shall now compute the average values of all the one- or two-particle operators for a system described by the density operator (1).

1-b. Grand canonical partition function, grand potential

In statistical mechanics, the "grand potential" Φ associated with the grand canonical equilibrium is defined as the (natural) logarithm of the partition function, multiplied by $-k_B T$ (*cf.* Appendix VI, § 1-c-β):

$$\Phi = -k_B T \ln Z \tag{6}$$

where Z is given by (2). The trace appearing in this equation is easily computed in the basis of the Fock states built from the individual states $\{|u_k\rangle\}$, as we now show. The trace of a tensor product of operators (Chapter II, § F-2-b) is simply the product of the traces of each operator. The Fock space has the structure of a tensor product of the spaces associated with each of the $|u_k\rangle$ (each being spanned by kets having a population n_k ranging from zero to infinity – see comment (i) of § A-1-c in Chapter XV); we must thus compute a product of traces in each of these spaces. For a fixed k, we sum all the diagonal elements over all the values of n_k, then take the product over all k's, which leads to:

$$Z = \prod_k \sum_{n_k} \exp\left[-\beta n_k \left(e_k - \mu\right)\right] \tag{7}$$

α. *Fermions*

For fermions, as n_k can only take the values 0 or 1 (two identical fermions never occupy the same individual state), we get:

$$Z_{\text{fermions}} = \prod_k \left[1 + e^{-\beta(e_k - \mu)}\right] \tag{8}$$

and:

$$\Phi_{\text{fermions}} = -k_B T \sum_k \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \tag{9}$$

The index k must be summed over all the individual states. In case these states are also labeled by orbital and spin subscripts, these must also be included in the summation. Let us consider for example particles having a spin S and contained in a box of volume $\mathcal{V}$ with periodic boundary conditions. The individual stationary states may be written as $|\mathbf{k}, \nu\rangle$, where $\mathbf{k}$ obeys the periodic boundary conditions (Complement C_{XIV}, § 1-c) and the subscript ν takes $(2S+1)$ values. Assuming the particles to be free in the box (no spin Hamiltonian), each ν value yields the same contribution to Φ_{fermions}; in the large volume limit, expression (9) then becomes:

$$\Phi_{\text{fermions}} = -\left(2S+1\right) k_B T \frac{\mathcal{V}}{\left(2\pi\right)^3} \int \mathrm{d}^3 k \, \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \tag{10}$$

β. *Bosons*

For bosons, the summation over n_k in (7) goes from $n_k = 0$ to infinity, which introduces a geometric series whose sum is readily computed. We therefore get:

$$Z_{\text{bosons}} = \prod_k \frac{1}{1 - e^{-\beta(e_k - \mu)}} \tag{11}$$

which leads to:

$$\Phi_{\text{bosons}} = k_B T \sum_k \ln \left[1 - e^{-\beta(e_k - \mu)}\right] \tag{12}$$

For a system of free particles with spin S, confined in a box with periodic boundary conditions, we obtain, in the large volume limit:

$$\Phi_{\text{bosons}} = (2S+1)\, k_B T \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3 k \, \ln \left[1 - e^{-\beta(e_k - \mu)}\right] \tag{13}$$

In a general way, for fermions as well as bosons, the grand potential directly yields the pressure P, as shown in relation (61) of AppendixVI:

$$\Phi = -P\mathcal{V} \tag{14}$$

Using the proper derivatives with respect to the equilibrium parameters (temperature, chemical potential, volume), it also yields the other thermodynamic quantities such as the energy, the specific heats, etc.

2. Average values of symmetric one-particle operators

Symmetric quantum operators for one, and then for two particles, were introduced in a general way in Chapter XV (§§ B and C). The general expression for a one-particle operator $\widehat{F}$ is given by equation (B-12) of that chapter. We can thus write:

$$\left\langle \widehat{F} \right\rangle = \sum_{i,j} \langle u_i | \, \widehat{f} \, | u_j \rangle \left\langle a_i^\dagger a_j \right\rangle \tag{15}$$

with, when the state of the system is given by the density operator (1):

$$\left\langle a_i^\dagger a_j \right\rangle = \mathrm{Tr} \left\{ \rho_{eq} a_i^\dagger a_j \right\} = Z^{-1} \, \mathrm{Tr} \left\{ e^{-\beta\left(\widehat{H} - \mu \widehat{N}\right)} a_i^\dagger a_j \right\} \tag{16}$$

This trace can be computed in the Fock state basis $|n_1, .., n_i .., n_j, ..\rangle$ associated with the eigenstates basis $\{|u_k\rangle\}$ of $\widehat{h}$. If $i \neq j$, operator $a_i^\dagger a_j$ destroys a particle in the individual state $|u_j\rangle$ and creates another one in the different state $|u_i\rangle$; it therefore transforms the Fock state $|n_1, .., n_i, .., n_j, ..\rangle$ into a different, hence orthogonal, Fock state $|n_1, .., n_i - 1.., n_j + 1, ..\rangle$. Operator ρ_{eq} then acts on this ket, multiplying it by a constant. Consequently, if $i \neq j$, all the diagonal elements of the operator whose trace is taken in (16) are zero; the trace is therefore zero. If $i = j$, this average value may be computed as for the partition function, since the Fock space has the structure of a tensor product of individual state's spaces. The trace is the product of the i value contribution by all the other k values contributions. We can thus write, in a general way:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij} \; Z^{-1} \left\{ \sum_{n_i} n_i \exp\left[-\beta n_i \left(e_i - \mu\right)\right] \right\} \times \prod_{k \neq i} \left\{ \sum_{n_k} \exp\left[-\beta n_k \left(e_k - \mu\right)\right] \right\} \tag{17}$$

For $i = j$, this expression yields the average particle number in the individual state $|u_i\rangle$.

2-a. Fermion distribution function

As the occupation number only takes the values 0 and 1, the first bracket in expression (17) is equal to $\left[e^{-\beta(e_i-\mu)}\right]$; as for the other modes ($k \neq i$) contribution, in the second bracket, it has already been computed when we determined the partition function. We therefore obtain:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij}\ Z^{-1} \left[e^{-\beta(e_i-\mu)}\right] \times \prod_{k\neq i} \left[1 + e^{-\beta(e_k-\mu)}\right] \tag{18}$$

Multiplying both the numerator and denominator by $1 + e^{-\beta(e_i-\mu)}$ allows reconstructing the function Z in the numerator, and, after simplification by Z, we get:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij} \frac{e^{-\beta(e_i-\mu)}}{1 + e^{-\beta(e_i-\mu)}} = \delta_{ij}\ f_\beta^{FD}(e_i - \mu) \tag{19}$$

We find again the Fermi-Dirac distribution function f_β^{FD} (§ 1-b of Complement C_{XIV}):

$$f_\beta^{FD}(e-\mu) = \frac{e^{-\beta(e-\mu)}}{1 + e^{-\beta(e-\mu)}} = \frac{1}{e^{\beta(e-\mu)} + 1} \tag{20}$$

This distribution function gives the average population of each individual state $|u_i\rangle$ with energy e; its value is always less than 1, as expected for fermions.

The average value at thermal equilibrium of any one-particle operator is now readily computed by using (19) in relation (15).

2-b. Boson distribution function

The mode $j = i$ contribution can be expressed as:

$$\begin{aligned}\sum_{n_i=0}^{\infty} n_i\ \exp\left[-\beta n_i\,(e_i-\mu)\right] &= \frac{1}{\beta}\frac{\partial}{\partial\mu}\sum_{n_i=0}^{\infty}\exp\left[-\beta n_i\,(e_i-\mu)\right] \\ &= \frac{1}{\beta}\frac{\partial}{\partial\mu}\frac{1}{1-\exp\left[-\beta\,(e_i-\mu)\right]}\end{aligned} \tag{21}$$

We then get:

$$\left\langle a_i^\dagger a_j \right\rangle = \frac{\delta_{i,j}}{Z}\frac{e^{-\beta(e_i-\mu)}}{\left[1-e^{-\beta(e_i-\mu)}\right]^2} \times \prod_{k\neq i}\frac{1}{1-e^{-\beta(e_k-\mu)}} \tag{22}$$

which, using (11), amounts to:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij}\ f_\beta^{BE}(e_i-\mu) \tag{23}$$

where the Bose-Einstein distribution function f_β^{BE} is defined as:

$$f_\beta^{BE}(e-\mu) = \frac{e^{-\beta(e-\mu)}}{1-e^{-\beta(e-\mu)}} = \frac{1}{e^{\beta(e-\mu)}-1} \tag{24}$$

This distribution function gives the average population of the individual state $|u_i\rangle$ with energy e. The only constraint of this population, for bosons, is to be positive. The chemical potential is always less than the lowest individual energy e_k. In case this energy is zero, μ must always be negative. This avoids any divergence of the function f_β^{BE}.

Hence for bosons, the average value of any one-particle operator is obtained by inserting (23) into relation (15).

2-c. Common expression

We define the function f_β as equal to either the function f_β^{FD} for fermions, or the function f_β^{BE} for bosons. We can write for both cases:

$$f_\beta(e-\mu) = \frac{1}{e^{\beta(e-\mu)} - \eta} \tag{25}$$

where the number η is defined as:

$$\begin{aligned} &\eta = -1 \text{ for fermions} \\ &\eta = +1 \text{ for bosons} \end{aligned} \tag{26}$$

2-d. Characteristics of Fermi-Dirac and Bose-Einstein distributions

We already gave in Complement C_{XIV} (Figure 3) the form of the Fermi-Dirac distribution. Figure 1 shows both the variations of this distribution and the Bose-Einstein distribution. For the sake of comparison, it also includes the variations of the classical Boltzmann distribution:

$$f_\beta^{\text{Boltzmann}}(e-\mu) = e^{-\beta(e-\mu)} \tag{27}$$

which takes on intermediate values between the two quantum distributions. For a non-interacting gas contained in a box with periodic boundary conditions, the lowest possible energy e is zero and all the others are positive. Exponential $e^{\beta(e_i-\mu)}$ is therefore always greater than $e^{-\beta\mu}$. We are now going to distinguish several cases, starting with the most negative values for the chemical potential.

(i) For a negative value of $\beta\mu$, with a modulus large compared to 1 (i.e. for $\mu \ll -k_BT$, which corresponds to the right-hand side of the figure), the exponential in the denominator of (25) is always much larger than 1 (whatever the energy e), and the distribution reduces to the classical Boltzmann distribution (27). Bosons and fermions have practically the same distribution; the gas is said to be "non-degenerate".

(ii) For a fermion system, the chemical potential has no upper boundary, but the population of an individual state can never exceed 1. If μ is positive, with $\mu \gg k_BT$:

– for low values of the energy, the factor 1 is much larger than the exponential term; the population of each individual state is almost equal to 1, its maximum value.

– if the energy e_i increases to values of the order μ, the population decreases and when $e_i \gg \mu$, it becomes practically equal to the value predicted by the Boltzmann exponential (27).

Most of the particles occupy, however, the individual states having an energy less or comparable to μ, whose population is close to 1. The fermion system is said to be "degenerate".

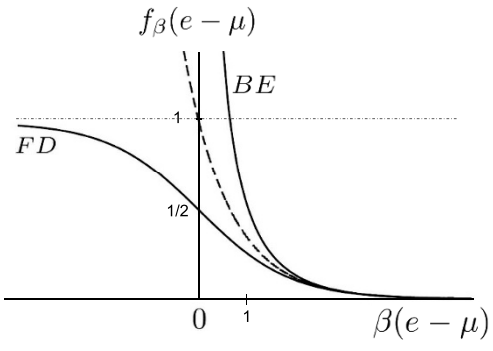


Figure 1: Quantum distribution functions of Fermi-Dirac f_β^{FD} (for fermions, lower curve) and of Bose-Einstein f_β^{BE} (for bosons, upper curve) as a function of the dimensionless variable $\beta(e-\mu)$; the dashed line intermediate curve represents the classical Boltzmann distribution $e^{-\beta(e-\mu)}$. In the right-hand side of the figure, corresponding to large negative values of μ, the particle number is small (the low density region) and the two distributions practically join the Boltzmann distribution. The system is said to be non-degenerate, or classical. As μ increases, we reach the central and left hand side of the figure, and the distributions become more and more different, reflecting the increasing gas degeneracy. For bosons, μ cannot be larger than the one-particle ground state energy, assumed to be zero in this case. The divergence observed for $\mu = 0$ corresponds to Bose-Einstein condensation. For fermions, the chemical potential μ can increase without limit, and for all the energy values, the distribution function tends towards 1 (but never exceeding 1 due to the Pauli exclusion principle).

(iii) For a boson system, the chemical potential cannot be larger than the lowest e_0 individual energy value, which we assumed to be zero. As μ tends towards zero through negative values and $-k_BT \ll \mu < 0$, the distribution function denominator becomes very small leading to very large populations of the corresponding states. The boson gas is then said to be "degenerate". On the other hand, for energies of the order or larger than μ, and as was the case for fermions, the boson distribution becomes practically equal to the Boltzmann distribution.

(iv) Finally, for situations intermediate between the extreme cases described above, the gas is said to be "partially degenerate".

3. Two-particle operators

For a two-particle symmetric operator $\widehat{G}$ we must use formula (C-16) of Chapter XV, which yields:

$$\left\langle \widehat{G} \right\rangle = \frac{1}{2} \sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j |\, \widehat{g}(1,2) \,| 1:u_k;\ 2:u_l \rangle \ \left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle \tag{28}$$

with:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \frac{1}{Z} \mathrm{Tr} \left\{ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} a_i^\dagger a_j^\dagger a_l a_k \right\} \tag{29}$$

As the exponential operator in the trace is diagonal in the Fock basis states $|n_1, .., n_i, .., n_j, ..\rangle$, this trace will be non-zero on the double condition that the states i and j associated with the creation operator be exactly the same as the states k and l associated with the annihilation operators, whatever the order. In other words, to get a non-zero trace, we must have either $i = l$ and $j = k$, or $i = k$ and $j = l$, or both.

3-a. Fermions

As two fermions cannot occupy the same quantum state, the product $a_i^\dagger a_j^\dagger$ is zero if $i = j$; we therefore assume $i \neq j$ which allows, using for ρ_{eq} expression (5) (which is a product), to perform independent calculations for the different modes. The case $i = l$ and $j = k$ yields, using the anticommutation relations:

$$a_i^\dagger a_j^\dagger a_i a_j = -\left(a_i^\dagger a_i\right)\left(a_j^\dagger a_j\right) \tag{30}$$

and the case $i = k$ and $j = l$ yields:

$$a_i^\dagger a_j^\dagger a_j a_i = +\left(a_i^\dagger a_i\right)\left(a_j^\dagger a_j\right) \tag{31}$$

We begin with term (30). As i and j are different, operators $a_i^\dagger a_i$ and $a_j^\dagger a_j$ act on different modes, which belong to different factors in the density operator (5). The average value of the product is thus simply the product of the average values:

$$\left\langle \left(a_i^\dagger a_i\right)\left(a_j^\dagger a_j\right) \right\rangle = \left\langle a_i^\dagger a_i \right\rangle \left\langle a_j^\dagger a_j \right\rangle \tag{32}$$

$$= f_\beta^{FD}(e_i - \mu) f_\beta^{FD}(e_j - \mu) \tag{33}$$

As for the second term (31), it is just the opposite of the first one. Consequently, we finally get:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \left[\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}\right] \; f_\beta^{FD}(e_i - \mu) f_\beta^{FD}(e_j - \mu) \tag{34}$$

The first term on the right-hand side is called the direct term. The second one is the exchange term, and has a minus sign, as expected for fermions.

3-b. Bosons

For bosons, the operators a commute with each other.

α. Average value calculation

If $i \neq j$, a calculation, similar to the one we just did, yields:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \left[\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right] \; f_\beta^{BE}(e_i - \mu) f_\beta^{BE}(e_j - \mu) \tag{35}$$

which differs in two ways from (34): the result now involves the Bose-Einstein distribution, and the exchange term is positive.

If $i = j$, only one individual state comes into a new calculation, which we now perform. Using for ρ_{eq} expression (5) we get, after summing as in (11) a geometric series:

$$\left\langle a_i^\dagger a_i^\dagger a_i a_i \right\rangle = \frac{1}{Z} \sum_{n_i=0}^{\infty} n_i \left(n_i - 1\right) \ \exp\left[-\beta n_i \left(e_i - \mu\right)\right] \prod_{k \neq i} \frac{1}{1 - e^{-\beta(e_k - \mu)}} \tag{36}$$

The sum appearing in this equation can be written as:

$$\begin{aligned} \sum_{n_i=0}^{\infty} n_i \left(n_i - 1\right) \ \exp\left[-\beta n_i \left(e_i - \mu\right)\right] &= \left[\frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} - \frac{1}{\beta} \frac{\partial}{\partial \mu}\right] \sum_{n_i=0}^{\infty} \exp\left[-\beta n_i \left(e_i - \mu\right)\right] \\ &= \left[\frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} - \frac{1}{\beta} \frac{\partial}{\partial \mu}\right] \frac{1}{1 - e^{-\beta(e_i - \mu)}} \end{aligned} \tag{37}$$

The first order derivative term yields:

$$-\frac{1}{\beta} \frac{\partial}{\partial \mu} \frac{1}{1 - e^{-\beta(e_i - \mu)}} = \frac{e^{-\beta(e_i - \mu)}}{\left[1 - e^{\beta(e_i - \mu)}\right]^2} \tag{38}$$

and the second order derivative term is:

$$\frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} \frac{1}{1 - e^{-\beta(e_i - \mu)}} = -\frac{e^{-\beta(e_i - \mu)}}{\left[1 - e^{-\beta(e_i - \mu)}\right]^2} + 2 \frac{\left[e^{-\beta(e_i - \mu)}\right]^2}{\left[1 - e^{-\beta(e_i - \mu)}\right]^3} \tag{39}$$

Summing these two terms yields:

$$\frac{2 \left[e^{-\beta(e_i - \mu)}\right]^2}{\left[1 - e^{-\beta(e_i - \mu)}\right]^3} = \frac{2}{1 - e^{-\beta(e_i - \mu)}} \left[f_\beta^{BE}(e_i - \mu)\right]^2 \tag{40}$$

Multiplying by $1/\left[1 - e^{-\beta(e_i - \mu)}\right]$ the product at the end of the right-hand side of (36) yields the partition function Z, which cancels out the first factor $1/Z$. We are then left with:

$$\left\langle a_i^\dagger a_i^\dagger a_l a_k \right\rangle = 2 \ \delta_{il} \delta_{ik} \left[f_\beta^{BE}(e_i - \mu)\right]^2 \tag{41}$$

This result proves that (35) remains valid even in the case $i = j$.

β. Physical discussion: occupation number fluctuations

For two different physical states i and j, the average value $\left\langle a_i^\dagger a_j^\dagger a_j a_i \right\rangle$ for an ideal gas is simply equal to the product of the average values $\left\langle a_i^\dagger a_i \right\rangle = f_\beta^{BE}(e_i - \mu)$ and $\left\langle a_j^\dagger a_j \right\rangle = f_\beta^{BE}(e_j - \mu)$; this is a consequence of the total absence of interaction between the particles. The same is true for the average value $\left\langle a_i^\dagger a_j^\dagger a_i a_j \right\rangle$.

Now if $i = j$, we note the factor 2 in relation (41). As we now show, this factor leads to the presence of strong fluctuations associated with the operator $\widehat{n}_i$, the particle number in the state $|u_i\rangle$. The calculations shows that:

$$\left\langle (\widehat{n}_i)^2 \right\rangle = \langle a_i^\dagger a_i a_i^\dagger a_i \rangle = \langle a_i^\dagger a_i^\dagger a_i a_i \rangle + \langle a_i^\dagger a_i \rangle$$
$$= 2\left[f_\beta^{BE}(e_i - \mu) \right]^2 + f_\beta^{BE}(e_i - \mu) \tag{42a}$$

The square of the root mean square deviation Δn_i is therefore given by:

$$(\Delta n_i)^2 = \left\langle (\widehat{n}_i)^2 \right\rangle - \langle \widehat{n}_i \rangle^2 = \left[f_\beta^{BE}(e_i - \mu) \right]^2 + f_\beta^{BE}(e_i - \mu) = \langle \widehat{n}_i \rangle^2 + \langle \widehat{n}_i \rangle \tag{42b}$$

The fluctuations of this operator are therefore larger than its average value, which implies that the population of each state $|u_i\rangle$ is necessarily poorly defined[1] at thermal equilibrium. This is particularly true for large $\langle \widehat{n}_i \rangle$: in an ideal boson gas, a largely populated individual state is associated with a very large population fluctuation. This is due to the shape of the Bose-Einstein distribution (24), a decreasing exponential which is maximum at the origin: the most probable occupation number is always $n_i = 0$. Hence it is impossible to get a very large average $\langle \widehat{n}_i \rangle$ without introducing a distribution spreading over many n_i values. Complement H_{XV} (§ 4-a) discusses certain consequences of these fluctuations for an ideal gas. It also shows that as soon as a weak repulsive particle interaction is introduced, the fluctuations greatly diminish and almost completely disappear, since their presence would lead to a very large increase in the potential energy.

3-c. Common expression

To summarize, we can write in all cases:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \left[\delta_{ik}\delta_{jl} + \eta\ \delta_{il}\delta_{jk} \right]\ f_\beta(e_i - \mu) f_\beta(e_j - \mu) \tag{43}$$

with:

$$\begin{cases} \text{for fermions} & \eta = -1, \quad f_\beta = f_\beta^{FD} \\ \text{for bosons} & \eta = +1, \quad f_\beta = f_\beta^{BE} \end{cases} \tag{44}$$

As shown in relation (C-19) of Chapter XV, this average value is simply the matrix element $\langle 1 : u_k;\ 2 : u_l | \widehat{\rho}_2 | 1 : u_i;\ 2 : u_j \rangle$ of the two-particle reduced density operator. To get the general expression for the average of any symmetric two-particle operator, we simply use (43) in (28). Consequently, for independent particles, the average values of all these operators are simply expressed in terms of the quantum Fermi-Dirac and Bose-Einstein distribution functions.

Complement C_{XVI} will show how the Wick theorem allows generalizing these results to operators dealing with any number of particles.

[1] A physical observable is said to have a well defined value in a given quantum state if, in this state, its root mean square deviation is small compared to its average value.

4. Total number of particles

The operator $\widehat{N}$ corresponding to the total number of particles is given by the sum over all the individual states:

$$\widehat{N} = \sum_{i=1}^{\infty} a_i^{\dagger} a_i \tag{45}$$

and its average value is given by:

$$\left\langle \widehat{N} \right\rangle = \mathrm{Tr}\left\{ \widehat{N} \rho_{eq} \right\} = \sum_{i=1}^{\infty} f_\beta(e_i - \mu) \tag{46}$$

As f_β increases as a function of μ, the total number of particles is controlled (for fixed β) by the chemical potential.

4-a. Fermions

For the sake of simplicity, we study the ideal gas properties without taking into account the spin, which assumes that all particles are in the same spin state (the spin can easily be accounted for by adding the contributions of the different individual spin states). For a large physical system, the energy levels are very close and the discrete sum in (46) can be replaced by an integral. This leads to:

$$\left\langle \widehat{N} \right\rangle = N_{ig}^{FD}(\beta, \mu) \tag{47}$$

where the function $N_{ig}^{FD}(\beta, \mu)$ is defined as (the subscript ig stands for ideal gas):

$$N_{ig}^{FD}(\beta, \mu) = \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3 k \frac{1}{e^{\beta(e_k - \mu)} + 1} \tag{48}$$

Figure 2 shows the variations of the function $N_{ig}^{FD}(\beta, \mu)$ as a function of μ, for fixed values of β and the volume $\mathcal{V}$.

To deal with dimensionless quantities, one often introduces the "thermal wavelength" λ_T as:

$$\lambda_T = \hbar \sqrt{\frac{2\pi}{m k_B T}} = \hbar \sqrt{\frac{2\pi\beta}{m}} \tag{49}$$

We can then use in the integral of (48) the dimensionless variable:

$$\kappa = \frac{k \lambda_T}{2\sqrt{\pi}} \tag{50}$$

and write:

$$N_{ig}^{FD}(\beta, \mu) = \frac{\mathcal{V}}{(\lambda_T)^3} \; I_{3/2}(\beta\mu) \tag{51}$$

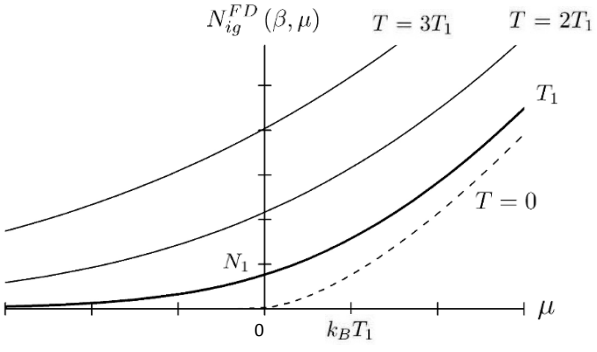


Figure 2: Variations of the particle number $N_{ig}^{FD}(\beta,\mu)$ for an ideal fermion gas, as a function of the chemical potential μ, and for different fixed temperatures T ($\beta = 1/(k_BT)$). For $T = 0$ (lower dashed line curve), the particle number is zero for negative values of μ, and proportional to $\mu^{3/2}$ for positive values of μ. For a non-zero temperature $T = T_1$ (thick line curve), the curve is above the previous one, and never goes to zero. Also shown are the curves obtained for temperatures twice ($T = 2T_1$) and three times ($T = 3T_1$) as large. The units chosen for the axes are the thermal energy k_BT_1 associated with the thick line curve, and the particle number $N_1 = \mathcal{V}/(\lambda_{T_1})^3$, where λ_{T_1} is the thermal wavelength at temperature T_1.
Largely negative values of μ correspond to the classical region where the fermion gas is not degenerate; the classical ideal gas equations are then valid to a good approximation. In the region where $\mu >> k_BT$, the gas is largely degenerate and a Fermi sphere shows up clearly in the momentum space; the total number of particles has only a slight temperature dependence and varies approximately as $\mu^{3/2}$.
This figure was kindly contributed by Geneviève Tastevin.

with[2]:

$$I_{3/2}(\beta\mu) = \pi^{-3/2}\int \mathrm{d}^3\kappa \frac{1}{e^{\kappa^2-\beta\mu}+1} = \frac{2}{\sqrt{\pi}}\int_0^\infty \mathrm{d}x \frac{\sqrt{x}}{e^{x-\beta\mu}+1} \tag{52}$$

where, in the second equality, we made the change of variable:

$$x = \kappa^2 \tag{53}$$

Note that the value of $I_{3/2}$ only depends on a dimensionless variable, the product $\beta\mu$.

If the particles have a spin 1/2, both contributions $\left\langle \widehat{N}_+ \right\rangle$ and $\left\langle \widehat{N}_- \right\rangle$ from the two spin states must be added to (46); in the absence of an external magnetic field, the

[2]The subscript 3/2 refers to the subscript used for more general functions $f_m(z)$, often called the Fermi functions in physics. They are defined by $f_m(z) = \sum_{l=1}^{\infty}(-1)^{l+1}z^l/l^m$, where z is the "fugacity" $z = e^{\beta\mu}$. Expanding in terms of z the function $1/\left[1+z^{-1}e^x\right] = ze^{-x}/\left[1+ze^{-x}\right]$ and using the properties of the Euler Gamma function, it can be shown that $I_{3/2}(\beta\mu) = f_{3/2}(z)$.

individual particle energies do not depend on their spin direction, and the total particle number is simply doubled:

$$\left\langle \widehat{N} \right\rangle = \left\langle \widehat{N}_+ \right\rangle + \left\langle \widehat{N}_- \right\rangle = 2N_{ig}^{FD}(\beta, \mu) \tag{54}$$

4-b. Bosons

For the sake of simplicity, we shall also start with spinless particles, but including several spin states is fairly straightforward. For bosons, we must use the Bose-Einstein distribution (24) and their average number is therefore:

$$\left\langle \widehat{N} \right\rangle = \sum_k f_\beta^{BE}(e_k - \mu) = \sum_k \frac{1}{e^{\beta(e_k-\mu)} - 1} \tag{55}$$

We impose periodic boundary conditions in a cubic box of edge length L. The lowest individual energy[3] is $e_k = 0$. Consequently, for expression (55) to be meaningful, μ must be negative or zero:

$$\mu \leq 0 \tag{56}$$

Two cases are possible, depending on whether the boson system is condensed or not.

α. *Non-condensed bosons*

When the parameter μ takes on a sufficiently negative value (much lower than the opposite of the individual energy e_1 of the first excited level), the function in the summation (55) is sufficiently regular for the discrete summation to be replaced by an integral (in the limit of large volumes). The average particle number is then written as:

$$\left\langle \widehat{N} \right\rangle = N_{ig}^{BE}(\beta, \mu) \tag{57}$$

with:

$$N_{ig}^{BE}(\beta, \mu) = \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3k \; \frac{1}{e^{\beta(e_k-\mu)} - 1} \tag{58}$$

Performing the same change of variables as above, this expression becomes:

$$N_{ig}^{BE}(\beta, \mu) = \frac{\mathcal{V}}{(\lambda_T)^3} K_{3/2}(\beta\mu) \tag{59}$$

with[4]:

$$K_{3/2}(\beta\mu) = \pi^{-3/2} \int \mathrm{d}^3\kappa \frac{1}{e^{\kappa^2-\beta\mu} - 1} = \frac{2}{\sqrt{\pi}} \int_0^\infty \mathrm{d}x \frac{\sqrt{x}}{e^{x-\beta\mu} - 1} \tag{60}$$

[3] Defining other boundary conditions on the box walls will lead in general to a non-zero ground state energy; choosing that value as the common origin for the energies and the chemical potential will leave the following computations unchanged.

[4] The subscript 3/2 refers to the subscript used for the functions $g_m(z)$, often called, in physics, the Bose functions (or the polylogarithmic functions). They are defined by the series $g_m(z) = \sum_{l=1}^{\infty} z^l/l^m$. The exact value of the number ζ defined in (61) is thus given by the series $\sum_{l=1}^{\infty} l^{-3/2}$.

The variations of $N_{ig}^{BE}(\beta,\mu)$ as a function of μ are shown in Figure 3. Note that the total particle number tends towards a limit $\zeta \times \left(\mathcal{V}/\lambda_T^3\right)$ as μ tends towards zero through negative values, where ζ is the number:

$$\zeta = K_{3/2}(0) = 2,612... \tag{61}$$

As the function increases with μ, we can write:

$$N_{ig}^{BE}(\beta,\mu) \leq \zeta \frac{\mathcal{V}}{(\lambda_T)^3} \tag{62}$$

There exists an insurmountable upper limit for the total particle number of a non-condensed ideal Bose gas.

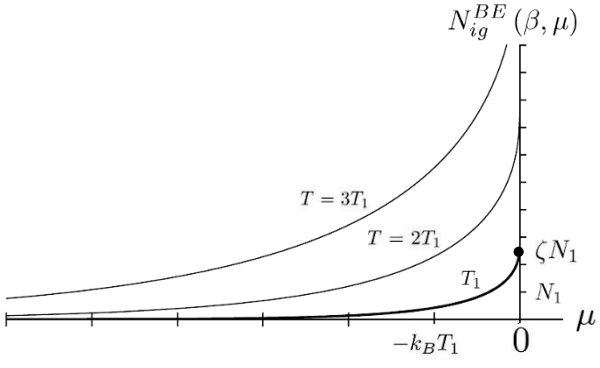


Figure 3: Variations of the total particle number $N_{ig}^{BE}(\beta,\mu)$ in a non-condensed ideal Bose gas, as a function of μ and for fixed $\beta = 1/(k_BT)$. The chemical potential is always negative, and the figure shows curves corresponding to several temperatures $T = T_1$ (thick line), $T = 2T_1$ and $T = 3T_1$. Units on the axes are the same as in Figure 2: the thermal energy k_BT_1 associated with curve $T = T_1$, and the particle number $N_1 = \mathcal{V}/(\lambda_{T_1})^3$, where λ_{T_1} is the thermal wavelength for this same temperature T_1. As the chemical potential tends towards zero, the particle numbers tend towards a finite value. For $T = T_1$, this value is equal to ζN_1 (shown as a dot on the vertical axis), where ζ is given by (61). This figure was kindly contributed by Geneviève Tastevin

β. Condensed bosons

As μ gets closer to zero, the population N_0 of the ground state becomes:

$$N_0 \equiv f_\beta^{BE}(e_0 - \mu) = \frac{1}{e^{-\beta\mu} - 1} \underset{\mu\to 0}{\sim} \frac{1}{\beta\,|\mu|} \tag{63}$$

This population diverges in the limit $\mu = 0$ and, when $|\mu|$ gets small enough, it can become arbitrarily large. It can, for example, become proportional[5] to the volume $\mathcal{V}$, in which case it adds a finite contribution $N_0/\mathcal{V}$ to the particle numerical density (particle number per unit volume) as $\mathcal{V} \to \infty$.

This particularity is limited to the ground state, which, in this case, plays a very different role than the other levels. Let us show, for example, that the first excited state population does not yield a similar effect. Assuming the system to be contained in a cubic box[6] of edge length L, the population of the first excited energy level $e_1 \sim \pi^2\hbar^2/(2mL^2)$ can be written as:

$$f_\beta^{BE}(e_1 - \mu) = \frac{1}{e^{\beta(e_1-\mu)} - 1} \underset{\mu\to 0}{\to} \frac{1}{\beta e_1} \sim \left(\frac{L}{\lambda_T}\right)^2 \tag{64}$$

(we assume the box to be large enough so that $L \gg \lambda_T$, which means $\beta e_1 \ll 1$); this population can therefore be proportional only to the square of L, i.e. to the volume to the power 2/3. It shows that this first excited level cannot make a contribution to the particle density in the limit $L \to \infty$; the same is true for all the other excited levels whose contributions are even smaller. The only arbitrary contribution to the density comes from the ground state.

This arbitrarily large value as $\mu \to 0$ obviously does not appear in relation (59), which predicts that the density $N_{ig}^{BE}(\beta, \mu)/\mathcal{V}$ is always less than a finite value as shown by (62). This is not surprising: as the population varies radically from the first energy level to the next, we can no longer compute the average particle number by replacing in (55) the discrete summation by an integral and a more precise calculation is necessary. Actually, only the ground state population must be treated separately, and the summation over all the excited states (of which none contributes to the density divergence) can still be replaced by an integral as before. Consequently, to get the total population of the physical system we simply add the integral on the right-hand side of (57) to the contribution N_0 of the ground level:

$$\left\langle \widehat{N} \right\rangle = N_{ig}^{BE}(\beta, 0) + N_0 \tag{65}$$

where N_0 is defined in (63).

As $\mu \to 0$, the total population of all the excited levels (others than the ground level) remains practically constant and equal to its upper limit (62); only the ground state has a continuously increasing population N_0, which becomes comparable to the total population of all the excited states when the right-hand sides of (63) and (62) are of the same order of magnitude:

$$\mu \gtrsim -k_B T \frac{\lambda_T^3}{\mathcal{V}} \Rightarrow N_0 \gtrsim \frac{\mathcal{V}}{\lambda_T^3} \tag{66}$$

(μ being of course always negative). When this condition is satisfied, a significant fraction of the particles accumulates in the individual ground level, which is said to have a

[5] The limit where $\mathcal{V} \to \infty$ while the density remains constant is often called the "thermodynamic limit".

[6] As above, we assume periodic conditions on the box walls. Another choice would be to impose zero values for the wave functions on the walls: the numerical coefficients of the individual energies would be changed, but not the line of reasoning.

"macroscopic population" (proportional to the volume). We can even encounter situations where the majority of the particles all occupy the same quantum state. This phenomenon is called "Bose-Einstein condensation" (it was predicted by Einstein in 1935, following Bose's studies of quantum statistics applicable to photons). It occurs when the total density n_{tot} reaches the maximum predicted by formula (62), that is:

$$n_{tot} = \frac{\zeta}{\lambda_T^3} \simeq \frac{2.612}{\lambda_T^3} \tag{67}$$

This condition means that the average distance between particles is of the order of the thermal wavelength λ_T.

Initially, Bose-Einstein condensation was considered to be a mathematical curiosity rather than an important physical phenomenon. Later on, people realized that it played an important role in superfluid liquid Helium 4, although this was a system with constantly interacting particles, hence far from an ideal gas. For a dilute gas, Bose-Einstein condensation was observed for the first time in 1995, and in a great number of later experiments.

5. Equation of state, pressure

The "equation of state" of a fluid at thermal equilibrium is the relation that links, for a given particle number N, its pressure P, volume V, and temperature $T = 1/k_B\beta$. We have just studied the variations of the total particle number. We shall now examine the pressure of a fermion or boson ideal gas.

5-a. Fermions

The grand canonical potential of a fermion ideal gas is given by (9). Equation (14) indicates that, for a system at thermal equilibrium, this grand potential is equal to the opposite of the product of the volume $\mathcal{V}$ and the pressure P. We thus have:

$$\begin{aligned} P\mathcal{V} &= k_B T \sum_k \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \\ &= \frac{\mathcal{V}}{(2\pi)^3} k_B T \int \mathrm{d}^3 k \, \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \end{aligned} \tag{68}$$

(where the second equality is valid in the limit of large volumes). Simplifying by $\mathcal{V} = L^3$, we get the pressure of a fermion system contained in a box of macroscopic dimension:

$$\begin{aligned} P &= (2\pi)^{-3} k_B T \int \mathrm{d}^3 k \, \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \\ &= \frac{1}{\beta \lambda_T^3} I_{5/2}(\beta\mu) \end{aligned} \tag{69}$$

with:

$$\begin{aligned} I_{5/2}(\beta\mu) &= \pi^{-3/2} \int \mathrm{d}^3 \kappa \, \ln \left[1 + e^{\beta\mu - \kappa^2}\right] \\ &= \frac{2}{\sqrt{\pi}} \int_0^\infty \mathrm{d}x \, \sqrt{x} \, \ln \left[1 + e^{\beta\mu - x}\right] \end{aligned} \tag{70}$$

where x has been defined in (53).

To obtain the equation of state, we must find a relation between the pressure P, the volume $\mathcal{V}$, and the temperature T of the physical system, assuming the particle number to be fixed. We have, however, used the grand canonical ensemble (cf. Appendix VI), where the temperature is determined by the parameter β and the volume is fixed, but where the particle number can vary: its average value is a function of a parameter, the chemical potential μ (for fixed values of β and $\mathcal{V}$). Mathematically, the pressure P appears as a function of $\mathcal{V}$, T and μ and not as the function of $\mathcal{V}$, T and the particle number we were looking for. We can nevertheless vary μ, and obtain values of the pressure and particle number of the system and consequently explore, point by point, the equation of state in this parametric form. To obtain an explicit form of the equation of state would require the elimination of the chemical potential using both (47) and (69); there is generally no algebraic solution, and people just use the parametric form of the equation of state, which allows computing all the possible state variables. There also exists a "virial expansion" in powers of the fugacity $e^{\beta\mu}$, which allows the explicit elimination of μ at all the successive orders; its description is beyond the scope of this book.

5-b. Bosons

The pressure of an ideal boson gas is derived from the grand potential (12), taking into account its relation (14) to the pressure P and volume $\mathcal{V}$:

$$\begin{aligned} P\mathcal{V} &= -k_B T \sum_k \ln \left[1 - e^{-\beta(e_k-\mu)}\right] \\ &= -\frac{\mathcal{V}}{(2\pi)^3} k_B T \int d^3k \, \ln \left[1 - e^{-\beta(e_k-\mu)}\right] \end{aligned} \tag{71}$$

(the second relation being valid in the limit of large volumes). This leads to:

$$\begin{aligned} P &= -\frac{k_B T}{(2\pi)^3} \int d^3k \, \ln \left[1 - e^{-\beta(e_k-\mu)}\right] \\ &= \frac{1}{\beta\lambda_T^3} K_{5/2}(\beta\mu) \end{aligned} \tag{72}$$

with:

$$\begin{aligned} K_{5/2}(\beta\mu) &= -\pi^{-3/2} \int \mathrm{d}^3\kappa \, \ln \left[1 - e^{\beta\mu-\kappa^2}\right] \\ &= \frac{-2}{\sqrt{\pi}} \int_0^\infty \mathrm{d}x \, \sqrt{x} \, \ln \left[1 - e^{\beta\mu-x}\right] \end{aligned} \tag{73}$$

As $\mu \to 0$, the contribution P_0 of the ground level to the pressure written in (71) is:

$$P_0 = -\frac{k_B T}{\mathcal{V}} \ln \left[1 - e^{\beta\mu}\right] \simeq -\frac{k_B T}{\mathcal{V}} \ln \left[-\beta\mu\right] \tag{74}$$

When the chemical potential tends towards zero as in (66), it leads to:

$$P_0 \simeq \frac{k_B T}{\mathcal{V}} \ln \frac{\mathcal{V}}{\lambda_T^3} \tag{75}$$

which therefore goes to zero in the limit of large volumes. For a large system, the ground level contribution to the pressure remains negligible compared to that of all the other individual energy levels, whose number gets bigger as the system gets larger. Contrary to what we encountered for the average total particle number, the condensed particles' contribution to the pressure goes to zero in the limit of large volumes.

As we have seen for fermions, the equation of state must be obtained by eliminating the chemical potential μ between equations (72) yielding the pressure and (65) yielding the total particle number. As opposed to an ideal fermion gas, whose particle number and pressure increase without limit as μ and the density increase, the pressure in a boson system is limited. As soon as the system condenses, only the particle number in the individual ground state continues to grow, but not the pressure. In other words, the physical system acquires an infinite compressibility, and becomes a "marginally pathological" system (a system whose pressure decreases with its volume is unstable). This pathology comes, however, from totally neglecting the bosons' interactions. As soon as repulsive interactions are introduced, no matter how small, the compressibility will take on a finite value and the pathology will disappear.

This complement is a nice illustration of the simplifications incurred by the systematic use, in the calculations, of the creation and annihilation operators. We shall see in the following complements that these simplifications still occur when taking into account the interactions, provided we stay in the framework of the mean field approximation. Complement B_{XVI} will even show that for an interacting system studied without using this approximation, the ideal gas distribution functions are still somewhat useful for expressing the average values of various physical quantities.

Complement C_{XV}

Condensed boson system, Gross-Pitaevskii equation

1 Notation, variational ket 1643

1-a Hamiltonian 1644

1-b Choice of the variational ket (or trial ket) 1644

2 First approach 1645

2-a Trial wave function for spinless bosons, average energy 1645

2-b Variational optimization 1646

3 Generalization, Dirac notation 1648

3-a Average energy 1648

3-b Energy minimization 1649

3-c Gross-Pitaevskii equation 1650

4 Physical discussion 1651

4-a Energy and chemical potential 1651

4-b Healing length 1652

4-c Another trial ket: fragmentation of the condensate 1654

The Bose-Einstein condensation phenomenon for an ideal gas (no interaction) of N identical bosons was introduced in § 4-b-β of Complement B_{XV}. We show in the present complement how to describe this phenomenon when the bosons interact. We shall look for the ground state of this physical system within the mean field approximation, using a variational method (see Complement E_{XI}). After introducing in § 1 the notation and the variational ket, we study in § 2 spinless bosons, for which the wave function formalism is simple and the introduction of the creation and annihilation operators does not lead to any major computation simplifications. This will lead us to a first version of the Gross-Pitaevskii equation. We will then come back in § 3 to Dirac notation and the creation operators, to deal with the more general case where each particle may have a spin. Defining the Gross-Pitaevskii potential operator, we shall obtain a more general version of that equation. Finally, some properties of the Gross-Pitaevskii equation will be discussed in § 4, as well as the role of the chemical potential, the existence of a relaxation (or "healing") length, and the energetic consequences of "condensate fragmentation" (these terms will be defined in § 4-c).

1. Notation, variational ket

We first define the notation and the variational family of state vectors that will lead to relatively simple calculations for a system of identical interacting bosons.

1-a. Hamiltonian

The Hamiltonian operator $\widehat{H}$ we consider is the sum of operators for the kinetic energy $\widehat{H}_0$, the one-body potential energy $\widehat{V}_{\rm ext}$, and the interaction energy $\widehat{W}_{\rm int}$:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\rm ext} + \widehat{W}_{\rm int} \tag{1}$$

The first term $\widehat{H}_0$ is simply the sum of the individual kinetic energy operators associated with each of the particles q:

$$\widehat{H}_0 = \sum_q K_0(q) \tag{2}$$

where :

$$K_0(q) = \frac{\mathbf{P}_q^2}{2m} \tag{3}$$

($\mathbf{P}_q$ is the momentum of particle q). Similarly, $\widehat{V}_{\rm ext}$ is the sum of the external potential operators $V_1(\mathbf{R}_q)$, each depending on the position operator $\mathbf{R}_q$ of particle q:

$$\widehat{V}_{\rm ext} = \sum_{q=1}^{N} V_1(\mathbf{R}_q) \tag{4}$$

Finally, $\widehat{W}_{\rm int}$ is the sum of the interaction energy associated with all the pairs of particles:

$$\widehat{W}_{\rm int} = \frac{1}{2} \sum_{q \neq q'=1}^{N} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{5}$$

(this summation can also be written as a sum over $q < q'$, while removing the prefactor $1/2$).

1-b. Choice of the variational ket (or trial ket)

Let us choose an arbitrary normalized quantum state $|\theta\rangle$:

$$\langle\theta\,|\theta\rangle = 1 \tag{6}$$

and call $a_\theta^\dagger$ the associated creation operator. The N-particle variational kets we consider are defined by the family of all the kets that can be written as:

$$\widetilde{|\Psi\rangle} = \frac{1}{\sqrt{N!}} \left[a_\theta^\dagger\right]^N |0\rangle \tag{7}$$

where $|\theta\rangle$ can vary, only constrained by (6). Consider a basis $\{|\theta_k\rangle\}$ of the individual state space whose first vector is $|\theta_1\rangle = |\theta\rangle$. Relation (A-17) of Chapter XV shows that this ket is simply a Fock state whose only non-zero occupation number is the first one:

$$\widetilde{|\Psi\rangle} = |n_1 = N, n_2 = 0, n_3 = 0, ...\rangle \tag{8}$$

An assembly of bosons that occupy the same individual state is called a "Bose-Einstein condensate".

Relation (8) shows that the kets $\widetilde{|\Psi\rangle}$ are normalized to 1. We are going to vary $|\theta\rangle$, and therefore $\widetilde{|\Psi\rangle}$, so as to minimize the average energy:

$$\widetilde{E} = \widetilde{\langle\Psi|}\widehat{H}\widetilde{|\Psi\rangle} \tag{9}$$

2. First approach

We start with a simple case where the bosons have no spin. We can then use the wave function formalism and keep the computations fairly simple.

2-a. Trial wave function for spinless bosons, average energy

Assuming one single individual state to be populated, the wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ is simply the product of N functions $\theta(\mathbf{r})$:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = \theta(\mathbf{r}_1)\ \theta(\mathbf{r}_2) ... \ \theta(\mathbf{r}_N) \tag{10}$$

with:

$$\theta(\mathbf{r}) = \langle \mathbf{r} \,|\theta\rangle \tag{11}$$

This wave function is obviously symmetric with respect to the exchange of all particles and can be used for a system of identical bosons.

In the position representation, each operator $K_0(q)$ defined by (3) corresponds to $\left(-\hbar^2/2m\right)\Delta_{\mathbf{r}_q}$, where $\Delta_{\mathbf{r}_q}$ is the Laplacian with respect to the position $\mathbf{r}_q$; consequently, we have:

$$\left\langle \widehat{H}_0 \right\rangle = \frac{-\hbar^2}{2m} \sum_{q=1}^{N} \int \mathrm{d}^3 r_1 ... \int \mathrm{d}^3 r_q ... \int \mathrm{d}^3 r_N \\ \times \theta^*(\mathbf{r}_1) ... \theta^*(\mathbf{r}_q) ... \theta^*(\mathbf{r}_N) \times \theta(\mathbf{r}_1) ... \ \Delta\theta(\mathbf{r}_q) ... \ \theta(\mathbf{r}_N) \tag{12}$$

In this expression, all the integral variables others than $\mathbf{r}_q$ simply introduce the square of the norm of the function $\theta(\mathbf{r})$, which is equal to 1. We are just left with one integral over $\mathbf{r}_q$, in which $\mathbf{r}_q$ plays the role of a dummy variable, and thus yields a result independent of q. Consequently, all the q values give the same contribution, and we can write:

$$\left\langle \widehat{H}_0 \right\rangle = \frac{-\hbar^2}{2m} N \int \mathrm{d}^3 r \ \theta^*(\mathbf{r}) \Delta\theta(\mathbf{r}) \tag{13}$$

As for the one-body potential energy, a similar calculation yields:

$$\left\langle \widehat{V}_{\text{ext}} \right\rangle = N \int \mathrm{d}^3 r \ \theta^*(\mathbf{r}) \ V_1(\mathbf{r}) \ \theta(\mathbf{r}) \tag{14}$$

Finally, the interaction energy calculation follows the same steps, but we must keep two integral variables instead of one. The final result is proportional to the number $N(N-1)/2$ of pairs of integral variables:

$$\left\langle \widehat{W}_2 \right\rangle = \frac{N(N-1)}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \ \theta^*(\mathbf{r}) \theta^*(\mathbf{r}') \ W_2(\mathbf{r}, \mathbf{r}') \ \theta(\mathbf{r}) \theta(\mathbf{r}') \tag{15}$$

The variational average energy $\widetilde{E}$ is the sum of these three terms:

$$\widetilde{E} = \left\langle \widehat{H}_0 \right\rangle + \left\langle \widehat{V}_{\text{ext}} \right\rangle + \left\langle \widehat{W}_2 \right\rangle \tag{16}$$

2-b. Variational optimization

We now optimize the energy we just computed, so as to determine the wave functions $\theta(\mathbf{r})$ corresponding to its minimum value.

α. Variation of the wave function

Let us vary the function $\theta(\mathbf{r})$ by a quantity:

$$\theta(\mathbf{r}) \Rightarrow \theta(\mathbf{r}) + e^{i\chi}\ \delta f(\mathbf{r}) \tag{17}$$

where $\delta f(\mathbf{r})$ is an infinitesimal function and χ an arbitrary number. A priori, $\delta f(\mathbf{r})$ must be chosen to take into account the normalization constraint (6), which forces the integral of the $\theta(\mathbf{r})$ modulus squared to remain constant. We can, however, use the Lagrange multiplier method (Appendix V) to impose this constraint. We therefore introduce the multiplier μ (we shall see in § 4-a that this factor can be interpreted as the chemical potential) and minimize the function:

$$\widetilde{F} = \widetilde{E} - \mu \int \mathrm{d}^3 r\ \theta^*(\mathbf{r})\theta(\mathbf{r}) \tag{18}$$

This allows considering the infinitesimal variation $\delta f(\mathbf{r})$ to be free of any constraint. The variation $\delta\widetilde{F}$ of the function $\widetilde{F}$ is now the sum of 4 variations, coming from the three terms of (16) and from the integral in (18). For example, the variation of $\left\langle \widehat{H}_0 \right\rangle$ yields:

$$\delta \left\langle \widehat{H}_0 \right\rangle = -\frac{\hbar^2}{2m} N \int \mathrm{d}^3 r\ \left[e^{-i\chi}\ \delta f^*(\mathbf{r})\ \Delta\theta(\mathbf{r}) + e^{i\chi}\ \theta^*(\mathbf{r})\ \Delta\delta f(\mathbf{r}) \right] \tag{19}$$

which is the sum of a term proportional to $e^{-i\chi}$ and another proportional to $e^{i\chi}$. This is true for all 4 variations and the total variation $\delta\widetilde{F}$ can be expressed as the sum of two terms:

$$\delta\widetilde{F} = \delta c_1 e^{-i\chi} + \delta c_2 e^{i\chi} \tag{20}$$

the first being the $\delta f^*(\mathbf{r})$ contribution and the second, that of $\delta f(\mathbf{r})$. Now if $\widetilde{F}$ is stationary, $\delta\widetilde{F}$ must be zero whatever the choice of χ, which is real. Choosing for example $\chi = 0$ imposes $\delta c_1 + \delta c_2 = 0$, and the choice $\chi = \pi/2$ leads (after multiplication by i) to $\delta c_1 - \delta c_2 = 0$. Adding and subtracting the two relations shows that both coefficients δc_1 and δc_2 must be zero. In other words, we can impose $\delta\widetilde{F}$ to be zero as just $\theta^*(\mathbf{r})$ varies but not $\theta(\mathbf{r})$ – or the opposite[1].

β. Stationary condition: Gross-Pitaevskii equation

We choose to impose the variation $\delta\widetilde{F}$ to be zero as only $\theta^*(\mathbf{r})$ varies and for $\chi = 0$. We must first add contributions coming from (13) and (14), then from (15). For this last contribution, we must add two terms, one coming from the variations due to $\theta^*(\mathbf{r})$, and the other from the variation due to $\theta^*(\mathbf{r}')$. These two terms only differ by the notation

[1] This means that the stationary condition may be found by varying indifferently the real or imaginary part of $\theta(\mathbf{r})$.

in the integral variable and are thus equal: we just keep one and double it. We finally add the term due to the variation of the integral in (18), and we get:

$$\delta\widetilde{F} = N \int \mathrm{d}^3 r \; \delta f^*(\mathbf{r}) \; \times \\ \left\{ \left[\frac{-\hbar^2}{2m}\Delta + V_1(\mathbf{r}) - \mu \right] + (N-1) \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \; \theta^*(\mathbf{r}') \; \theta(\mathbf{r}') \right\} \theta(\mathbf{r}) \qquad (21)$$

This variation must be zero for any value of $\delta f^*(\mathbf{r})$; this requires the function that multiplies $\delta f^*(\mathbf{r})$ in the integral to be zero, and consequently that $\theta(\mathbf{r})$ be the solution of the following equation, written for $\varphi(\mathbf{r})$:

$$\boxed{\left\{ \left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) \right] + (N-1) \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \;\; |\varphi(\mathbf{r}')|^2 \right\} \varphi(\mathbf{r}) = \mu \; \varphi(\mathbf{r})} \qquad (22)$$

This is the time-independent Gross-Pitaevskii equation. It is similar to an eigenvalue Schrödinger equation, but with a potential term:

$$V_1(\mathbf{r}) + (N-1) \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \;\; |\varphi(\mathbf{r}')|^2 \qquad (23)$$

which actually contains the wave function φ in the integral over $\mathrm{d}^3 r'$; it is therefore a nonlinear equation. The physical meaning of the potential term in W_2 is simply that, in the mean field approximation, each particle moves in the mean potential created by all the others, each of them being described by the same wave function $\varphi(\mathbf{r}')$; the factor $(N-1)$ corresponds to the fact that each particle interacts with $(N-1)$ other particles. The Gross-Pitaevskii equation is often used to describe the properties of a boson system in its ground state (Bose-Einstein condensate).

γ. *Zero-range potential*

The Gross-Pitaevskii equation is often written in conjunction with an approximation where the particle interaction potential has a microscopic range, very small compared to the distances over which the wave function $\varphi(\mathbf{r})$ varies. We can then substitute:

$$W_2(\mathbf{r},\mathbf{r}') \Longrightarrow g \; \delta(\mathbf{r}-\mathbf{r}') \qquad (24)$$

where the constant g is called the "coupling constant"; such a potential is sometimes known as a "contact potential" or, in other contexts, a "Fermi potential". We then get:

$$\left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + (N-1)\, g\, |\varphi(\mathbf{r})|^2 \right] \varphi(\mathbf{r}) = \mu\varphi(\mathbf{r}) \qquad (25)$$

Whether in this form[2] or in its more general form (22), the equation includes a cubic term in $\varphi(\mathbf{r})$. It may render the problem difficult to solve mathematically, but it also is the source of many interesting physical phenomena. This equation explains, for example, the existence of quantum vortices in superfluid liquid helium.

[2]Strictly speaking, in what is generally called the Gross-Pitaevskii equation, the coupling constant g is replaced by $4\pi\hbar^2 a_0/m$, where a_0 is the "scattering length"; this length is defined when studying the collision phase shift $\delta_l(k)$ (Chapter VIII, § *C*), as the limit of $\delta_0(k) \sim -ka_0$ when $k \to 0$. This scattering length is a function of the interaction potential $W_2(\mathbf{r},\mathbf{r}')$, but generally not merely proportional to it, as opposed to the matrix elements of $W_2(\mathbf{r},\mathbf{r}')$. It is then necessary to make a specific demonstration for this form of the Gross-Pitaevskii equation, using for example the "pseudo-potential" method.

δ. Other normalization

Rather than normalizing the wave function $\varphi(\mathbf{r})$ to 1 in the entire space, one sometimes chooses a normalization taking into account the particle number by setting:

$$\int \mathrm{d}^3 r \, |\varphi(\mathbf{r})|^2 = N \tag{26}$$

This amounts to multiplying by $\sqrt{N}$ the wave function we have used until now. At each point $\mathbf{r}$ of space, the particle (numerical) density $n(\mathbf{r})$ is then given by:

$$n(\mathbf{r}) = |\varphi(\mathbf{r})|^2 \tag{27}$$

With this normalization, the factor $(N-1)$ in (25) is replaced by $(N-1)/N$, which can generally be taken equal to 1 for large N. The Gross-Pitaevskii equation then becomes:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + g\,|\varphi(\mathbf{r})|^2\right]\varphi(\mathbf{r}) = \mu\varphi(\mathbf{r})} \tag{28}$$

As already mentioned, we shall see in § 4-a that μ is simply the chemical potential.

3. Generalization, Dirac notation

We now go back to the previous line of reasoning, but in a more general case where the bosons may have spins. The variational family is the set of the N-particle state vectors written in (7). The one-body potential may depend on the position $\mathbf{r}$, and, at the same time, act on the spin (particles in a magnetic field gradient, for example).

3-a. Average energy

To compute the average energy value $\widetilde{\langle\Psi|}\widehat{H}\widetilde{|\Psi\rangle}$, we use a basis $\{|\theta_k\rangle\}$ of the individual state space, whose first vector is $|\theta_1\rangle = |\theta\rangle$.

Using relation (B-12) of Chapter XV, we can write the average value $\left\langle \widehat{H}_0 \right\rangle$ as:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{k,l} \langle\theta_k|\, K_0\, |\theta_l\rangle \;\; \widetilde{\langle\Psi|} a_k^\dagger a_l \widetilde{|\Psi\rangle} \tag{29}$$

Since $\widetilde{|\Psi\rangle}$ is a Fock state whose only non-zero population is that of the state $|\theta_1\rangle$, the ket $a_k^\dagger a_l \widetilde{|\Psi\rangle}$ is non-zero only if $l = 1$; it is then orthogonal to $\widetilde{|\Psi\rangle}$ if $k \neq 1$. Consequently, the only term left in the summation corresponds to $k = l = 1$. As the operator $a_1^\dagger a_1$ multiplies the ket by its population N, we get:

$$\left\langle \widehat{H}_0 \right\rangle = N \;\; \langle\theta_1|\, K_0\, |\theta_1\rangle \tag{30}$$

With the same argument, we can write:

$$\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle = N \;\; \langle\theta_1|\, V_1\, |\theta_1\rangle \tag{31}$$

Using relation (C-16) of Chapter XV, we can express the average value of the interaction energy as[3]:

$$\left\langle \widehat{W_2} \right\rangle = \frac{1}{2} \sum_{k,l,m,n} \langle 1:\theta_k; 2:\theta_l |\, W_2(1,2) \,| 1:\theta_m; 2:\theta_n \rangle \ \widetilde{\langle \Psi |} a_k^\dagger a_l^\dagger a_n a_m \widetilde{| \Psi \rangle} \tag{32}$$

In this case, for the second matrix element to be non-zero, both subscripts m and n must be equal to 1 and the same is true for both subscripts k and l (otherwise the operator will yield a Fock state orthogonal to $\widetilde{|\Psi\rangle}$). When all the subscripts are equal to 1, the operator multiplies the ket $\widetilde{|\Psi\rangle}$ by $N(N-1)$. This leads to:

$$\left\langle \widehat{W_2} \right\rangle = \frac{N(N-1)}{2} \langle 1:\theta_1; 2:\theta_1 |\, W_2(1,2) \,| 1:\theta_1; 2:\theta_1 \rangle \tag{33}$$

The average interaction energy is therefore simply the product of the number of pairs $N(N-1)/2$ that can be formed with N particles and the average interaction energy of a given pair.

We can replace $|\theta_1\rangle$ by $|\theta\rangle$, since they are equal. The variational energy, obtained as the sum of (30), (31) and (33), then reads:

$$\widetilde{E} = N \ \langle \theta | \left[K_0 + V_1 \right] | \theta \rangle + \frac{N(N-1)}{2} \langle 1:\theta; 2:\theta |\, W_2(1,2) \,| 1:\theta; 2:\theta \rangle \tag{34}$$

3-b. Energy minimization

Consider a variation of $|\theta\rangle$:

$$|\theta\rangle \Rightarrow |\theta\rangle + e^{i\chi} |\delta\alpha\rangle \tag{35}$$

where $|\delta\alpha\rangle$ is an arbitrary infinitesimal ket of the individual state space, and χ an arbitrary real number. To ensure that the normalization condition (6) is still satisfied, we impose $|\delta\alpha\rangle$ and $|\theta\rangle$ to be orthogonal:

$$\langle \delta\alpha \,|\theta\rangle = 0 \tag{36}$$

so that $\langle \theta \,|\theta\rangle$ remains equal to 1 (to the first order in $|\delta\alpha\rangle$). Inserting (35) into (34) to obtain the variation $\mathrm{d}\widetilde{E}$ of the variational energy, we get the sum of two terms: the first one comes from the variation of the ket $|\theta\rangle$, and is proportional to $e^{i\chi}$; the second one comes from the variation of the bra $\langle\theta|$, and is proportional to $e^{-i\chi}$. The result has the form:

$$\delta\widetilde{E} = \delta c_1 e^{i\chi} + \delta c_2 e^{-i\chi} \tag{37}$$

The stationarity condition for $\widetilde{E}$ must hold for any arbitrary real value of χ. As before (§ 2-b-α), it follows that both δc_1 and δc_2 are zero. Consequently, we can impose the variation $\widetilde{E}$ to be zero as just the bra $\langle\theta|$ varies (but not the ket $|\theta\rangle$), or the opposite.

Varying only the bra, we get the condition:

$$\begin{aligned} 0 = N \ \langle \delta\alpha | \left[K_0 + V_1 - \mu \right] | \theta \rangle + \frac{N(N-1)}{2} \big[& \langle 1:\delta\alpha; 2:\theta |\, W_2(1,2) \,| 1:\theta; 2:\theta \rangle \\ & + \langle 1:\theta; 2:\delta\alpha |\, W_2(1,2) \,| 1:\theta; 2:\theta \rangle \big] \end{aligned} \tag{38}$$

[3]We use the simpler notation $W_2(1,2)$ for $W_2(\mathbf{R}_1, \mathbf{R}_2)$.

As the interaction operator $W_2(1,2)$ is symmetric, the last two terms within the bracket in this equation are equal. We get (after simplification by N):

$$0 = \langle \delta\alpha | \left[K_0 + V_1 - \mu \right] | \theta \rangle + (N-1) \langle 1 : \delta\alpha; 2 : \theta | \, W_2(1,2) \, | 1 : \theta; 2 : \theta \rangle \tag{39}$$

3-c. Gross-Pitaevskii equation

To deal with equation (39), we introduce the Gross-Pitaevskii operator V_{GP}^{θ}, defined as a one-particle operator whose matrix elements in an arbitrary basis $\{|u_i\rangle\}$ are given by:

$$\langle u_i | \, V_{GP}^{\theta} \, | u_j \rangle = (N-1) \langle 1 : u_i; 2 : \theta | \, W_2(1,2) \, | 1 : u_j; 2 : \theta \rangle \tag{40}$$

which leads to:

$$\langle v | \, V_{GP}^{\theta} \, | v' \rangle = (N-1) \langle 1 : v; 2 : \theta | \, W_2(1,2) \, | 1 : v'; 2 : \theta \rangle \tag{41}$$

where $|v\rangle$ and $|v'\rangle$ are two arbitrary one-particle kets – this can be shown by expanding these two kets on the basis $\{|u_i\rangle\}$ and using relation (40). Note that this potential operator does not include an exchange term; this term does not exist when the two interacting particles are in the same individual quantum state. Equation (39) then becomes:

$$0 = \langle \delta\alpha | \left[K_0 + V_1 + V_{GP}^{\theta} \right] | \theta \rangle \tag{42}$$

This stationarity condition must be verified for any value of the bra $\langle \delta\alpha|$, with only the constraint that it must be orthogonal to $|\theta\rangle$ (according to relation (36)). This means that the ket resulting from the action of the operator $\left[K_0 + V_1 + V_{GP}^{\theta} \right]$ on $|\theta\rangle$ must have zero components on all the vectors orthogonal to $|\theta\rangle$; its only non-zero component must be on the ket $|\theta\rangle$ itself, which means it is necessarily proportional to $|\theta\rangle$. In other words, $|\theta\rangle$ must be an eigenvector of that operator, with eigenvalue μ (real since the operator is Hermitian):

$$\left[K_0 + V_1 + V_{GP}^{\theta} \right] | \theta \rangle = \mu \, | \theta \rangle \tag{43}$$

We have just shown that the optimal value $|\varphi\rangle$ of $|\theta\rangle$ is the solution of the Gross-Pitaevskii equation:

$$\left[K_0 + V_1 + V_{GP}^{\varphi} \right] | \varphi \rangle = \mu \, | \varphi \rangle \tag{44}$$

which is a generalization of (28) to particles with spin, and is valid for one- or two-body arbitrary potentials. For each particle, the operator V_{GP}^{φ} represents the mean field created by all the others in the same state $|\varphi\rangle$.

Comment:

The Gross-Pitaevskii operator V_{GP}^{θ} is simply a partial trace over the second particle:

$$V_{GP}^{\theta}(1) = (N-1) \, \mathrm{Tr}_2 \left\{ P^{\theta}(2) W_2(1,2) \right\} \tag{45}$$

where $P^{\theta}(2)$ is the projection operator $P^{\theta}(2)$ of the state of particle 2 onto $|\theta\rangle$:

$$P^{\theta}(2) = \sum_k |1 : u_k\rangle \langle 1 : u_k| \otimes |2 : \theta\rangle \langle 2 : \theta| = \sum_k |1 : u_k; 2 : \theta\rangle \langle 1 : u_k; 2 : \theta| \tag{46}$$

To show this, let us compute the partial trace on the right-hand side of (45). To obtain this trace (Complement E_{III}, § 5-b), we choose for particle 2 a set of basis states $\{|\theta_n\rangle\}$ whose first vector $|\theta_1\rangle$ coincides with $|\theta\rangle$:

$$\langle u_i|\,\mathrm{Tr}_2\left\{P^\theta(2)W_2(1,2)\right\}|u_j\rangle = \sum_n \langle 1:u_i;2:\theta_n|\,P^\theta(2)W_2(1,2)\,|1:u_j;2:\theta_n\rangle \tag{47}$$

Replacing $P^\theta(2)$ by its value (46) yields the product of δ_{ik} (for the scalar product associated with particle 1) and δ_{n1} (for the one associated with particle 2). This leads to:

$$\langle u_i|\,\mathrm{Tr}_2\left\{P^\theta(2)W_2(1,2)\right\}|u_j\rangle = \langle 1:u_i;2:\theta|\,W_2(1,2)\,|1:u_j;2:\theta\rangle \tag{48}$$

which is simply the initial definition (40) of V^θ_{GP}. Relation (45) is therefore another possible definition for the Gross-Pitaevskii potential.

4. Physical discussion

We have established which conditions the variational wave function must obey to make the energy stationary, but we have yet to study the actual value of this energy. This will allow us to show that the parameter μ is in fact the chemical potential associated with the system of interacting bosons. We shall then introduce the concept of a relaxation (or "healing") length, and discuss the effect, on the final energy, of the fragmentation of a single condensate into several condensates, associated with distinct individual quantum states.

4-a. Energy and chemical potential

Since the ket $|\varphi\rangle$ is normalized, multiplying (44) by the bra $\langle\varphi|$ and by N, we get:

$$N\,\langle\varphi|\,[K_0+V_1+V^\varphi_{GP}]\,|\varphi\rangle = N\mu \tag{49}$$

We recognize the first two terms of the left-hand side as the average values of the kinetic energy and the external potential. As for the last term, using definition (41) for V^φ_{GP}, we can write it as:

$$N\,\langle\varphi|\,V^\varphi_{GP}\,|\varphi\rangle = N\,(N-1)\,\langle 1:\varphi;2:\varphi|\,W_2(1,2)\,|1:\varphi;2:\varphi\rangle \tag{50}$$

which is simply twice the potential interaction energy given in (33) when $|\theta_1\rangle = |\varphi\rangle$. This leads to:

$$N\mu = \langle H_0\rangle + \langle V_1\rangle + 2\left\langle\widehat{W}_2\right\rangle = \widetilde{E} + \left\langle\widehat{W}_2\right\rangle \tag{51}$$

To find the energy $\widetilde{E}$, note that $N\mu/2$ is the sum of $\left\langle\widehat{W}_2\right\rangle$ and of half the kinetic and external potential energies. Adding the missing halves, we finally get for $\widetilde{E}$:

$$\widetilde{E} = \frac{N}{2}\left[\mu + \langle\varphi|\,[K_0+V_1]\,|\varphi\rangle\right] \tag{52}$$

An advantage of this formula is to involve only one- (and not two-) particle operators, which simplifies the computations. The interaction energy is implicitly contained in the factor μ.

The quantity μ does not yield directly the average energy, but it is related to it, as we now show. Taking the derivative, with respect to N, of equation (34) written for $|\theta\rangle = |\varphi\rangle$, we get:

$$\frac{\mathrm{d}\widetilde{E}}{\mathrm{d}N} = \langle\varphi|\,[K_0 + V_1]\,|\varphi\rangle + \left(N - \frac{1}{2}\right)\langle\varphi,\varphi|\,W_2(1,2)\,|\varphi,\varphi\rangle \tag{53}$$

For large N, one can safely replace in this equation $(N - 1/2)$ by $(N - 1)$; after multiplication by N, we obtain a sum of average energies:

$$N\frac{\mathrm{d}\widetilde{E}}{\mathrm{d}N} = \langle H_0\rangle + \langle V_1\rangle + 2\left\langle\widehat{W}_2\right\rangle \tag{54}$$

Taking relation (51) into account, this leads to:

$$\frac{\mathrm{d}\widetilde{E}}{\mathrm{d}N} = \mu \tag{55}$$

We know (Appendix VI, § 2-b) that in the grand canonical ensemble, and at zero temperature, the derivative of the energy with respect to the particle number (for a fixed volume) is equal to the chemical potential. The quantity μ, introduced mathematically as a Lagrange multiplier, can therefore be simply interpreted as this chemical potential.

4-b. Healing length

The "healing length" is an important concept that characterizes the way a solution of the time-independent Gross-Pitaevskii equation reacts to a spatial constraint (for example, the solution can be forced to be zero along a wall, or along the line of a vortex core). We now calculate an approximate order of magnitude for this length.

Assuming the potential $V_1(\mathbf{r})$ to be zero in the region of interest, we divide equation (28) by $\varphi(\mathbf{r})$ and get:

$$-\frac{\hbar^2}{2m}\frac{\Delta\varphi(\mathbf{r})}{\varphi(\mathbf{r})} + g\,|\varphi(\mathbf{r})|^2 = \mu \tag{56}$$

Consequently, the left-hand side of this equation must be independent of $\mathbf{r}$. Let us assume $\varphi(\mathbf{r})$ is constant in an entire region of space where the density is n_0, independent of $\mathbf{r}$:

$$n_0 = |\varphi(\mathbf{r})|^2 \tag{57}$$

but constrained by the boundary conditions to be zero along its border. For the sake of simplicity, we shall treat the problem in one dimension, and assume $\varphi(\mathbf{r})$ only depends on the first coordinate x of $\mathbf{r}$; the wave function must then be zero along a plane (supposed to be at $x = 0$). We are looking for an order of magnitude of the distance ξ over which the wave function goes from a practically constant value to zero, i.e. for the spatial range of the wave function transition regime. In the region where $\varphi(\mathbf{r})$ is constant, relation (56) yields:

$$\mu = gn_0 \tag{58}$$

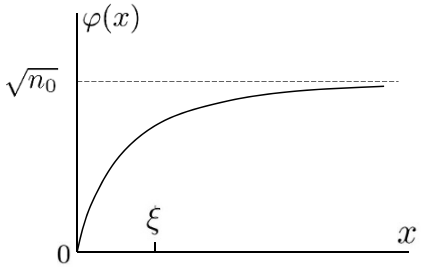


Figure 1: Variation as a function of the position x of the wave function $\varphi(x)$ in the vicinity of a wall (at $x = 0$) where it is forced to be zero. This variation occurs over a distance of the order of the healing length ξ defined in (61); the stronger the particle interactions, the shorter that length. As x increases, the wave function tends towards a constant plateau, of coordinate $\sqrt{n_0}$, represented as a dashed line.

On the other hand, in the whole region where $\varphi(\mathbf{r})$ has significantly decreased, and in particular close to the origin, we have:

$$-\frac{\hbar^2}{2m}\frac{\Delta\varphi(\mathbf{r})}{\varphi(\mathbf{r})} \sim \mu = gn_0 \tag{59}$$

In one dimension[4], we then get the differential equation:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi(x) \sim gn_0\varphi(x) \tag{60}$$

whose solutions are sums of exponential functions $e^{\pm ix/\xi}$, with:

$$\xi = \sqrt{\frac{\hbar^2}{2mgn_0}} \tag{61}$$

The solution that is zero for $x = 0$ is the difference between these two exponentials; it is proportional to $\sin(x/\xi)$, a function that starts from zero and increases over a characteristic length ξ. Figure 1 shows the wave function variation in the vicinity of the wall where it is forced to be zero.

The stronger the interactions, the shorter this "healing length" ξ; it varies as the inverse of the square root of the product of the coupling constant g and the density n_0. From a physical point of view, the healing length results from a compromise between the repulsive interaction forces, which try to keep the wave function as constant as possible in space, and the kinetic energy, which tends to minimize its spatial derivative (while the wave function is forced to be zero at $x = 0$); ξ is equal (except for a 2π coefficient) to the de Broglie wavelength of a free particle having a kinetic energy comparable to the repulsion energy gn_0 in the boson system.

[4] A more precise derivation can be given by verifying that $\varphi(x) = \sqrt{n_0}\,\tanh x/\xi\sqrt{2}$ is a solution of the one-dimensional equation (56).

4-c. Another trial ket: fragmentation of the condensate

We now show that repulsive interactions do stabilize a boson "condensate" where all the particles occupy the same individual state, as opposed to a "fragmented" state where some particles occupy a different state, which can be very close in energy. Instead of using a trial ket (7), where all the particles form a perfect Bose-Einstein condensate in a single quantum state $|\theta\rangle$, we can "fragment" this condensate by distributing the N particles in two distinct individual states. Consequently, we take a trial ket where N_a particles are in the state $|\theta_a\rangle$ and $N_b = N - N_a$ in the orthogonal state $|\theta_b\rangle$:

$$\widetilde{|\Psi\rangle} = \frac{1}{\sqrt{N_a! N_b!}} \left[a^\dagger_{\theta_a}\right]^{N_a} \left[a^\dagger_{\theta_b}\right]^{N_b} |0\rangle \tag{62}$$

We now compute the change in the average variational energy. In formula (29) giving the average kinetic energy, for the operator $a^\dagger_k a_l$ to yield a Fock state identical to $\widetilde{|\Psi\rangle}$, we must have either $k = l = a$, or $k = l = b$. This leads to:

$$\left\langle \widehat{H}_0 \right\rangle = N_a \ \langle\theta_a| K_0 |\theta_a\rangle + N_b \ \langle\theta_b| K_0 |\theta_b\rangle \tag{63}$$

The computation of the one-body potential energy is similar and leads to:

$$\left\langle \widehat{V}_{\text{ext}} \right\rangle = N_a \ \langle\theta_a| V_1 |\theta_a\rangle + N_b \ \langle\theta_b| V_1 |\theta_b\rangle \tag{64}$$

In both cases, the contributions of two populated states are proportional to their respective populations, as expected for energies involving a single particle.

As for the two-body interaction energy, we use again relation (32). It contains the operator $a^\dagger_k a^\dagger_l a_n a_m$, which will reconstruct the Fock state $\widetilde{|\Psi\rangle}$ in the following three cases:

- $k = l = m = n = a$ or b yields the contribution:

$$\frac{N_a (N_a - 1)}{2} \langle 1:\theta_a; 2:\theta_a| W_2(1,2) |1:\theta_a; 2:\theta_a\rangle + \frac{N_b (N_b - 1)}{2} \langle 1:\theta_b; 2:\theta_b| W_2(1,2) |1:\theta_b; 2:\theta_b\rangle \tag{65}$$

- $k = m = a$ and $l = n = b$, or $k = m = b$ and $l = n = a$; these two possibilities yield the same contribution (since the W_2 operator is symmetric), and the $1/2$ factor disappears, leading to the direct term:

$$N_a N_b \langle 1:\theta_a; 2:\theta_b| W_2(1,2) |1:\theta_a; 2:\theta_b\rangle \tag{66}$$

- Finally, $k = n = a$ and $l = m = b$, or $k = n = b$ and $l = m = a$, yield two contributions whose sum introduces the exchange term (here again without the factor $1/2$):

$$N_a N_b \langle 1:\theta_a; 2:\theta_b| W_2(1,2) |1:\theta_b; 2:\theta_a\rangle \tag{67}$$

The direct and exchange terms have been schematized in Figure 3 in Chapter XV (replacing $|u_i\rangle$ by $|\theta_a\rangle$, and $|u_j\rangle$ by $|\theta_b\rangle$), with the direct term on the left, and the exchange term on the right.

The variational energy can thus be written as:

$$\begin{aligned}\widetilde{E} &= N_a\ \left[\langle\theta_a|\left[K_0+V_1\right]|\theta_a\rangle\right] + N_b\ \left[\langle\theta_b|\left[K_0+V_1\right]|\theta_b\rangle\right] \\ &+ \frac{N_a\left(N_a-1\right)}{2}\langle 1:\theta_a;2:\theta_a|\,W_2(1,2)\,|1:\theta_a;2:\theta_a\rangle \\ &+ \frac{N_b\left(N_b-1\right)}{2}\langle 1:\theta_b;2:\theta_b|\,W_2(1,2)\,|1:\theta_b;2:\theta_b\rangle \\ &+ N_aN_b\langle 1:\theta_a;2:\theta_b|\,W_2(1,2)\,|1:\theta_a;2:\theta_b\rangle \\ &+ N_aN_b\langle 1:\theta_a;2:\theta_b|\,W_2(1,2)\,|1:\theta_b;2:\theta_a\rangle \end{aligned} \tag{68}$$

As above, the interaction between particles in the same state $|\theta_a\rangle$ contributes a term proportional to $N_a\left(N_a-1\right)/2$, the number of pairs of particles in that state; the same is true for the interaction term between particles in the same state $|\theta_b\rangle$. The direct term associated with the interaction between two particles in distinct states is proportional to N_aN_b, the number of such pairs. But to this direct term we must add an exchange term, also proportional to N_aN_b, corresponding to an additional interaction. This increased interaction is due to the bunching effect of two bosons in different quantum states, that will be discussed in more detail in § 3-b of Complement $\mathrm{A_{XVI}}$. As they are indistinguishable, two bosons occupying individual orthogonal states show correlations in their positions; this increases the probability of finding them at the same point in space. This increase does not occur when the two bosons occupy the same individual quantum state.

We now assume the diagonal matrix elements of $[K_0+V_1]$ between the two states $|\theta_a\rangle$ and $|\theta_b\rangle$ to be practically the same. For example, if these two states are the lowest energy levels of spinless particles in a cubic box of edge L, the corresponding energy difference is proportional to $1/L^2$ – hence very small in the limit of large L. We also assume all the matrix elements of $W_2(1,2)$ to be equal, which is the case if the (microscopic) range of the particle interaction potential is very small compared to the distances over which the wave functions of the two states vary. We can therefore replace in all the matrix elements the kets $|\theta_a\rangle$ and $|\theta_b\rangle$ by the same ket $|\theta\rangle$. Since $N_a+N_b=N$, we obtain:

$$\begin{aligned}\widetilde{E} &= N\ \langle\theta|\left[K_0+V_1\right]|\theta\rangle \\ &+ \frac{1}{2}\left[N_a\left(N_a-1\right)+N_b\left(N_b-1\right)+2N_aN_b\right]\langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle \\ &+ N_aN_b\langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle \end{aligned} \tag{69}$$

However:

$$N\left(N-1\right)=\left(N_a+N_b\right)\left(N_a+N_b-1\right)=N_a\left(N_a-1\right)+N_b\left(N_b-1\right)+2N_aN_b \tag{70}$$

so that:

$$\widetilde{E}=N\ \langle\theta|\left[K_0+V_1\right]|\theta\rangle+\frac{N\left(N-1\right)}{2}\langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle+\Delta\widetilde{E} \tag{71}$$

with:

$$\Delta\widetilde{E}=N_aN_b\ \langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle \tag{72}$$

We find again result (34), but with an additional term $\Delta\widetilde{E}$, the exchange term. Two cases are then possible, depending on whether the particle interactions are attractive or repulsive. In the first case, the fragmentation of the condensate lowers the energy and leads to a more stable state. Consequently, when the particle interactions are attractive, a condensate where only one individual state is occupied tends to split into two condensates, which might each split again, and so on. This means that the initial single condensate is unstable (we will come back and discuss this instability in § 2-b of Complement F_{XV} for the more general case of thermal equilibrium at non-zero temperature). On the contrary, for repulsive interactions the fragmentation increases the energy and leads to a less stable state: repulsive interactions therefore tend to stabilize the condensate in a single individual quantum state[5]. This result will be interpreted in § 3-b of Complement A_{XVI} in terms of changes of the particle position correlation function (bunching effect of bosons). As for the ideal gas, an intermediate case between the two previous ones, it is a marginal borderline case: adding any infinitesimal attractive interaction, no matter how small, destabilizes any condensate.

[5]We are discussing here the simple case of spinless bosons, contained in a box. When the bosons have several internal quantum states, and in other geometries, more complex situations may arise where the ground state is fragmented [4].

Complement D_{XV}

Time-dependent Gross-Pitaevskii equation

1 Time evolution 1657
1-a Functional variation 1658
1-b Variational computation: the time-dependent Gross-Pitaevskii equation 1659
1-c Phonons and Bogolubov spectrum 1660
2 Hydrodynamic analogy 1664
2-a Probability current 1664
2-b Velocity evolution 1665
3 Metastable currents, superfluidity 1667
3-a Toroidal geometry, quantization of the circulation, vortex . . 1667
3-b Repulsive potential barrier between states of different l . . . 1669
3-c Critical velocity, metastable flow 1671
3-d Generalization; topological aspects 1674

In this complement, we return to the calculations of Complement C_{XV}, concerning a system of bosons all in the same individual state. We now consider the more general case where that state is time-dependent. Using a variational method similar to the one we used in Complement C_{XV}, we shall study the time variations of the N-particle state vector. This amounts to using a time-dependent mean field approximation. We shall establish in § 1 a time-dependent version of the Gross-Pitaevskii equation, and explore some of its predictions such as the small oscillations associated with Bogolubov phonons. In § 2, we shall study local conservation laws derived from this equation for which we will give a hydrodynamic analogy, introducing a characteristic relaxation length. Finally, we will show in § 3 how the Gross-Pitaevskii equation predicts the existence of metastable flows and superfluidity.

1. Time evolution

We assume that the ket describing the physical system of N bosons can be written using relation (7) of Complement C_{XV}:

$$\widetilde{|\Psi(t)\rangle} = \frac{1}{\sqrt{N!}} \left[a_\theta^\dagger(t)\right]^N |0\rangle \tag{1}$$

but we now suppose that the individual ket $|\theta\rangle$ is a function of time $|\theta(t)\rangle$. The creation operator $a_\theta^\dagger(t)$ in the corresponding individual state is then time-dependent:

$$a_\theta^\dagger(t)\,|0\rangle = |\theta(t)\rangle \tag{2}$$

We will let the ket $|\theta(t)\rangle$ vary arbitrarily, as long as it remains normalized at all times:

$$\langle\theta(t)\,|\theta(t)\rangle = 1 \tag{3}$$

We are looking for the time variations of $|\theta(t)\rangle$ that will yield for $\widetilde{|\Psi(t)\rangle}$ variations as close as possible to those predicted by the exact N-particle Schrödinger equation. As the one-particle potential V_1 may also be time-dependent, it will be written as $V_1(t)$.

1-a. Functional variation

Let us introduce the functional of $|\Psi(t)\rangle$:

$$S\left[|\Psi(t)\rangle\right] = \left\{ \int_{t_0}^{t_1} \mathrm{d}t \, \langle \Psi(t)| \left[i\hbar \frac{d}{dt} - H(t) \right] |\Psi(t)\rangle \right\} + \frac{i\hbar}{2} \left[\langle \Psi(t_0) \, |\Psi(t_0)\rangle - \langle \Psi(t_1) \, |\Psi(t_1)\rangle \right] \tag{4}$$

It can be shown that this functional is stationary when $|\Psi(t)\rangle$ is solution of the exact Schrödinger equation (an explicit demonstration of this property is given in § 2 of Complement F_{XV}. If $|\Psi(t)\rangle$ belongs to a variational family, imposing the stationarity of this functional allows selecting, among all the family kets, the one closest to the exact solution of the Schrödinger equation. We shall therefore try and make this functional stationary, choosing as the variational family the set of kets $\widetilde{|\Psi(t)\rangle}$ written as in (1) where the individual ket $|\theta(t)\rangle$ is time-dependent.

As condition (3) means that the norm of $\widetilde{|\Psi(t)\rangle}$ remains constant, the second bracket in expression (4) must be zero. We now have to evaluate the average value of the Hamiltonian $H(t)$ that, actually, has been already computed in (34) of Complement C_{XV}:

$$\widetilde{\langle \Psi(t)|} \left[H(t) \right] \widetilde{|\Psi(t)\rangle} = N \, \langle \theta(t)| \left[K_0 + V_1(t) \right] |\theta(t)\rangle + \frac{N(N-1)}{2} \langle 1:\theta(t); 2:\theta(t)| \, W_2(1,2) \, |1:\theta(t); 2:\theta(t)\rangle \tag{5}$$

The only term left to be computed in (4) contains the time derivative.

This term includes the diagonal matrix element:

$$\widetilde{\langle \Psi(t)|} \left[i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \right] \widetilde{|\Psi(t)\rangle} = \frac{i\hbar}{N!} \langle 0| \left[a_\theta(t) \right]^N \sum_{k=0}^{N-1} \left[a_\theta^\dagger(t) \right]^k \frac{\mathrm{d}a_\theta^\dagger}{\mathrm{d}t} \left[a_\theta^\dagger(t) \right]^{N-k-1} |0\rangle \tag{6}$$

For an infinitesimal time dt, the operator $da_\theta^\dagger/dt$ is proportional to the difference $a_\theta^\dagger(t+dt) - a_\theta^\dagger(t)$, hence to the difference between two creation operators associated with two slightly different orthonormal bases. Now, for bosons, all the creation operators commute with each other, regardless of their associated basis. Therefore, in each term of the summation over k, we can move the derivative of the operator to the far right, and obtain the same result, whatever the value of k. The summation is therefore equal to N times the expression:

$$\frac{1}{N!} \langle 0| \left[a_\theta(t) \right]^N \left[a_\theta^\dagger(t) \right]^{N-1} \frac{\mathrm{d}a_\theta^\dagger}{\mathrm{d}t} |0\rangle \tag{7}$$

Now, we know that:

$$\left[a_\theta(t) \right]^{N-1} \left[a_\theta^\dagger(t) \right]^N |0\rangle = N! \, |1:\theta(t)\rangle = N! \left[a_\theta^\dagger(t) \right] |0\rangle \tag{8}$$

Using in (6) the bra associated with that expression, multiplied by N, we get:

$$\widetilde{\langle\Psi(t)|}\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\right]\widetilde{|\Psi(t)\rangle} = i\hbar\ N\langle 0|\,a_\theta(t)\frac{\mathrm{d}a_\theta^\dagger}{\mathrm{d}t}\,|0\rangle = N\langle\theta(t)|\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\right]|\theta(t)\rangle \tag{9}$$

Regrouping all these results, we finally obtain:

$$S\left[\widetilde{|\Psi(t)\rangle}\right] = -\int_{t_0}^{t_1}\mathrm{d}t\left\{N\ \langle\theta(t)|\left[[K_0+V_1(t)]-i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\right]|\theta(t)\rangle + \frac{N(N-1)}{2}\langle 1:\theta(t);2:\theta(t)|\,W_2(1,2)\,|1:\theta(t);2:\theta(t)\rangle\right\} \tag{10}$$

1-b. Variational computation: the time-dependent Gross-Pitaevskii equation

We now make an infinitesimal variation of $|\theta(t)\rangle$:

$$|\theta(t)\rangle \Rightarrow |\theta(t)\rangle + e^{i\chi}\,|\delta\theta(t)\rangle \tag{11}$$

in order to find the kets $|\theta(t)\rangle$ for which the previous expression will be stationary. As in the search for a stationary state in Complement $\mathrm{C_{XV}}$, we get variations coming from the infinitesimal ket $e^{i\chi}\,|\delta\theta(t)\rangle$ and others from the infinitesimal bra $e^{-i\chi}\,\langle\delta\theta(t)|$; as χ is chosen arbitrarily, the same argument as before leads us to conclude that each of these variations must be zero. Writing only the variation associated with the infinitesimal bra, we see that the stationarity condition requires $|\theta(t)\rangle$ to be a solution of the following equation, written for $|\varphi(t)\rangle$:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\varphi(t)\rangle = [K_0+V_1(t)+V_{GP}^{\varphi}(t)]\,|\varphi(t)\rangle \tag{12}$$

The mean field operator $V_{GP}^{\varphi}(t)$ is defined as in relations (45) and (46) of Complement $\mathrm{C_{XV}}$ by a partial trace:

$$V_{GP}^{\varphi}(1,t) = (N-1)\,\mathrm{Tr}_2\left\{P^{\varphi(t)}(2)\ W_2(1,2)\right\} \tag{13}$$

where $P^{\varphi(t)}$ is the projector onto the ket $|\varphi(t)\rangle$:

$$P^{\varphi(t)} = |\varphi(t)\rangle\langle\varphi(t)| \tag{14}$$

As we take the trace over particle 2 whose state is time-dependent, the mean field is also time-dependent. Relation (12) is the general form of the time-dependent Gross-Pitaevskii equation.

Let us return, as in § 2 of Complement $\mathrm{C_{XV}}$, to the simple case of spinless bosons, interacting through a contact potential:

$$W_2(\mathbf{r},\mathbf{r}') = g\ \delta(\mathbf{r}-\mathbf{r}') \tag{15}$$

Using definition (13) of the Gross-Pitaevskii potential, we can compute its effect in the position representation, as in Complement $\mathrm{C_{XV}}$. The same calculations as in §§ 2-b-β and 2-b-γ of that complement allow showing that relation (12) becomes the Gross-Pitaevskii

time-dependent equation (N is supposed to be large enough to permit replacing $N-1$ by N):

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V_1(\mathbf{r},t)+Ng\left|\varphi(\mathbf{r},t)\right|^2\right]\varphi(\mathbf{r},t) \tag{16}$$

Normalizing the wave function $\varphi(\mathbf{r},t)$ to N:

$$\int \mathrm{d}^3r\ \left|\varphi(\mathbf{r},t)\right|^2=N \tag{17}$$

equation (16) simply becomes:

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V_1(\mathbf{r},t)+g\left|\varphi(\mathbf{r},t)\right|^2\right]\varphi(\mathbf{r},t) \tag{18}$$

Comment:

It can be shown that this time evolution does conserve the norm of $|\varphi(t)\rangle$, as required by (3). Without the nonlinear term of (16), it would be obvious since the usual Schrödinger equation conserves the norm. With the nonlinear term present, it will be shown in § 2-a that the norm is still conserved.

1-c. Phonons and Bogolubov spectrum

Still dealing with spinless bosons, we consider a uniform system, at rest, of particles contained in a cubic box of edge length L. The external potential $V_1(\mathbf{r})$ is therefore zero inside the box and infinite outside. This potential may be accounted for by forcing the wave function to be zero at the walls. In many cases, it is however more convenient to use periodic boundary conditions (Complement C_{XIV}, § 1-c), for which the wave function of the individual lowest energy state is simply a constant in the box. We thus consider a system in its ground state, whose Gross-Pitaevskii wave function is independent of $\mathbf{r}$:

$$\varphi(\mathbf{r},t)=\varphi_0(t)=\frac{1}{L^{3/2}}e^{-i\mu t/\hbar} \tag{19}$$

with a μ value that satisfies equation (16):

$$\mu=\frac{Ng}{L^3}=gn_0 \tag{20}$$

where $n_0=N/L^3$ is the system density. Comparing this expression with relation (58) of Complement C_{XV} allows us to identify μ with the ground state chemical potential. We assume in this section that the interactions between the particles are repulsive (see the comment at the end of the section):

$$g\geq 0 \tag{21}$$

α. Excitation propagation

Let us see which excitations can propagate in this physical system, whose wave function is no longer the function (19), uniform in space. We assume:

$$\varphi(\mathbf{r},t) = \varphi_0(t) + \delta\varphi(\mathbf{r},t) \tag{22}$$

where $\delta\varphi(\mathbf{r},t)$ is sufficiently small to be treated to first order. Inserting this expression in the right-hand side of (16), and keeping only the first-order terms, we find in the interaction term the first-order expression:

$$\begin{aligned}\delta\left[Ng\ \varphi^2(\mathbf{r},t)\varphi^*(\mathbf{r},t)\right] &= Ng\left[(2\varphi_0\delta\varphi)\,\varphi_0^* + \varphi_0^2\delta\varphi^*\right] \\ &= gn_0\left[2\delta\varphi + e^{-2i\mu t/\hbar}\delta\varphi^*\right]\end{aligned} \tag{23}$$

We therefore get, to first-order:

$$i\hbar\frac{\partial}{\partial t}\delta\varphi(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\Delta + 2gn_0\right]\delta\varphi(\mathbf{r},t) + gn_0 e^{-2i\mu t/\hbar}\delta\varphi^*(\mathbf{r},t) \tag{24}$$

which shows that the evolution of $\delta\varphi(\mathbf{r},t)$ is coupled to that of $\delta\varphi^*(\mathbf{r},t)$. The complex conjugate equation can be written as:

$$i\hbar\frac{\partial}{\partial t}\delta\varphi^*(\mathbf{r},t) = \left[\frac{\hbar^2}{2m}\Delta - 2gn_0\right]\delta\varphi^*(\mathbf{r},t) - gn_0 e^{2i\mu t/\hbar}\delta\varphi(\mathbf{r},t) \tag{25}$$

We can make the time-dependent exponentials on the right-hand side disappear by defining:

$$\begin{aligned}\delta\varphi(\mathbf{r},t) &= \overline{\delta\varphi}(\mathbf{r},t)\ e^{-i\mu t/\hbar} \\ \delta\varphi^*(\mathbf{r},t) &= \overline{\delta\varphi}^*(\mathbf{r},t)\ e^{i\mu t/\hbar}\end{aligned} \tag{26}$$

This leads us to a differential equation with constant coefficients, which can be simply expressed in a matrix form:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\overline{\delta\varphi}(\mathbf{r},t)\\ \overline{\delta\varphi}^*(\mathbf{r},t)\end{pmatrix} = \begin{pmatrix}\left[-\frac{\hbar^2}{2m}\Delta + gn_0\right] & gn_0\\ -gn_0 & \left[\frac{\hbar^2}{2m}\Delta - gn_0\right]\end{pmatrix}\begin{pmatrix}\overline{\delta\varphi}(\mathbf{r},t)\\ \overline{\delta\varphi}^*(\mathbf{r},t)\end{pmatrix} \tag{27}$$

where we have used definition (20) for μ to replace $2gn_0 - \mu$ by gn_0. If we now look for solutions having a plane wave spatial dependence:

$$\begin{aligned}\overline{\delta\varphi}(\mathbf{r},t) &= \widetilde{\delta\varphi}(\mathbf{k},t)\ e^{i\mathbf{k}\cdot\mathbf{r}} \\ \overline{\delta\varphi}^*(\mathbf{r},t) &= \widetilde{\delta\varphi}^*(\mathbf{k},t)\ e^{-i\mathbf{k}\cdot\mathbf{r}}\end{aligned} \tag{28}$$

the differential equation can be written as:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\widetilde{\delta\varphi}(\mathbf{k},t)\\ \widetilde{\delta\varphi}^*(\mathbf{k},t)\end{pmatrix} = \begin{pmatrix}\left[\frac{\hbar^2k^2}{2m} + gn_0\right] & gn_0\\ -gn_0 & \left[-\frac{\hbar^2k^2}{2m} - gn_0\right]\end{pmatrix}\begin{pmatrix}\widetilde{\delta\varphi}(\mathbf{k},t)\\ \widetilde{\delta\varphi}^*(\mathbf{k},t)\end{pmatrix} \tag{29}$$

The eigenvalues $\hbar\omega(\mathbf{k})$ of this matrix satisfy the equation:

$$\left[\frac{\hbar^2k^2}{2m}+gn_0-\hbar\omega(\mathbf{k})\right]\left[-\frac{\hbar^2k^2}{2m}-gn_0-\hbar\omega(\mathbf{k})\right]+(gn_0)^2=0 \tag{30}$$

that is:

$$[\hbar\omega(\mathbf{k})]^2-\left[\frac{\hbar^2k^2}{2m}+gn_0\right]^2+(gn_0)^2=0 \tag{31}$$

The solution of this equation is:

$$\hbar\omega(\mathbf{k})=\sqrt{\left[\frac{\hbar^2k^2}{2m}+gn_0\right]^2-(gn_0)^2}=\sqrt{\frac{\hbar^2k^2}{2m}\left[\frac{\hbar^2k^2}{2m}+2gn_0\right]} \tag{32}$$

(the opposite value is also a solution, as expected since we calculate at the same time the evolution of $\widetilde{\delta\varphi}(t)$ and of its complex conjugate; we only use here the positive value). Setting:

$$k_0=\frac{2}{\hbar}\sqrt{gn_0\ m} \tag{33}$$

relation (32) can be written:

$$\omega(\mathbf{k})=\frac{\hbar}{2m}\sqrt{k^2\left(k^2+k_0^2\right)} \tag{34}$$

The spectrum given by (32) is plotted in Figure 1, where one sees the intermediate regime between the linear region at low energy, and the quadratic region at higher energy. It is called the "Bogolubov spectrum" of the boson system.

β. Discussion

Let us compute the spatial and time evolution of the particle density $n(\mathbf{r},\mathbf{t})$ when $\delta\varphi(\mathbf{r},t)$ obeys relation (28). The particle density at each point $\mathbf{r}$ of space is the sum of the densities associated with each particle, that is N times the squared modulus of the wave function $\varphi(\mathbf{r},t)$. To first-order in $\delta\varphi(\mathbf{r},t)$, we obtain:

$$\delta n(\mathbf{r},\mathbf{t})=N\left[\varphi_0(t)e^{-i\mu t/\hbar}\right]^*[\delta\varphi(\mathbf{r},t)]+\text{c.c.} \tag{35}$$

(where c.c. stands for "complex conjugate"). Using (26) and (28), we can finally write:

$$\begin{aligned}\delta n(\mathbf{r},\mathbf{t}) &= N\left[\varphi_0^*(t)e^{i\mu t/\hbar}\right]\left[e^{-i\mu t/\hbar}\ \widetilde{\delta\varphi}(\mathbf{k},0)\ e^{i[\mathbf{k}\cdot\mathbf{r}-\omega(\mathbf{k})t]}\right]+\text{c.c.}\\ &= \frac{N}{L^{3/2}}\left[\widetilde{\delta\varphi}(\mathbf{k},0)\ e^{i[\mathbf{k}\cdot\mathbf{r}-\omega(\mathbf{k})t]}+\text{c.c.}\right]\end{aligned} \tag{36}$$

Consequently, the excitation spectrum we have calculated corresponds to density waves propagating in the system with a phase velocity $\omega(\mathbf{k})/k$.

In the absence of interactions, $(g=k_0=0)$, this spectrum becomes:

$$\hbar\omega(\mathbf{k})=\frac{\hbar^2k^2}{2m} \tag{37}$$

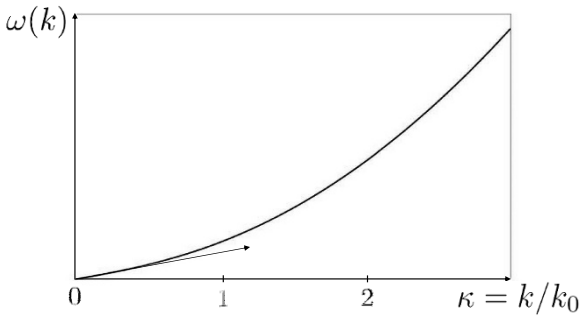


Figure 1: Bogolubov spectrum: variations of the function $\omega(\mathbf{k})$ given by equation (32) as a function of the dimensionless variable $\kappa = k/k_0$. When $\kappa \ll 1$, we get a linear spectrum (the arrow in the figure shows the tangent to the curve at the origin), whose slope is equal to the sound velocity c; when $\kappa \gg 1$, the spectrum becomes quadratic, as for a free particle.

which simply yields the usual quadratic relation for a free particle. Physically, this means that the boson system can be excited by transferring a particle from the individual ground state, with wave function $\varphi_0(\mathbf{r})$ and zero kinetic energy, to any state $\varphi_{\mathbf{k}}(\mathbf{r})$ having an energy $\hbar^2 k^2/2m$.

In the presence of interactions, it is no longer possible to limit the excitation to a single particle, which immediately transmits it to the others. The system's excitations become what we call "elementary excitations", involving a collective motion of all the particles, and hence oscillations in the density of the boson system. If $k \ll k_0$, we see from (34) that:

$$\omega(\mathbf{k}) \sim kc \tag{38}$$

where c is defined as:

$$c = \frac{\hbar}{2m} k_0 = \sqrt{\frac{gn_0}{m}} \tag{39}$$

For small values of k, the interactions have the effect of replacing the quadratic spectrum (37) by a linear spectrum. The phase velocity of all the excitations in this k value domain is a constant c. It is called the "sound velocity " in the interacting boson system, by analogy with a classical fluid where the sound wave dispersion relation is linear, as predicted by the Helmholtz equation. We shall see in § 3 that the quantity c plays a fundamental role in the computations related to superfluidity, especially for the critical velocity determination. If, on the other hand, $k \gg k_0$, the spectrum becomes:

$$\hbar\omega(\mathbf{k}) \sim \frac{\hbar^2 k^2}{2m} + gn_0 + .. \tag{40}$$

(the following corrections being in k_0^2/k^2, k_0^4/k^4, etc.). We find again, within a small correction, the free particle spectrum: exciting the system with enough energy allows

exciting individual particles almost as if they were independent. Figure 1 shows the complete variation of the spectrum (32), with the transition from the linear region at low energies, to the quadratic region at high energies.

Comment:

As we assumed the interactions to be repulsive in (21), the square roots in (32) and (39) are well defined. If the coupling constant g becomes negative, the sound velocity c will become imaginary, and, as seen from (32), so will the frequencies $\omega(k)$ (at least for small values of k). This will lead, for the evolution equation (29), to solutions that are exponentially increasing or decreasing in time, instead of oscillating. An exponentially increasing solution corresponds to an instability of the system. As already encountered in § 4-c of Complement C_{XV}, we see that a boson system becomes unstable in the presence of attractive interactions, however small they might be. In § 4-b of Complement H_{XV}, we shall see that this instability persists even for non-zero temperature. In a general way, an attractive condensate occupying a large region in space tends to collapse onto itself, concentrating into an ever smaller region. However, when it is confined in a finite region (as is the case for experiments where cold atoms are placed in a magneto-optical trap), any change in the wave function that brings the system closer to the instability also increases the gas energy; this results in an energy barrier, which allows the system of condensed attractive bosons to remain in a metastable state.

2. Hydrodynamic analogy

Let us return to the study of the time evolution of the Gross-Pitaevskii wave function and of the density variations $n(\mathbf{r},t)$, without assuming as in § 1-c that the boson system stays very close to uniform equilibrium. We will show that the Gross-Pitaevskii equation can take a form similar to the hydrodynamic equation describing a fluid's evolution. In this discussion, it is useful to normalize the Gross-Pitaevskii wave function to the particle number, as in equation (17). Equation (16) can then be written as:

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V_1(\mathbf{r},t)+gn(\mathbf{r},t)\right]\varphi(\mathbf{r},t) \tag{41}$$

where the local particle density $n(\mathbf{r},t)$ is given by:

$$n(\mathbf{r},t)=|\varphi(\mathbf{r},t)|^2 \tag{42}$$

2-a. Probability current

Since:

$$\frac{\partial}{\partial t}n(\mathbf{r},t)=\varphi^*(\mathbf{r},t)\frac{\partial}{\partial t}\varphi(\mathbf{r},t)+\varphi(\mathbf{r},t)\frac{\partial}{\partial t}\varphi^*(\mathbf{r},t) \tag{43}$$

the time variation of the density may be obtained by first multiplying (41) by $\varphi^*(\mathbf{r},t)$, then its complex conjugate by $\varphi(\mathbf{r},t)$, and then adding the two results; the potential terms in $V_1(\mathbf{r},t)$ and $g\ n(\mathbf{r},t)$ cancel out, and we get:

$$\frac{\partial}{\partial t}n(\mathbf{r},t)=-\frac{\hbar}{2mi}\left[\varphi^*(\mathbf{r},t)\Delta\varphi(\mathbf{r},t)-\varphi(\mathbf{r},t)\Delta\varphi^*(\mathbf{r},t)\right] \tag{44}$$

Let us now define a vector $\mathbf{J}(\mathbf{r},t)$ by:

$$\mathbf{J}(\mathbf{r},t) = \frac{\hbar}{2mi}\left[\varphi^*(\mathbf{r},t)\boldsymbol{\nabla}\varphi(\mathbf{r},t) - \varphi(\mathbf{r},t)\boldsymbol{\nabla}\varphi^*(\mathbf{r},t)\right] \tag{45}$$

If we compute the divergence of this vector, the terms in $\boldsymbol{\nabla}\varphi^* \cdot \boldsymbol{\nabla}\varphi$ cancel out and we are left with terms identical to the right-hand side of (44), with the opposite sign. This leads to the conservation equation:

$$\frac{\partial}{\partial t} n(\mathbf{r},t) + \boldsymbol{\nabla} \cdot \mathbf{J}(\mathbf{r},t) = 0 \tag{46}$$

$\mathbf{J}(\mathbf{r},t)$ is thus the probability current associated with our boson system. Integrating over all space, using the divergence theorem, and assuming $\varphi(\mathbf{r},t)$ (hence the current) goes to zero at infinity, we obtain:

$$\frac{\partial}{\partial t}\int \mathrm{d}^3r\ n(\mathbf{r},t) = \frac{\partial}{\partial t}\int \mathrm{d}^3r\ \ |\varphi(\mathbf{r},t)|^2 = 0 \tag{47}$$

This shows, as announced earlier, that the Gross-Pitaevskii equation conserves the norm of the wave function describing the particle system.

We now set:

$$\varphi(\mathbf{r},t) = \sqrt{n(\mathbf{r},\mathbf{t})}\ e^{i\alpha(\mathbf{r},t)} \tag{48}$$

The gradient of this function is written as:

$$\boldsymbol{\nabla}\varphi(\mathbf{r},t) = e^{i\alpha(\mathbf{r},\mathbf{t})}\left[\boldsymbol{\nabla}\sqrt{n(\mathbf{r},t)} + i\sqrt{n(\mathbf{r},t)}\boldsymbol{\nabla}\alpha(\mathbf{r},t)\right] \tag{49}$$

Inserting this result in (45), we get:

$$\mathbf{J}(\mathbf{r},t) = \frac{\hbar}{m} n(\mathbf{r},t)\boldsymbol{\nabla}\alpha(\mathbf{r},t) \tag{50}$$

or, defining the particle local velocity $\mathbf{v}(\mathbf{r},t)$ as the ratio of the current to the density:

$$\mathbf{v}(\mathbf{r},t) = \frac{\mathbf{J}(\mathbf{r},t)}{n(\mathbf{r},t)} = \frac{\hbar}{m}\boldsymbol{\nabla}\alpha(\mathbf{r},t) \tag{51}$$

We have defined a velocity field, similar to the velocity field of a fluid in motion in a certain region of space; this field velocity is irrotational (zero curl everywhere).

2-b. Velocity evolution

We now compute the time derivative of this velocity. Taking the derivative of (48), we get:

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t) = e^{i\alpha(\mathbf{r},\mathbf{t})} i\hbar\frac{\partial}{\partial t}\sqrt{n(\mathbf{r},t)} - \hbar\sqrt{n(\mathbf{r},t)}e^{i\alpha(\mathbf{r},t)}\frac{\partial}{\partial t}\alpha(\mathbf{r},t) \tag{52}$$

so that we can isolate the time derivative of $\alpha(\mathbf{r},\mathbf{t})$ by the following combination:

$$i\hbar\left[\varphi^*(\mathbf{r},t)\frac{\partial}{\partial t}\varphi(\mathbf{r},t) - \varphi(\mathbf{r},t)\frac{\partial}{\partial t}\varphi^*(\mathbf{r},t)\right] = -2\hbar\ n(\mathbf{r},t)\frac{\partial}{\partial t}\alpha(\mathbf{r},t) \tag{53}$$

The left-hand side of this relation can be computed with the Gross-Pitaevskii equation (18) and its complex conjugate, as we now show. We first take the divergence of the gradient (49) to obtain the Laplacian:

$$\Delta\varphi(\mathbf{r},t) = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\varphi(\mathbf{r},t) = e^{i\alpha(\mathbf{r},t)}\left[\Delta\sqrt{n(\mathbf{r},t)} + 2i\ \left(\boldsymbol{\nabla}\sqrt{n(\mathbf{r},t)}\right)\cdot\boldsymbol{\nabla}\alpha(\mathbf{r},t)\right.$$
$$\left. +i\ \sqrt{n(\mathbf{r},t)}\Delta\alpha(\mathbf{r},t) - \sqrt{n(\mathbf{r},t)}\,(\boldsymbol{\nabla}\alpha(\mathbf{r},t))^2\right] \tag{54}$$

We then insert the time derivative of $\varphi(\mathbf{r},t)$ given by the Gross-Pitaevskii equation (18) in the left-hand side of relation (53), which becomes:

$$-\frac{\hbar^2}{2m}\left[\varphi^*(\mathbf{r},t)\Delta\varphi(\mathbf{r},t) + \varphi(\mathbf{r},t)\Delta\varphi^*(\mathbf{r},t)\right] + 2\left[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t)\right]|\varphi(\mathbf{r},t)|^2$$
$$= -\frac{\hbar^2}{2m}\left[2\sqrt{n(\mathbf{r},t)}\ \Delta\left(\sqrt{n(\mathbf{r},t)}\right) - 2n(\mathbf{r},t)\,(\boldsymbol{\nabla}\alpha(\mathbf{r},t))^2\right]$$
$$+ 2\Big[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t)\Big]n(\mathbf{r},t) \tag{55}$$

This result must be equal to the right-hand side of (53). We therefore get, after dividing both sides by $-2n(\mathbf{r},t)$:

$$\hbar\frac{\partial}{\partial t}\alpha(\mathbf{r},\mathbf{t}) = \frac{\hbar^2}{2m}\left[\frac{1}{\sqrt{n(\mathbf{r},t)}}\Delta\left(\sqrt{n(\mathbf{r},t)}\right) - (\boldsymbol{\nabla}\alpha(\mathbf{r},t))^2\right] - \left[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t)\right] \tag{56}$$

Using (51), we finally obtain the evolution equation for the velocity $\mathbf{v}(\mathbf{r},t)$:

$$m\frac{\partial}{\partial t}\mathbf{v}(\mathbf{r},t) = -\boldsymbol{\nabla}\left[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t) + \frac{m\mathbf{v}^2(\mathbf{r},t)}{2} + \frac{\hbar^2}{2m}\frac{1}{\sqrt{n(\mathbf{r},t)}}\Delta\left(\sqrt{n(\mathbf{r},t)}\right)\right] \tag{57}$$

This equation looks like the classical Newton equation. Its right-hand side includes the sum of the forces corresponding to the external potential $V_1(\mathbf{r},t)$, and to the mean interaction potential with the other particles $gn(\mathbf{r},t)$; the third term in the gradient is the classical kinetic energy gradient[1] (as in Bernoulli's equation of classical hydrodynamics). The only purely quantum term is the last one, as shown by its explicit dependence on $\hbar^2$. It involves spatial derivatives of $n(\mathbf{r},t)$, and is only important if the relative variations $\delta n/n$ of the density occur over small enough distances (for example, this term is zero for a uniform density). This term is sometimes called "quantum potential", or "quantum pressure term" or, in other contexts, "Bohm potential". A frequently used approximation is to consider the spatial variations of $n(\mathbf{r},t)$ to be slow, which amounts to ignoring this quantum potential term: this is the so-called Thomas-Fermi approximation.

We have found for a system of N particles a series of properties usually associated with the wave function of a single particle, and in particular a local velocity directly proportional to its phase gradient[2]. The only difference is that, for the N-particle case,

[1] It is a "total derivative" term (the derivative describing, in a fluid, the motion of each particle). As the velocity field has a zero curl according to (51), a simple vector analysis calculation shows this term to be equal to $m\,(\mathbf{v}\cdot\boldsymbol{\nabla})\,\mathbf{v}$; it can therefore be accounted for by replacing on the left-hand side of (57) the partial derivative $\partial/\partial t$ by the total derivative $\mathrm{d}/\mathrm{d}t = \partial/\partial t + \mathbf{v}\cdot\boldsymbol{\nabla}$.

[2] The quantum potential is still present for a single particle, since making $g = 0$ in (57) does not change this potential. For $g = 0$, the Gross-Pitaevskii equation simply reduces to the standard Schrödinger equation, valid for a single particle.

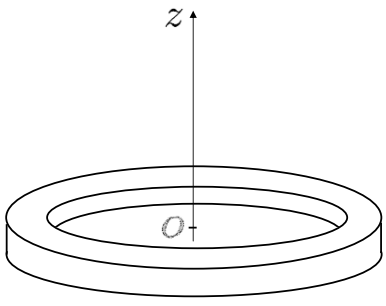


Figure 2: A repulsive boson gas is contained in a toroidal box. All the bosons are supposed to be initially in the same quantum state describing a rotation around the Oz axis. As we explain in the text, this rotation can only slow down if the system overcomes a potential energy barrier that comes from the repulsive interactions between the particles. This prevents any observable damping of the rotation over any accessible time scale; the fluid rotates indefinitely, and is said to be superfluid.

we must add to the external potential $V_1(\mathbf{r},t)$ a local interaction potential $gn(\mathbf{r},t)$, which does not significantly change the form of the equations but introduces some nonlinearity that can lead to completely new physical effects.

3. Metastable currents, superfluidity

Consider now a system of repulsive bosons contained in a toroidal box with a rotational axis Oz (Figure 2); the shape of the torus cross-section (circular, rectangular or other) is irrelevant for our argument and we shall use cylindrical coordinates r, φ and z. We first introduce solutions of the Gross-Pitaevskii equation that correspond to the system rotating inside the toroidal box, around the Oz axis. We will then show that these rotational states are metastable, as they can only relax towards lower energy rotational states by overcoming a macroscopic energy barrier: this is the physical origin of superfluidity.

3-a. Toroidal geometry, quantization of the circulation, vortex

To prevent any confusion with the azimuthal angle φ, we now call χ the Gross-Pitaevskii wave function. The time-independent Gross-Pitaevskii equation then becomes (in the absence of any potential except the wall potentials of the box):

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\chi}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2\chi}{\partial\varphi^2}+\frac{\partial^2\chi}{\partial z^2}\right]+g\left|\chi(\mathbf{r})\right|^2\chi(\mathbf{r})=\mu\chi(\mathbf{r}) \tag{58}$$

We look for solutions of the form:

$$\chi_l(\mathbf{r})=u_l(r,z)e^{il\varphi} \tag{59}$$

where l is necessarily an integer (otherwise the wave function would be multi-valued). Such a solution has an angular momentum with a well defined component along Oz, equal to $l\hbar$ per atom. Inserting this expression in (58), we obtain the equation for $u_l(r,z)$:

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_l(r,z)}{\partial r}\right)+\frac{\partial^2 u_l(r,z)}{\partial z^2}\right]+\left[g\left|u_l(r,z)\right|^2+\frac{l^2\hbar^2}{2mr^2}\right]u_l(r,z) = \mu\ u_l(r,z) \tag{60}$$

which must be solved with the boundary conditions imposed by the torus shape to obtain the ground state (associated with the lowest value of μ). The term in $l^2\hbar^2/2mr^2$ is simply the rotational kinetic energy around Oz. If the tore radius R is very large compared to the size of its cross-section, the term l^2/r^2 may, to a good approximation, be replaced by the constant l^2/R^2. It follows that the same solution of (60) is valid for any value of l as long as the chemical potential is increased accordingly. Each value of the angular momentum thus yields a ground state and the larger l, the higher the corresponding chemical potential. All the coefficients of the equation being real, we shall assume, from now on, the functions $u_l(r,z)$ to be real.

As the wave function is of the form (59), its phase only depends on φ, and expression (51) for the fluid velocity is written as:

$$\mathbf{v}=\frac{1}{r}\frac{l\hbar}{m}\mathbf{e}_\varphi \tag{61}$$

where $\mathbf{e}_\varphi$ is the tangential unit vector (perpendicular both to $\mathbf{r}$ and the Oz axis). Consequently, the fluid rotates along the toroidal tube, with a velocity proportional to l. As $\mathbf{v}$ is a gradient, its circulation along a closed loop "equivalent to zero" (i.e. which can be contracted continuously to a point) is zero. If the closed loop goes around the tore, the path is no longer equivalent to zero and its circulation may be computed along a circle where r and z remain constant, and φ varies from 0 to 2π; as the path length equals $2\pi r$, we get:

$$\int \mathbf{v}\cdot \mathrm{d}\mathbf{s}=\pm l\frac{2\pi\hbar}{m} \tag{62}$$

(with a + sign if the rotation is counterclockwise and a − sign in the opposite case). As l is an integer, the velocity circulation around the center of the tore is quantized in units of h/m. This is obviously a pure quantum property (for a classical fluid, this circulation can take on a continuous set of values).

To simplify the calculations, we have assumed until now that the fluid rotates as a whole inside the toroidal ring. More complex fluid motions, with different geometries, are obviously possible. An important case, which we will return to later, concerns the rotation around an axis still parallel to Oz, but located inside the fluid. The Gross-Pitaevskii wave function must then be zero along a line inside the fluid itself, which thus contains a singular line. This means that the phase may change by 2π as one rotates around this line. This situation corresponds to what is called a "vortex", a little swirl of fluid rotating around the singular line, called the "vortex core line". As the circulation of the velocity only depends on the phase change along the path going around the vortex core, the quantization relation (62) remains valid. Actually, from a historical point of view, the Gross-Pitaevskii equation was first introduced for the study of superfluidity and the quantization of the vortices circulation.

3-b. Repulsive potential barrier between states of different l

A classical rotating fluid will always come to rest after a certain time, due to the viscous dissipation at the walls. In such a process, the macroscopic rotational kinetic energy of the whole fluid is progressively degraded into numerous smaller scale excitations, which end up simply heating the fluid. Will a rotating quantum fluid of repulsive bosons, described by a wave function $\chi_l(\mathbf{r})$, behave in the same way? Will it successively evolve towards the state $\chi_{l-1}(\mathbf{r})$, then $\chi_{l-2}(\mathbf{r})$, etc., until it comes to rest in the state $\chi_0(\mathbf{r})$?

We have seen in § 4-c of Complement $\mathrm{C_{XV}}$ that, to avoid the energy cost of fragmentation, the system always remains in a state where all the particles occupy the same quantum state. This is why we can use the Gross-Pitaevskii equation (18).

α. *A simple geometry*

Let us first assume that the wave function $\chi(\mathbf{r},t)$ changes smoothly from $\chi_l(\mathbf{r})$ to $\chi_{l'}(\mathbf{r})$ according to:

$$\chi(\mathbf{r},t) = c_l(t)\,\chi_l(\mathbf{r}) + c_{l'}(t)\,\chi_{l'}(\mathbf{r}) \tag{63}$$

where the modulus of $c_l(t)$ decreases with time from 1 to 0, whereas $c_{l'}(t)$ does the opposite. Normalization imposes that at all times t:

$$|c_l(t)|^2 + |c_{l'}(t)|^2 = 1 \tag{64}$$

In such a state, let us show that the numerical density $n(r,\varphi,z;t)$ now depends on φ (this was not the case for either states l or l' separately). The transverse dependence of the density as a function of the variables r and z, is barely affected[3]. The variations of $n(r,\varphi,z;t)$ are given by:

$$\begin{aligned} n(r,\varphi,z;t) &= \left| c_l(t)\,u_l(r,z)e^{il\varphi} + c_{l'}(t)\,u_{l'}(r,z)e^{il'\varphi} \right|^2 \\ &= |c_l(t)|^2\,|u_l(r,z)|^2 + |c_{l'}(t)|^2\,|u_{l'}(r,z)|^2 \\ &\quad + c_l(t)\,c_{l'}^*(t)\,u_l(r,z)u_{l'}(r,z)e^{i(l-l')\varphi} + \text{c.c.} \end{aligned} \tag{65}$$

where c.c. stands for the complex conjugate of the preceding factor. The first two terms are independent of φ, and are just a weighted average of the densities associated with each of the states l and l'. The last term oscillates as a function of φ with an amplitude $|c_l(t)| \times |c_{l'}(t)|$, which is only zero if one of the two coefficients $c_l(t)$ or $c_{l'}(t)$ is zero. Calling φ_l the phase of the coefficient $c_l(t)$, this last term is proportional to:

$$c_l(t)\,c_{l'}^*(t)\,e^{i(l-l')\varphi} + \text{c.c.} = 2\,|c_l(t)| \times |c_{l'}(t)| \cos\left[(l-l')\,\varphi + \varphi_l - \varphi_{l'}\right] \tag{66}$$

Whatever the phases of the two coefficients $c_l(t)$ and $c_{l'}(t)$, the cosine will always oscillate between -1 and 1 as a function of φ. Adjusting those phases, one can deliberately change the value of φ for which the density is maximum (or minimum), but this will always occur somewhere on the circle. Superposing two states necessarily modulates the density.

Let us evaluate the consequences of this density modulation on the internal repulsive interaction energy of the fluid. As we did in relation (15), we use for the interaction

[3] or not at all, if we suppose the functions $u_l(r,z)$ and $u_{l'}(r,z)$ to be equal.

energy the zero range potential approximation, and insert it in expression (15) of Complement C_{XV}. Taking into account the normalization (17) of the wave function, we get:

$$\left\langle \widehat{W}_2 \right\rangle_\chi = \frac{g}{2} \int \mathrm{d}^3 r \ \left|\chi\left(\mathbf{r},t\right)\right|^4 = \frac{g}{2} \int_0^\infty \mathrm{d}r \int_0^{2\pi} \mathrm{d}\varphi \int_{-\infty}^{+\infty} \mathrm{d}z \ \left[n(r,\varphi,z;t)\right]^2 \tag{67}$$

We must now include the square of (65) in this expression, which will yield several terms. The first one, in $\left|c_l\left(t\right)\right|^4$, leads to the contribution:

$$\left|c_l\left(t\right)\right|^4 \left\langle \widehat{W}_2 \right\rangle_l \tag{68}$$

where $\left\langle \widehat{W}_2 \right\rangle_l$ is the interaction energy for the state $\chi_l(\mathbf{r})$. The second contribution is the similar term for the state l', and the third one, a cross term in $2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2$. Assuming, to keep things simple, that the densities associated with the states l and l' are practically the same, the sum of these three terms is just:

$$\left[\left|c_l\left(t\right)\right|^2 + \left|c_{l'}\left(t\right)\right|^2\right]^2 \left\langle \widehat{W}_2 \right\rangle_l = \left\langle \widehat{W}_2 \right\rangle_l \tag{69}$$

Up to now, the superposition has had no effect on the repulsive internal interaction energy. As for the cross terms between the terms independent of φ in (65) and the terms in $e^{\pm i\left(l-l'\right)\varphi}$, they will cancel out when integrated over φ. We are then left with the cross terms in $e^{\pm i\left(l-l'\right)\varphi} \times e^{\mp i\left(l-l'\right)\varphi}$, whose integral over φ yields:

$$2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2 \left|u_l(r,z)\right|^2 \left|u_{l'}(r,z)\right|^2 \tag{70}$$

Assuming as before that the densities associated with the states l and l' are practically the same, we obtain, after integration over r and z:

$$2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2 \left\langle \widehat{W}_2 \right\rangle_l \tag{71}$$

Adding (69), we finally obtain:

$$\left\langle \widehat{W}_2 \right\rangle_\chi = \left[1 + 2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2\right] \left\langle \widehat{W}_2 \right\rangle_l \tag{72}$$

We have shown that the density modulation associated with the superposition of states always increases the internal repulsion energy: this modulation does lower the energy in the low density region, but the increase in the high energy region outweighs the decrease (since the repulsive energy is a quadratic function of the density). The internal energy therefore varies between $\left\langle \widehat{W}_2 \right\rangle_l$ and the maximum $(3/2)\left\langle \widehat{W}_2 \right\rangle_l$, reached when the moduli of $c_l\left(t\right)$ and $c_{l'}\left(t\right)$ are both equal to $1/\sqrt{2}$.

β. Other geometries, different relaxation channels

There are many other ways for the Gross-Pitaevskii wave function to go from one rotational state to another. We have limited ourselves to the simplest geometry to introduce the concept of energy barriers with minimal mathematics. The fluid could

transit, however, through more complex geometries, such as the frequently observed creation of a vortex on the wall, the little swirl we briefly talked about at the end of § 3-a. A vortex introduces a 2π phase shift around a singular line along which the wave function is zero. Once the vortex is created, and contrary to what was the case in (62), the velocity circulation along a loop going around the torus is no longer independent of its path: it will change by $2\pi\hbar/m$ depending on whether the vortex is included in the loop or not. Furthermore, as the vortex moves in the fluid from one wall to another, it can be shown that the proportion of fluid conserving the initial circulation decreases while the proportion having a circulation where the quantum number l differs by one unit increases. Consequently, this vortex motion changes progressively the rotational angular momentum. Once the vortex has vanished on the other wall, the final result is a decrease by one unit of the quantum number l associated with the fluid rotation.

The continuous passage of vortices from one wall to another therefore yields another mechanism that allows the angular moment of the fluid to decrease. The creation of a vortex, however, is necessarily accompanied by a non-uniform fluid density, described by the Gross-Pitaevskii equation (this density must be zero along the vortex core). As we have seen above, this leads to an increase in the average repulsive energy between the particles (the fluid elastic energy). This process thus also encounters an energy barrier (discussed in more detail in the conclusion). In other words, the creation and motion of vortices provide another "relaxation channel" for the fluid velocity, with its own energy barrier, and associated relaxation time.

Many other geometries can be imagined for changing the fluid flow. Each of them is associated with a potential barrier, and therefore a certain lifetime. The relaxation channel with the shortest lifetime will mainly determine the damping of the fluid velocity, which may take, in certain cases, an extraordinarily long time (dozens of years or more), hence the name of "superfluid".

3-c. Critical velocity, metastable flow

For the sake of simplicity, we will use in our discussion the simple geometry of § 3-a. The transposition to other geometries involving, for example, the creation of vortices in the fluid would be straightforward. The main change would concern the height of the energy barrier[4].

With this simple geometry, the potential to be used in (60) is the sum of a repulsive potential $g\left|u_l(r,z)\right|^2$ and a kinetic energy of rotation around Oz, equal to $l^2\hbar^2/2mr^2$. We now show that, in a given l state, these two contributions can be expressed as a function of two velocities. First, relation (61) yields the rotation velocity v_l associated with state l:

$$v_l = \frac{1}{r}\frac{l\hbar}{m} \tag{73}$$

and the rotational energy is simply written as:

$$\langle E_{rot}\rangle_l = \frac{l^2\hbar^2}{2mr^2} = \frac{1}{2}m\left(v_l\right)^2 \tag{74}$$

[4]When several relaxation channels are present, the one associated with the lowest barrier mainly determines the time evolution.

As for the interaction term (term in g on the left-hand side), we can express it in a more convenient way, defining as before the numerical density n_0:

$$n_0 = |u_l(r,z)|^2 \tag{75}$$

and using the definition (39) for the sound velocity c. It can then be written in a form similar to (74):

$$gn_0 = mc^2 \tag{76}$$

The two velocities v_l and c allow an easy comparison of the respective importance of the kinetic and potential energies in a state l.

We now compare the contributions of these two terms either for states with a given l, or for a superposition of states (63). To clarify the discussion and be able to draw a figure, we will use a continuous variable defined as the average $\langle J_z \rangle$ of the component along Oz of the angular momentum:

$$\langle J_z \rangle = l\hbar \ |c_l(t)|^2 + l'\hbar \ |c_{l'}(t)|^2 \tag{77}$$

This expression varies continuously between $l\hbar$ and $l'\hbar$ when the relative weights of $|c_l(t)|^2$ and $|c_{l'}(t)|^2$ are changed while imposing relation (64); the continuous variable:

$$x = \langle J_z \rangle / \hbar \tag{78}$$

allows making interpolations between the discrete integer values of l.

Using the normalization relation (64) of the wave function (63), we can express x as a function of $|c_l(t)|^2$:

$$x = (l - l')\,|c_l(t)|^2 + l' \tag{79}$$

The variable x characterizes the modulus of each of the two components of the variational function (63). A second variable is needed to define the relative phase between these two components, which comes into play for example in (66). Instead of studying the time evolution of the fluid state vector inside this variational family, we shall simply give a qualitative argument, for several reasons. First of all, it is not easy to characterize precisely the coupling between the fluid and the environment by a Hamiltonian that can change the fluid rotational angular momentum (for example, the wall's irregularities may transfer energy and angular momentum from the fluid to the container). Furthermore, as the time-dependent Gross-Pitaevskii equation is nonlinear, its precise solutions are generally found numerically. This is why we shall only qualitatively discuss the effects of the potential barrier found in §3-b. The higher this barrier, the more difficult it is for x to go from l to l'. Let us evaluate the variation of the average energy as a function of x.

For integer values of x, relation (74) shows that the average rotational kinetic energy varies as the square of x; in between, its value can be found by interpolation as in (77). As for the potential energy, we saw that a continuous variation of $c_l(t)$ and $c_{l'}(t)$ necessarily involves a coherent superposition, which has an energy cost and increases the repulsive potential interaction. In particular, this interaction energy is multiplied by the factor 3/2 when the moduli of $c_l(t)$ and $c_{l'}(t)$ are equal (i.e. when x is an integer plus 1/2). As a result, to the quadratic variation of the rotational kinetic energy, we must

add an oscillating variation of the potential energy, minimum for all the integer values of x, and maximum half-way between. The oscillation amplitude is given by:

$$\frac{gn_0}{2} = \frac{1}{2}mc^2 \tag{80}$$

Figure 3 shows three plots of the variation of the system energy as a function of the average value $\langle J_z \rangle$. The lowest one, shown as a dotted line, corresponds to a superposition of the state l with the state $l' = l-1$, for a very small value of the coupling constant g (weak interactions, gas almost ideal). In this case and according to (39), the sound velocity is also very small and we are in the case $c \ll v_l$. Comparing (74) and (80) then shows that the potential energy contribution is negligible compared to the variation of the rotational kinetic energy between the two states. As a result, the modulation on this dotted line is barely perceptible, and this curve presents a single minimum at $x = 0$: whatever the initial rotational state, no potential barrier prevents the fluid rotational velocity from returning to zero (for example under the effect of the interactions with the irregularities of the walls containing the fluid).

The other two curves in Figure 3 correspond to a much larger value of g, hence, according to (39), to a much higher value of c. There are now several values of l for which v_l is small compared to c. The dashed line corresponds, as for the previous curve, to a superposition of the two states $l = 1$ and $l' = l-1$; the solid line (for the same value of g) to a superposition of $l = 3$ and $l' = 0$, corresponding to the case where the system goes directly from the state $l = 3$ to the rotational ground state in the torus, with $l' = 0$. It is obviously this last curve that presents the lowest energy barrier starting from $l = 3$ (shown with a circle in the figure). This is normal since this is the curve that involves the largest variation in the kinetic energy, in a sense opposite to that of the potential energy variation. It is thus the direct transition from $l = 3$ to $l' = 0$ that will determine the possibility for the system to relax towards a state of slower rotation. Let us again use (74) and (80) to compare the kinetic energy variation and the height of the repulsive potential barrier. All the states l, with velocities v_l much larger than c, have a kinetic energy much bigger than the maximum value of the potential energy: no energy barrier can be formed. On the other hand, all the states l with velocities v_l much smaller than c cannot lower their rotational state without going over a potential barrier.

In between these two extreme cases, there exists (for a given g) a "critical" value l_c corresponding to the onset of the barrier. It is associated with a "critical velocity" $v_c = l_c\hbar/mr$, of the order of the sound velocity c, fixing the maximum value of v_l for which this potential barrier exists. If the fluid rotational velocity in the torus is greater than v_c, the liquid can slow down its rotation without going over an energy barrier, and dissipation occurs as in an ordinary viscous liquid – the fluid is said to be "normal". If, however, the fluid velocity is less than the critical velocity, the physical system must necessarily go over a potential barrier (or more) to continuously tend towards $l = 0$. As this barrier results from the repulsion between all the particles and their neighbors, it has a macroscopic value. In principle, any barrier can be overcome, be it by thermal excitation, or by the quantum tunnel effect. However the time needed for this passage may take a gigantic value. First of all, it is extremely unlikely for a thermal fluctuation to reach a macroscopic energy value. As for the tunnel effect, its transition probability decreases exponentially with the barrier height and becomes extremely low for a macroscopic object. Consequently, the relaxation times of the fluid velocity may

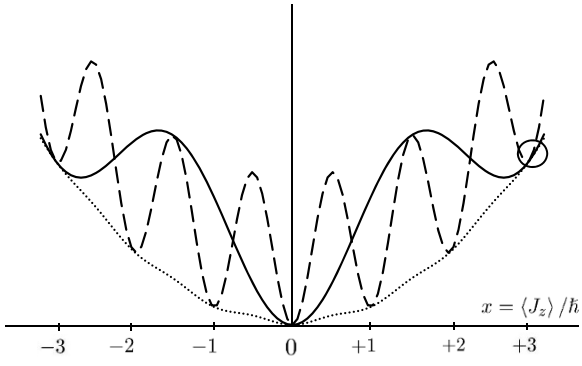


Figure 3: Plots of the energy of a rotating repulsive boson system, in a coherent superposition of the state l and the state l', as a function of its average angular momentum $\langle J_z \rangle$, expressed in units of $\hbar$. The lower dotted curve corresponds to the case where $l' = l - 1$ and the interaction constant g is small (almost ideal gas). The potential energy is then negligible and the total energy presents a single minimum in $\langle J_z \rangle = 0$. Consequently, whatever the initial rotational state of the fluid, it will relax to a motionless state $l = 0$ without having to go over any energy barrier, and its rotational kinetic energy will dissipate: it behaves as a normal fluid. The other two curves correspond to a much larger value of g – therefore, according to (39) to a much higher value of c. The dashed curve still corresponds to a superposition of the rotational states l and $l' = l - 1$, and the solid line to the direct superposition of the state $l = 3$ (shown with a circle in the figure) and the ground state $l' = 0$. The solid line curve presents the smallest barrier, hence determining the metastability of the current.
The higher the coupling constant g, the more l states presenting a minimum in the potential energy appear. They correspond to flow velocities in the torus that are smaller than the critical velocity. To go from the rotational state $l = 1$ to the motionless state $l = 0$, the system must go over a macroscopic energy barrier, which only occurs with a probability so small it can be considered equal to zero. The rotational current is therefore permanent, lasting for years, and the system is said to be superfluid. On the other hand, the states with higher values of l, for which the curve presents no minima, correspond to a normal fluid, whose rotation can slow down because of the viscosity (dissipation of the kinetic energy into heat).

become extraordinarily large, and, on the human scale, the rotation can be considered to last indefinitely. This phenomenon is called "superfluidity".

3-d. Generalization; topological aspects

Our argument remained qualitative for several reasons. To begin with, we showed the existence for the fluid of a critical velocity v_c, of the order of c, without giving its precise value. It would require a more detailed study of the potential curves such as the ones plotted in Figure 3, to obtain the precise values of the parameters for which the potential barrier appears or disappears. We also limited ourselves to simple geometries that could be described by a single variable φ, not taking into account other possible

deformations of the wave function. Various situations could occur, such as the creation of vortices or more complex processes, which would require a more elaborate mathematical treatment. In other words, we would have to take into account the existence of other relaxation channels for the moving fluid to come to rest, and look for the one leading to the lowest potential barrier, thereby determining the lifetime of the superfluid current.

There is, however, a more general way to address the problem, which shows that our basic conclusions are not limited to the particular case we have studied. It is based on the topological aspects of the wave function phase. When this phase varies by $2l\pi$ as we go around the torus, it expresses a topological property characterized by the winding number l, which is an integer and cannot vary continuously. This is why, as long as the phase is well defined everywhere – i.e. as long as the wave function does not go to zero – we cannot go continuously from l to $l \pm 1$. We already saw this in the particular example of the wave function (63): when the modulus of $c_l\,(t)$ varies in time from 1 to 0, while the modulus of $c_{l'}\,(t)$ does the opposite, we necessarily went through a situation where the wave function went to zero through interference, in a plane corresponding to a certain value of φ; but the phase of the wave function is undetermined in this plane, and as we cross it, the phase undergoes a discontinuous jump. Now the canceling of the wave function of a great number of condensed bosons means the density must also be zero at that point, hence larger in other points of space. This spatial density variation introduces an energy increase, due to the finite compressibility of the fluid (as we saw in § 3-b, the energy increase in the high density regions is larger than the energy decrease in low density regions). This means there is an energy barrier opposing the change in the number of turns l of the phase. The height of this barrier must now be compared with the kinetic energy variation. As seen above, there is a drastic change in the flow regime, depending on whether the fluid velocity is smaller or larger than a certain critical velocity v_c. In the first case, superfluidity allows a current to flow without dissipation, lasting practically indefinitely. In the second, no energy consideration opposes dissipation, and the rotation slows down progressively, as in an ordinary liquid.

The essential idea to remember is that superfluidity comes from the repulsive interactions, and for two reasons. First of all, they explain the presence of the energy barrier, responsible for the metastability. The second reason, even more essential, is that the repulsion between bosons constantly tends to put all the fluid particles in the same quantum state - see § 4-c of Complement C_{XV}; thanks to this property, we were able to characterize the intermediate rotational states by a very simple wave function (63). This implies that the quantum fluid can only occupy a very limited number of states, compared to a situation where the particles would be distinguishable. Consequently, it has a hard time dissipating its kinetic energy into heat, as a classical fluid would do, and it therefore maintains its rotation over such long times that a slowing down is practically impossible to observe.

Complement E_{XV}

Fermion system, Hartree-Fock approximation

1 Foundation of the method 1678
1-a Trial family and Hamiltonian 1678
1-b Energy average value 1680
1-c Optimization of the variational wave function 1682
1-d Equivalent formulation for the average energy stationarity . . 1684
1-e Variational energy 1685
1-f Hartree-Fock equations 1686
2 Generalization: operator method 1688
2-a Average energy 1689
2-b Optimization of the one-particle density operator 1692
2-c Mean field operator 1693
2-d Hartree-Fock equations for electrons 1695
2-e Discussion 1698

Introduction

Computing the energy levels of a system of N electrons, interacting with each other through the Coulomb force, and placed in an external potential $V_1(\mathbf{r})$ is a very important problem in physics and chemistry. It is encountered in the determination of the energy levels of atoms (in which case the external potential for the electrons[1] is the Coulomb potential created by the nucleus $-Zq^2/4\pi\varepsilon_0 r$), or of molecules as well, or of electrons in a solid (submitted to a periodic potential), or in an aggregate or a nanocristal, etc. It is a problem where two ingredients simultaneously play an essential role: the fermionic character of the electrons, which forbids them to occupy the same individual state, and the effects of their mutual interactions. Ignoring the Coulomb repulsion between electrons would make the calculation fairly simple, and similar to that of § 1 in Complement C_{XIV}, concerning free fermions in a box; the free plane wave individual states would have to be replaced by the energy eigenstates of a single particle placed in the potential $V_1(\mathbf{r})$. This would lead to a 3-dimensional Schrödinger equation, which can be solved with very good precision, although not necessarily analytically.

However, be it in atoms or in solids, the repulsion between electrons plays an essential role. Neglecting it would lead us to conclude, for example, that, as Z increases, the size of atoms decreases due to the attractive effect of the nucleus, whereas the opposite occurs[2]! For N interacting particles, even without taking the spin into account, an exact

[1] We assume the nucleus mass to be infinitely larger than the electron mass. The electronic system can then be studied assuming the nucleus fixed and placed at the origin.

[2] The Pauli exclusion principle is not sufficient to explain why an atom's size increases with its atomic number Z. One can evaluate the approximate size of a hypothetical atom with non-interacting electrons (we consider the atom's size to be given by the size of the outermost occupied orbit). The Bohr radius

computation would require solving a Schrödinger equation in a $3N$-dimensional space; this is clearly impossible when N becomes large, even with the most powerful computer. Hence, approximation methods are needed, and the most common one is the Hartree-Fock method, which reduces the problem to solving a series of 3-dimensional equations. It will be explained in this complement for fermionic particles.

The Hartree-Fock method is based on the variational approximation (Complement E_{XI}), where we choose a trial family of state vectors, and look for the one that minimizes the average energy. The chosen family is the set of all possible Fock states describing the system of N fermions. We will introduce and compute the "self-consistent" mean field in which each electron moves; this mean field takes into account the repulsion due to the other electrons, hence justifying the central field method discussed in Complement A_{XIV}. This method applies not only to the atom's ground state but also to all its stationary states. It can also be generalized to many other systems such as molecules, for example, or to the study of the ground level and excited states of nuclei, which are protons and neutrons in bound systems.

This complement presents the Hartree-Fock method in two steps, starting in § 1 with a simple approach in terms of wave functions, which is then generalized in § 2 by using Dirac notation and projector operators. The reader may choose to go through both steps or go directly to the second. In § 1, we deal with spinless particles, which allows discussing the basic physical ideas and introducing the mean field concept keeping the formalism simple. A more general point of view is exposed in § 2, to clarify a number of points and to introduce the concept of a one-particle (with or without spin) effective Hartree-Fock Hamiltonian. This Hamiltonian reduces the interactions with all the other particles to a mean field operator. More details on the Hartree-Fock methods, and in particular their relations with the Wick theorem, can be found in Chapters 7 and 8 of reference [5].

1. Foundation of the method

Let us first expose the foundation of the Hartree-Fock method in a simple case where the particles have no spin (or are all in the same individual spin state) so that no spin quantum number is needed to define their individual states, specified by their wave functions. We introduce the notation and define the trial family of the N-particle state vectors.

1-a. Trial family and Hamiltonian

We choose as the trial family for the state of the N-fermion system all the states that can be written as:

$$\widetilde{|\Psi\rangle} = a^\dagger_{\theta_1} a^\dagger_{\theta_2} ... \; a^\dagger_{\theta_N} |0\rangle \tag{1}$$

where $a^\dagger_{\theta_1}$, $a^\dagger_{\theta_2}$,..., $a^\dagger_{\theta_N}$ are the creation operators associated with a set of normalized individual states $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$, all orthogonal to each other (and hence distinct). The state $\widetilde{|\Psi\rangle}$ is therefore normalized to 1. This set of individual states is, at the moment, arbitrary; it will be determined by the following variational calculation.

a_0 varies as $1/Z$, whereas the highest value of the principal quantum number n of the occupied states varies approximately as $Z^{1/3}$. The size $n^2 a_0$ we are looking for varies approximately as $Z^{-1/3}$.

For spinless particles, the corresponding wave function $\widetilde{\Psi}(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_N)$ can be written in the form of a Slater determinant (Chapter XIV, § C-3-c-β):

$$\widetilde{\Psi}(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \theta_1(\mathbf{r}_1) & \theta_2(\mathbf{r}_1) & .. & \theta_N(\mathbf{r}_1) \\ \theta_1(\mathbf{r}_2) & \theta_2(\mathbf{r}_2) & .. & \theta_N(\mathbf{r}_2) \\ .. & .. & .. & .. \\ \theta_1(\mathbf{r}_N) & \theta_2(\mathbf{r}_N) & .. & \theta_N(\mathbf{r}_N) \end{vmatrix} \tag{2}$$

The system Hamiltonian is the sum of the kinetic energy, the one-body potential energy and the interaction energy:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\text{ext}} + \widehat{W}_{\text{int}} \tag{3}$$

The first term, $\widehat{H}_0$, is the operator associated with the fermion kinetic energy, sum of the individual kinetic energies:

$$\widehat{H}_0 = \sum_{q=1}^{N} \frac{(\mathbf{P}_q)^2}{2m} \tag{4}$$

where m is the particle mass and $\mathbf{P}_q$, the momentum operator of particle q. The second term, $\widehat{V}_{\text{ext}}$, is the operator associated with their energy in an applied external potential $V_{1.}$:

$$\widehat{V}_{\text{ext}} = \sum_{q=1}^{N} V_1(\mathbf{R}_q) \tag{5}$$

where $\mathbf{R}_q$ is the position operator of particle q. For electrons with charge q_e placed in the attractive Coulomb potential of a nucleus of charge $-Zq_e$ positioned at the origin (Z is the nucleus atomic number), this potential is attractive and equal to:

$$V_1(\mathbf{r}) = -\frac{Zq_e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r}|} \tag{6}$$

where ε_0 is the vacuum permittivity. Finally, the term $\widehat{W}_{\text{int}}$ corresponds to their mutual interaction energy:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{q \neq q'} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{7}$$

For electrons, the function W_2 is given by the Coulomb repulsive interaction:

$$W_2(\mathbf{r}, \mathbf{r}') = \frac{q_e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \tag{8}$$

The expressions given above are just examples; as mentioned earlier, the Hartree-Fock method is not limited to the computation of the electronic energy levels in an atom.

1-b. Energy average value

Since state (1) is normalized, the average energy in this state is given by:

$$\widetilde{E} = \widetilde{\langle\Psi|}\widehat{H}\widetilde{|\Psi\rangle} \tag{9}$$

Let us evaluate successively the contributions of the three terms of (3), to obtain an expression which we will eventually vary.

α. Kinetic energy

Let us introduce a complete orthonormal basis $\{|\theta_s\rangle\}$ of the one-particle state space by adding to the set of states $|\theta_i\rangle$ ($i = 1, 2, ..., N$) other orthonormal states; the subscript s now ranges from 1 to D, dimension of this space (D may be infinite). We can then expand $\widehat{H}_0$ as in relation (B-12) of Chapter XV:

$$\widehat{H}_0 = \sum_{r,s} \langle\theta_r| \frac{\mathbf{P}^2}{2m} |\theta_s\rangle \; a^\dagger_{\theta_r} a_{\theta_s} \tag{10}$$

where the two summations over r and s range from 1 to D. The average value in $\widetilde{|\Psi\rangle}$ of the kinetic energy can then be written:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{r,s} \langle\theta_r| \frac{\mathbf{P}^2}{2m} |\theta_s\rangle \;\; \langle 0| \, a_{\theta_N} ... \; a_{\theta_2} a_{\theta_1} \left(a^\dagger_{\theta_r} a_{\theta_s} \right) a^\dagger_{\theta_1} a^\dagger_{\theta_2} ... \; a^\dagger_{\theta_N} |0\rangle \tag{11}$$

which contains the scalar product of the ket:

$$\left(a_{\theta_s} \right) \; a^\dagger_{\theta_1} a^\dagger_{\theta_2} ... \; a^\dagger_{\theta_N} |0\rangle = \left(a_{\theta_s} \right) \;\; |\theta_1, \theta_2, \; ... \; , \theta_N\rangle \tag{12}$$

by the bra:

$$\langle 0| \, a_{\theta_N} ... \; a_{\theta_2} a_{\theta_1} \;\; \left(a^\dagger_{\theta_r} \right) = \langle \theta_1, \theta_2, \; ... \; , \theta_N| \;\; \left(a^\dagger_{\theta_r} \right) \tag{13}$$

Note however that in the ket, the action of the annihilation operator a_{θ_s} yields zero unless it acts on a ket where the individual state is already occupied; consequently, the result will be different from zero only if the state $|\theta_s\rangle$ is included in the list of the N states $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$. Taking the Hermitian conjugate of (13), we see that the same must be true for the state $|\theta_r\rangle$, which must be included in the same list. Furthermore, if $r \neq s$ the resulting kets have different occupation numbers, and are thus orthogonal. The scalar product will therefore only differ from zero if $r = s$, in which case it is simply equal to 1. This can be shown by moving to the front the state $|\theta_r\rangle$ both in the bra and in the ket; this will require two transpositions with two sign changes which cancel out, or none if the state $|\theta_r\rangle$ was already in the front. Once the operators have acted, the bra and the ket correspond to exactly the same occupied states and their scalar product is 1. We finally get:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{i=1}^{N} \langle\theta_i| \frac{\mathbf{P}^2}{2m} |\theta_i\rangle \tag{14}$$

Consequently, the average value of the kinetic energy is simply the sum of the average kinetic energy in each of the occupied states $|\theta_i\rangle$.

For spinless particles, the kinetic energy operator is actually a differential operator $-\hbar^2\Delta/2m$ acting on the individual wave functions. We therefore get:

$$\left\langle \widehat{H}_0 \right\rangle = -\frac{\hbar^2}{2m} \int \mathrm{d}^3 r \ \sum_{i=1}^{N} \theta_i^*(\mathbf{r}) \ \Delta\theta_i(\mathbf{r}) \tag{15}$$

β. Potential energy

As the potential energy $\widehat{V}_1$ is also a one-particle operator, its average value can be computed in a similar way. We obtain:

$$\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle = \sum_{i=1}^{N} \langle\theta_i| \, V_1(\mathbf{R}) \, |\theta_i\rangle \tag{16}$$

that is, for spinless particles:

$$\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle = \int \mathrm{d}^3 r \ V_1(\mathbf{r}) \sum_{i=1}^{N} \ |\theta_i(\mathbf{r})|^2 \tag{17}$$

As before, the result is simply the sum of the average values associated with the individual occupied states.

γ. Interaction energy

The average value of the interaction energy $\widehat{W}_2$ in the state $\widetilde{|\Psi\rangle}$ has already been computed in § C-5 of Chapter XV. We just have to replace, in the relations (C-28) as well as (C-32) to (C-34) of that chapter, the n_i by 1 for all the occupied states $|\theta_i\rangle$, by zero for the others, and to rename the wave functions $u_i(\mathbf{r})$ as $\theta_i(\mathbf{r})$. We then get:

$$\begin{aligned} \left\langle \widehat{W}_{\mathrm{int}} \right\rangle = \widetilde{\langle\Psi|} \widehat{W}_{\mathrm{int}} \widetilde{|\Psi\rangle} = \frac{1}{2} \ \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \ W_2(\mathbf{r}, \mathbf{r}') \\ \times \sum_{i,j=1}^{N} \left[|\theta_i(\mathbf{r})|^2 \, |\theta_j(\mathbf{r}')|^2 - \theta_i^*(\mathbf{r}) \theta_j^*(\mathbf{r}') \theta_i(\mathbf{r}') \theta_j(\mathbf{r}) \right] \end{aligned} \tag{18}$$

We have left out the condition $i \neq j$, no longer useful since the $i = j$ terms are zero. The second line of this equation contains the sum of the direct and the exchange terms.

The result can be written in a more concise way by introducing the projector P_N over the subspace spanned by the N kets $|\theta_i\rangle$:

$$P_N = \sum_{i=1}^{N} |\theta_i\rangle \, \langle\theta_i| \tag{19}$$

Its matrix elements are:

$$\langle\mathbf{r}| \, P_N \, |\mathbf{r}'\rangle = \sum_{i=1}^{N} \theta_i(\mathbf{r}) \theta_i^*(\mathbf{r}') \tag{20}$$

This leads to:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \left[\langle \mathbf{r}| \, P_N \, |\mathbf{r}\rangle \, \langle \mathbf{r}'| \, P_N \, |\mathbf{r}'\rangle - \langle \mathbf{r}| \, P_N \, |\mathbf{r}'\rangle \, \langle \mathbf{r}'| \, P_N \, |\mathbf{r}\rangle \right] \tag{21}$$

Comment:

The matrix elements of P_N are actually equal to the spatial non-diagonal correlation function $G_1(\mathbf{r},\mathbf{r}')$, which will be defined in Chapter XVI (§ B-3-a). This correlation function can be expressed as the average value of the product of field operators $\Psi(\mathbf{r})$:

$$G_1(\mathbf{r},\mathbf{r}') = \left\langle \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r}') \right\rangle \tag{22}$$

For a system of N fermions in the states $|\theta_1\rangle$, $|\theta_2\rangle$, .., $|\theta_N\rangle$, we can write:

$$\begin{array}{l} G_1(\mathbf{r},\mathbf{r}') = \langle \theta_1, \theta_2, .., \; \theta_N| \, \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r}') \, |\theta_1, \theta_2, .., \; \theta_N\rangle \\ = \sum_{i,j} \langle \theta_1, \theta_2, .., \; \theta_N| \left[\theta_i^*(\mathbf{r}) a_i^\dagger \right] \left[\theta_j(\mathbf{r}') a_j \right] |\theta_1, \theta_2, .., \; \theta_N\rangle = \sum_{i=1}^N \theta_i^*(\mathbf{r}) \theta_i(\mathbf{r}') \end{array} \tag{23}$$

Inserting this relation in (18) we get:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \left[G_1(\mathbf{r},\mathbf{r}) G_1(\mathbf{r}',\mathbf{r}') - G_1(\mathbf{r},\mathbf{r}') G_1(\mathbf{r}',\mathbf{r}) \right] \tag{24}$$

Comparison with relation (C-28) of Chapter XV, which gives the same average value, shows that the right-hand side bracket contains the two-particle correlation function $G_2(\mathbf{r},\mathbf{r}')$. For a Fock state, this function can therefore be simply expressed as two products of one-particle correlation functions at two points:

$$G_2(\mathbf{r},\mathbf{r}') = G_1(\mathbf{r},\mathbf{r}) G_1(\mathbf{r}',\mathbf{r}') - G_1(\mathbf{r},\mathbf{r}') G_1(\mathbf{r}',\mathbf{r}) \tag{25}$$

1-c. Optimization of the variational wave function

We now vary $\widetilde{|\Psi\rangle}$ to determine the conditions leading to a stationary value of the total energy $\widetilde{E}$:

$$\widetilde{E} = \left\langle \widehat{H}_0 \right\rangle + \left\langle \widehat{V}_{\text{ext}} \right\rangle + \left\langle \widehat{W}_{\text{int}} \right\rangle \tag{26}$$

where the three terms in this summation are given by (15), (16) and (18). Let us vary one of the kets $|\theta_k\rangle$, k being arbitrarily chosen between 1 and N:

$$|\theta_k\rangle \Rightarrow |\theta_k\rangle + |\delta\theta_k\rangle \tag{27}$$

or, in terms of an individual wave function:

$$\theta_k(\mathbf{r}) \Rightarrow \theta_k(\mathbf{r}) + \delta\theta_k(\mathbf{r}) \tag{28}$$

This will yield the following variations:

$$\delta \left\langle \widehat{H}_0 \right\rangle = \frac{-\hbar^2}{2m} \int \mathrm{d}^3 r \; \left[\delta\theta_k^*(\mathbf{r}) \; \Delta\theta_k(\mathbf{r}) + \theta_k^*(\mathbf{r}) \; \Delta\delta\theta_k(\mathbf{r}) \right] \tag{29}$$

and:

$$\delta \left\langle \widehat{V}_{\text{ext}} \right\rangle = \int \mathrm{d}^3 r \; V_1(\mathbf{r}) \left[\delta\theta_k^*(\mathbf{r}) \; \theta_k(\mathbf{r}) + \theta_k^*(\mathbf{r}) \; \delta\theta_k(\mathbf{r}) \right] \tag{30}$$

As for the variation of $\left\langle \widehat{W}_{\text{int}} \right\rangle$, we must take from (18) two contributions: the first one from the terms $i = k$, and the other from the terms $j = k$. These contributions are actually equal as they only differ by the choice of a dummy subscript. The factor 1/2 disappears and we get:

$$\begin{aligned} \delta \left\langle \widehat{W}_{\text{int}} \right\rangle = \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{j=1}^{N} \Big[\Big[\delta\theta_k^*(\mathbf{r}) \theta_k(\mathbf{r}) + \theta_k^*(\mathbf{r}) \delta\theta_k(\mathbf{r}) \Big] \left| \theta_j(\mathbf{r}') \right|^2 \\ - \delta\theta_k^*(\mathbf{r}) \theta_j^*(\mathbf{r}') \theta_k(\mathbf{r}') \theta_j(\mathbf{r}) - \theta_k^*(\mathbf{r}) \theta_j^*(\mathbf{r}') \delta\theta_k(\mathbf{r}') \theta_j(\mathbf{r}) \Big] \end{aligned} \tag{31}$$

The variation of $\widetilde{E}$ is simply the sum of (29), (30) and (31).

We now consider variations $\delta\theta_k$, which can be written as:

$$\delta\theta_k(\mathbf{r}) = \delta\varepsilon \; e^{i\chi} \theta_l(\mathbf{r}) \qquad \text{with } l > N \tag{32}$$

(where $\delta\varepsilon$ is a first order infinitely small parameter). These variations are proportional to the wave function of one of the non-occupied states, which was added to the occupied states to form a complete orthonormal basis; the phase χ is an arbitrary parameter. Such a variation does not change, to first order, either the norm of $|\theta_k\rangle$, or its scalar product with all the occupied states $l \leq N$; it therefore leaves unchanged our assumption that the occupied states basis is orthonormal. The first order variation of the energy $\delta\widetilde{E}$ is obtained by inserting $\delta\theta_k$ and its complex conjugate $\delta\theta_k^*$ into (29), (30) and (31); we then get terms in $e^{i\chi}$ in the first case, and terms in $e^{-i\chi}$ in the second. For $\widetilde{E}$ to be stationary, its variation must be zero to first order for any value of χ; now the sum of a term in $e^{i\chi}$ and another in $e^{-i\chi}$ will be zero for any value of χ only if both terms are zero. It follows that we can impose $\delta\widetilde{E}$ to be zero (stationary condition) considering the variations of $\delta\theta_k$ and $\delta\theta_k^*$ to be independent. Keeping only the terms in $\delta\theta_k^*$, we obtain the stationary condition of the variational energy:

$$\begin{aligned} \int \mathrm{d}^3 r \; \delta\theta_k^*(\mathbf{r}) \; \Bigg\{ & \frac{-\hbar^2}{2m} \Delta\theta_k(\mathbf{r}) + V_1(\mathbf{r}) \theta_k(\mathbf{r}) + \\ & + \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{j=1}^{N} \Big[\theta_k(\mathbf{r}) \left| \theta_j(\mathbf{r}') \right|^2 - \theta_j^*(\mathbf{r}') \theta_k(\mathbf{r}') \theta_j(\mathbf{r}) \Big] \Bigg\} = 0 \end{aligned} \tag{33}$$

or, taking (20) into account:

$$\begin{aligned} \int \mathrm{d}^3 r \; \delta\theta_k^*(\mathbf{r}) \; \Bigg\{ & \frac{-\hbar^2}{2m} \Delta\theta_k(\mathbf{r}) + V_1(\mathbf{r}) \theta_k(\mathbf{r}) + \\ & + \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \left[\langle \mathbf{r}' | \, P_N \, | \mathbf{r}' \rangle \, \theta_k(\mathbf{r}) - \langle \mathbf{r} | \, P_N \, | \mathbf{r}' \rangle \, \theta_k(\mathbf{r}') \right] \Bigg\} = 0 \end{aligned} \tag{34}$$

This relation can also be written as::

$$\int \mathrm{d}^3 r \; \delta\theta_k^*(\mathbf{r}) \; \mathcal{D} \left[\theta_k(\mathbf{r}) \right] = 0 \tag{35}$$

where the integro-differential operator $\mathcal{D}$ is defined by its action on an arbitrary function $\theta(\mathbf{r})$:

$$\mathcal{D}\left[\theta(\mathbf{r})\right] = \left\{ \frac{-\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \left\langle \mathbf{r}' \right| P_N \left| \mathbf{r}' \right\rangle \right\} \theta(\mathbf{r}) - \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \; \left\langle \mathbf{r} \right| P_N \left| \mathbf{r}' \right\rangle \theta(\mathbf{r}') \tag{36}$$

This operator depends on the diagonal $\left\langle \mathbf{r}' \right| P_N \left| \mathbf{r}' \right\rangle$ and non-diagonal $\left\langle \mathbf{r} \right| P_N \left| \mathbf{r}' \right\rangle$ spatial correlation functions associated with the set of states occupied by the N fermions.

Relation (35) thus shows that the action of the differential operator $\mathcal{D}$ on the function $\theta_k(\mathbf{r})$ yields a function orthogonal to all the functions $\theta_l(\mathbf{r})$ for $l > N$. This means that the function $\mathcal{D}\left[\theta_k(\mathbf{r})\right]$ only has components on the wave functions of the occupied states: it is a linear combination of these functions. Consequently, for the energy $\widetilde{E}$ to be stationary there is a simple condition: the invariance under the action of the integro-differential operator $\mathcal{D}$ of the N-dimensional vector space $\mathscr{F}_N$, spanned by all the linear combinations of the functions $\theta_i(\mathbf{r})$ with $i = 1, 2, ..N$.

Comment:

One could wonder why we limited ourselves to the variations $\delta\theta_k$ written in (32), proportional to non-occupied individual states. The reason will become clearer in § 2, where we use a more general method that shows directly which variations of each individual states are really useful to consider (see in particular the discussion at the end of § 2-a). For now, it can be noted that choosing a variation $\delta\theta_k$ proportional to the same wave function $\theta_k(\mathbf{r})$ would simply change its norm or phase, and therefore have no impact on the associated quantum state (in addition, a change of norm would not be compatible with our hypotheses, as in the computation of the average values we always assumed the individual states to remain normalized). If the state does not change, the energy $\widetilde{E}$ must remain constant and writing a stationary condition is pointless. Similarly, to give $\theta_k(\mathbf{r})$ a variation proportional to another occupied wave function $\theta_l(\mathbf{r})$ (where l is included between 1 and N) is just as useless, as we now show. In this operation, the creation operator $a_k^\dagger$ acquires a component on $a_l^\dagger$ (Chapter XV, § A-6), but the state vector expression (1) remains unchanged. The state vector thus acquire a component including the square of a creation operator, which is zero for fermions. Consequently, the stationarity of the energy is automatically ensured in this case.

1-d. Equivalent formulation for the average energy stationarity

Operator $\mathcal{D}$ can be diagonalized in the subspace $\mathscr{F}_N$, as can be shown[3] from its definition (36) – a more direct demonstration will be given in § 2. We call $\varphi_n(\mathbf{r})$ its

[3] As any Hermitian operator can be diagonalized, we simply show that (36) leads to matrix elements obeying the Hermitian conjugation relation. Let us verify that the two integrals $\int d^3r \; \theta_1^*(\mathbf{r}) \; \mathcal{D}\left[\theta_2(\mathbf{r})\right]$ and $\int d^3r \; \theta_2^*(\mathbf{r}) \; \mathcal{D}\left[\theta_1(\mathbf{r})\right]$ are complex conjugates of each other. For the contributions to these matrix elements of the kinetic and potential (in V_1) energy, we simply find the usual relations insuring the corresponding operators are Hermitian. As for the interaction term, the complex conjugation is obvious for the direct term; for the exchange term, a simple inversion of the integral variables d^3r and d^3r', plus the fact that $W_2(\mathbf{r},\mathbf{r}')$ is equal to $W_2(\mathbf{r}',\mathbf{r})$ allows verifying the conjugation.

eigenfunctions. These functions $\varphi_n(\mathbf{r})$ are linear combinations of the $\theta_j(\mathbf{r})$ corresponding to the states appearing in the trial ket (1), and therefore lead to the same N-particle state, because of the antisymmetrization[4]. The basis change from the $\theta_i(\mathbf{r})$ to the $\varphi_n(\mathbf{r})$ has no effect on the projector P_N onto to the subspace $\mathcal{F}_N$, whose matrix elements appearing in (36) can be expressed in a way similar to those in (20):

$$\langle \mathbf{r}| \, P_N \, |\mathbf{r}'\rangle = \sum_{n=1}^{N} \varphi_n(\mathbf{r}) \varphi_n^*(\mathbf{r}') \tag{37}$$

Consequently, the eigenfunctions of the operator $\mathcal{D}$ obey the equations:

$$\boxed{\begin{aligned} &\left[\frac{-\hbar^2}{2m} \Delta + V_1(\mathbf{r}) + \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{p=1}^{N} |\varphi_p(\mathbf{r}')|^2 \right] \varphi_n(\mathbf{r}) \\ &\qquad - \left[\int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{p=1}^{N} \varphi_p^*(\mathbf{r}') \varphi_p(\mathbf{r}) \varphi_n(\mathbf{r}') \right] = \widetilde{e}_n \varphi_n(\mathbf{r}) \end{aligned}} \tag{38}$$

where $\widetilde{e}_n$ are the associated eigenvalues. These relations are called the "Hartree-Fock equations".

For the average total energy associated with a state such as (1) to be stationary, it is therefore necessary for this state to be built from N individual states whose orthogonal wave functions φ_1, φ_2, .. , φ_N are solutions of the Hartree-Fock equations (38) with $n = 1$, 2, .. , N. Conversely, this condition is sufficient since, replacing the $\theta_k(\mathbf{r})$ by solutions $\varphi_n(\mathbf{r})$ of the Hartree-Fock equations in the energy variation (34) yields the result:

$$\widetilde{e}_n \int \mathrm{d}^3 r \; \delta\varphi_k^*(\mathbf{r}) \; \varphi_n(\mathbf{r}) \tag{39}$$

which is zero for all $\delta\varphi_k(\mathbf{r})$ variations, since, according to (32), they must be orthogonal to the N solutions $\varphi_n(\mathbf{r})$. Conditions (38) are thus equivalent to energy stationarity.

1-e. Variational energy

Assume we found a series of solutions for the Hartree-Fock equations, i.e. a set of N eigenfunctions $\varphi_n(\mathbf{r})$ with the associated eigenvalues $\widetilde{e}_i$. We still have to compute the minimal variational energy of the N-particle system. This energy is given by the sum (26) of the three terms of kinetic, potential and interaction energies obtained by replacing in (15), (16) and (18) the $\theta_i(\mathbf{r})$ by the eigenfunctions $\varphi_n(\mathbf{r})$:

$$\widetilde{E}_{HF} = \left\langle \widehat{H}_0 \right\rangle_{HF} + \left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF} + \left\langle \widehat{W}_{\mathrm{int}} \right\rangle_{HF} \tag{40}$$

[4] A determinant value does not change if one adds to one of its column a linear combination of the others. Hence we can add to the first column of the Slater determinant (2) the linear combination of the $\theta_2(\mathbf{r})$, $\theta_3(\mathbf{r})$, ... that makes it proportional to $\varphi_1(\mathbf{r})$. One can then add to the second column the combination that makes it proportional to $\varphi_2(\mathbf{r})$, etc. Step by step, we end up with a new expression for the original wave function $\widetilde{\Psi}(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N)$, which now involves the Slater determinant of the $\varphi_i(\mathbf{r})$. It is thus proportional to this determinant. A demonstration of the strict equality (within a phase factor) will be given in § 2.

(the subscripts HF indicate we are dealing with the average energies after the Hartree-Fock optimization, which minimizes the variational energy). Intuitively, one could expect this total energy to be simply the sum of the energies $\widetilde{e}_n$, but, as we are going to show, this is not the case. Multiplying the left-hand side of equation (38) by $\varphi_n^*(\mathbf{r})$ and after integration over d^3r, we get:

$$\widetilde{e}_n = \int \mathrm{d}^3r\ \varphi_n^*(\mathbf{r}) \left\{ \left[\frac{-\hbar^2}{2m}\Delta + V_1(\mathbf{r}) \right] \varphi_n(\mathbf{r}) + \int \mathrm{d}^3r'\ W_2(\mathbf{r},\mathbf{r}') \sum_{p=1}^{N} \left[\left|\varphi_p(\mathbf{r}')\right|^2 \varphi_n(\mathbf{r}) - \varphi_p^*(\mathbf{r}')\varphi_n(\mathbf{r}')\varphi_p(\mathbf{r}) \right] \right\} \tag{41}$$

We then take a summation over the subscript n, and use (15), (16) and (18), the θ being replaced by the φ:

$$\sum_{n=1}^{N} \widetilde{e}_n = \left\langle \widehat{H}_0 \right\rangle_{HF} + \left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF} + 2 \left\langle \widehat{W}_{\mathrm{int}} \right\rangle_{HF} \tag{42}$$

This expression does not yield the stationary value of the total energy, but rather a sum where the particle interaction energy is counted twice. From a physical point of view, it is clear that if each particle energy is computed taking into account its interaction with all the others, and if we then add all these energies, we get an expression that includes twice the interaction energy associated with each pair of particles.

The sum of the $\widetilde{e}_n$ does contain, however, useful information that enables us to avoid computing the interaction energy contribution to the variational energy. Eliminating $\left\langle \widehat{W}_{\mathrm{int}} \right\rangle$ between (40) and (42), we get:

$$\widetilde{E}_{HF} = \frac{1}{2} \left[\sum_{n=1}^{N} \widetilde{e}_n + \left\langle \widehat{H}_0 \right\rangle_{HF} + \left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF} \right] \tag{43}$$

where the interaction energy is no longer present. One can then compute $\left\langle \widehat{H}_0 \right\rangle_{HF}$ and $\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF}$ using the solutions of the Hartree-Fock equations (38), without worrying about the interaction energy. Using (15) and (17) in this relation, we can write the total energy $\widetilde{E}$ as:

$$\widetilde{E}_{HF} = \frac{1}{2} \left[\sum_{n=1}^{N} \widetilde{e}_n + \sum_{n=1}^{N} \int \mathrm{d}^3r'\ \left|\varphi_n(\mathbf{r}')\right|^2 V_1(\mathbf{r}') - \frac{\hbar^2}{2m} \sum_{n=1}^{N} \int \mathrm{d}^3r' \varphi_n^*(\mathbf{r}')\ \Delta\varphi_n(\mathbf{r}') \right] \tag{44}$$

The total energy is thus half the sum of the $\widetilde{e}_i$, of the average kinetic energy, and finally of the one-body average potential energy.

1-f. Hartree-Fock equations

Equation (38) may be written as:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + V^{\mathrm{dir}}(\mathbf{r}) \right] \varphi_n(\mathbf{r}) - \int \mathrm{d}^3r'\ V^{\mathrm{ex}}(\mathbf{r},\mathbf{r}')\varphi_n(\mathbf{r}') = \widetilde{e}_n\ \varphi_n(\mathbf{r})} \tag{45}$$

where the direct $V^{\rm dir}(\mathbf{r})$ and exchange $V^{\rm ex}(\mathbf{r},\mathbf{r}')$ potentials are defined as:

$$\boxed{\begin{aligned} V^{\rm dir}(\mathbf{r}) &= \sum_{p=1}^{N} \int \mathrm{d}^3 r' \ \left|\varphi_p(\mathbf{r}')\right|^2 \ W_2(\mathbf{r},\mathbf{r}') \\ V^{\rm ex}(\mathbf{r},\mathbf{r}') &= \sum_{p=1}^{N} \varphi_p^*(\mathbf{r}')\varphi_p(\mathbf{r}) \ W_2(\mathbf{r},\mathbf{r}') \end{aligned}} \tag{46}$$

Note that the terms $p = n$ coming from the two potentials cancel each other; hence they can be eliminated from the two summations, without changing the final result. The contribution of the direct potential is sometimes called the "Hartree term", and the contribution of the exchange potential, the "Fock term". The first is easy to understand: with the exception of the term $p = n$, it corresponds to the interaction of a particle at point $\mathbf{r}$ with all the others at points $\mathbf{r}'$, averaged for each of them by its density distribution $|\varphi_p(\mathbf{r}')|^2$. As for the exchange potential, and in spite of its name, this term is not, strictly speaking, a potential; it is not diagonal in the position representation, even though it basically comes from a particle interaction which is diagonal in that representation. This peculiar non-diagonal form actually comes from the combination of the fermion antisymmetrization and the variational approximation. This exchange potential is homogeneous to a potential divided by the cube of a length. It is obviously a Hermitian operator as it is derived from a potential $W_2(\mathbf{r},\mathbf{r}')$ which is real and symmetric with respect to $\mathbf{r}$ and $\mathbf{r}'$.

A more intuitive and simplified version of these equations was suggested by Hartree, in which the exchange potentials are ignored in (45). Without the integral term, these equations become very similar to a series of Schrödinger equations for independent particles, each of them moving in the mean potential created by all the others (still with the exception of the term $p = n$ in the summation). Including the Fock term should, however, lead to more precise calculations.

Using for the potentials their expressions (46), the Hartree-Fock equations (45) become a set of N coupled equations. They are nonlinear, since the direct and exchange potentials depend on the functions $\varphi_p(\mathbf{r})$. Even though they look like linear eigenvalue equations with eigenfunctions $\varphi_n(\mathbf{r})$ as solutions, a linear resolution would actually require knowing in advance the solutions, since these functions also appear in the potentials (46). The term "self-consistent" is used to characterize this type of situation and the solutions $\varphi_n(\mathbf{r})$ it leads to.

There are no general analytical methods to solve nonlinear self-consistent equations of this type, even in their simplified Hartree version, and numerical methods using successive approximations are commonly used. We start from a series of plausible functions $\varphi_n^{(0)}(\mathbf{r})$, and compute with (46) the associated potentials. Considering them to be fixed, we obtain linear eigenvalue equations which can be solved quite readily with computers (the single very complicated equation in a $3N$-dimensional space has been replaced by N independent 3-dimensional equations); we have to diagonalize a Hermitian operator to get a new series of orthonormal functions, resulting from the first iteration, and called $\varphi_n^{(1)}(\mathbf{r})$ and $\widetilde{e}_n^{(1)}$. The second iteration starts from these $\varphi_n^{(1)}(\mathbf{r})$, to compute the new potential values, and get new linear differential equations. Solving these equations yields the next order $\varphi_n^{(2)}(\mathbf{r})$, $\widetilde{e}_n^{(2)}$, etc. After a few iterations, one expects the $\varphi_n^{(n)}(\mathbf{r})$ and $\widetilde{e}_n^{(n)}$ to vary only slightly with the iteration order (n), in which case the Hartree-Fock

equations have been solved to a good approximation. Using (44) we can then compute the energy we were looking for. It is also possible that physical arguments can help us choose directly adequate trial functions $\varphi_n(\mathbf{r})$ without any iteration. Inserting them in (44) then directly provides the energy.

Comments:

(i) The solutions of the Hartree-Fock equations may not be unique. Using the iteration process described above, one can easily wind up with different solutions, depending on the initial choice for the $\varphi_n^{(0)}(\mathbf{r})$ functions. This multiplicity of solutions is actually one of the method's advantages, as it can help us find not only the ground level but also the excited levels.

(ii) As we shall see in § 2, taking into account the 1/2 spin of the electrons in an atom does not bring major complications to the Hartree-Fock equations. It is generally assumed that the one-body potential is diagonal in a basis of the two spin states, labeled + and −, and that the interaction potential does not act on the spins. We then simply assemble N_+ equations, for N_+ wave functions $\varphi_n^+(\mathbf{r})$ associated with the spin + particles, with N_- other equations, for N_- wave functions $\varphi_n^-(\mathbf{r})$ associated with spin − particles. These two sets of equations are not independent, since they contain the same direct potential (computed using (46), whose first line includes a summation over p of all the $N = N_+ + N_-$ wave functions). As for the exchange potential, it does not lead to any coupling between the two sets of equations: in the second line of (46), the summation over p only includes particles in the same spin state for the following reason. If the particles have opposite spins, they can be recognized by the direction of their spin (the interaction does not act on the spins), and they no longer behave as indistinguishable particles. The exchange effects only arise for particles having the same spin.

2. Generalization: operator method

We now describe the method in a more general way, using an operator method that leads to more concise expressions, while taking into account explicitly the possible existence of a spin – which plays an essential role in the atomic structure. We will identify more precisely the mathematical object, actually a projector, which we vary to optimize the energy. Physically, this projector is simply the one-particle density operator defined in § B-4 of Chapter XV. This will lead to expressions both more compact and general for the Hartree-Fock equations. They contain a Hartree-Fock operator acting on a single particle, as if it were alone, but which includes a potential operator defined by a partial trace which reflects the interactions with the other particles in the mean field approximation. Thanks to this operator we can get an approximate value of the entire system energy, computing only individual energies; these energies are obtained with calculations similar to the one used for a single particle placed in a mean field. With this approach, we have a better understanding of the way the mean field approximately represents the interaction with all the other particles; this approach can also suggest ways to make the approximations more precise.

We assume as before that the N-particle variational ket $\widetilde{|\Psi\rangle}$ is written as:

$$\widetilde{|\Psi\rangle} = a_{\theta_1}^\dagger a_{\theta_2}^\dagger \ldots a_{\theta_N}^\dagger |0\rangle \tag{47}$$

This ket is derived from N individual orthonormal kets $|\theta_k\rangle$, but these kets can now describe particles having an arbitrary spin. Consider the orthonormal basis $\{|\theta_k\rangle\}$ of the one-particle state space, in which the set of $|\theta_i\rangle$ ($i = 1, 2, ...N$) was completed by other orthonormal states. The projector P_N onto the subspace $\mathcal{F}_N$ is the sum of the projections onto the first N kets $|\theta_i\rangle$:

$$P_N = \sum_{i=1}^{N} |\theta_i\rangle \langle\theta_i| \tag{48}$$

This is simply the one-particle density operator defined in § B-4 of Chapter XV (normalized by a trace equal to the particle number N and not to one), as we now show. Relation (B-24) of that chapter can be written in the $|\theta_k\rangle$ basis:

$$\langle\theta_l| \widehat{\rho}_1 |\theta_k\rangle = \left\langle a_k^\dagger a_l \right\rangle \tag{49}$$

where the average value $\left\langle a_k^\dagger a_l \right\rangle$ is taken in the quantum state (47). In this kind of Fock state, the average value is different from zero only when the creation operator reconstructs the population destroyed by the annihilation operator, hence if $k = l$, in which case it is equal to the population n_k of the individual states $|\theta_k\rangle$. In the variational ket (47), all the populations are zero except for the first N states $|\theta_i\rangle$ ($i = 1, 2, ...N$), where they are equal to one. Consequently, the one-particle density operator is represented by a matrix, diagonal in the basis $|\theta_i\rangle$, and whose N first elements on the diagonal are all equal to one. It is indeed the matrix associated with the projector P_N, and we can write:

$$\widehat{\rho}_1 = P_N \tag{50}$$

As we shall see, all the average values useful in our calculation can be simply expressed as a function of this operator.

2-a. Average energy

We now evaluate the different terms included in the average energy, starting with the terms containing one-particle operators.

α. Kinetic and external potential energy

Using relation (B-12) of Chapter XV, we obtain for the average kinetic energy $\left\langle \widehat{H}_0 \right\rangle$:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{r,s} \langle\theta_r| \frac{\mathbf{P}^2}{2m} |\theta_s\rangle \ \ \left\langle a_r^\dagger a_s \right\rangle \tag{51}$$

The same argument as that for the evaluation of the matrix elements (49) shows that the average value $\left\langle a_r^\dagger a_s \right\rangle$ in the state (47) is only different from zero if $r = s$; in that case, it is equal to one when $r \leq N$, and to zero otherwise. This leads to:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{i=1}^{N} \langle\theta_i| \frac{\mathbf{P}^2}{2m} |\theta_i\rangle = \mathrm{Tr}_1 \left\{ P_N \ \frac{\mathbf{P}^2}{2m} \right\} \tag{52}$$

The subscript 1 was added to the trace to underline the fact that this trace is taken in the one-particle state space and not in the Fock space. The two operators included in the trace only act on that same particle, numbered arbitrarily 1; the subscript 1 could obviously be replaced by the subscript of any other particle, since they all play the same role. The average potential energy coming from the external potential is computed in a similar way and can be written as:

$$\left\langle \widehat{V}_{\text{ext}} \right\rangle = \sum_{i=1}^{N} \langle \theta_i | \, V_1 \, | \theta_i \rangle = \text{Tr}_1 \left\{ P_N \; V_1 \right\} \tag{53}$$

β. Average interaction energy, Hartree-Fock potential operator

The average interaction energy $\widehat{W}_{\text{int}}$ can be computed using the general expression (C-16) of Chapter XV for any two-particle operator, which yields:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \sum_{r,s,n,q} \langle 1:\theta_r; 2:\theta_s | \, W_2(1,2) \, | 1:\theta_n; 2:\theta_q \rangle \;\; \left\langle a_r^\dagger a_s^\dagger a_q a_n \right\rangle \tag{54}$$

For the average value $\left\langle a_r^\dagger a_s^\dagger a_q a_n \right\rangle$ in the Fock state $\widetilde{|\Psi\rangle}$ to be different from zero, the operator must leave unchanged the populations of the individual states $|\theta_n\rangle$ and $|\theta_q\rangle$. As in § C-5-b of Chapter XV, two possibilities may occur: either $r = n$ and $s = q$ (the direct term), or $r = q$ and $s = n$ (the exchange term). Commuting some of the operators, we can write:

$$\begin{aligned} \left\langle a_r^\dagger a_s^\dagger a_q a_n \right\rangle &= \delta_{rn} \; \delta_{sq} \left\langle a_r^\dagger a_s^\dagger a_s a_r \right\rangle + \delta_{rq} \; \delta_{sn} \left\langle a_r^\dagger a_s^\dagger a_r a_s \right\rangle \\ &= \left[\delta_{rn} \; \delta_{sq} - \delta_{rq} \; \delta_{sn} \right] n_r n_s \end{aligned} \tag{55}$$

where n_r and n_s are the respective populations of the states $|\theta_r\rangle$ and $|\theta_s\rangle$. Now these populations are different from zero only if the subscripts r and s are between 1 and N, in which case they are equal to 1 (note also that we must have $r \neq s$ to avoid a zero result). We finally get[5]:

$$\begin{aligned} \left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \sum_{i \neq j}^{N} \Big[& \langle 1:\theta_i; 2:\theta_j | \, W_2(1,2) \, | 1:\theta_i; 2:\theta_j \rangle \\ & - \langle 1:\theta_i; 2:\theta_j | \, W_2(1,2) \, | 1:\theta_j; 2:\theta_i \rangle \Big] \end{aligned} \tag{56}$$

(the constraint $i \neq j$ may be ignored since the right-hand side is equal to zero in this case). Here again, the subscripts 1 and 2 label two arbitrary, but different particles, that could have been labeled arbitrarily. We can therefore write:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \sum_{i,j=1}^{N} \langle 1:\theta_i; 2:\theta_j | \, W_2(1,2) \left[1 - P_{\text{ex}}(1,2) \right] | 1:\theta_i; 2:\theta_j \rangle \tag{57}$$

where $P_{\text{ex}}(1,2)$ is the exchange operator between particle 1 and 2 (the transposition which permutes them). This result can be written in a way similar to (53) by introducing a

[5] As in the previous complement, we have replaced $W_2(\mathbf{R}_1, \mathbf{R}_2)$ by $W_2(1,2)$ to simplify the notation

"Hartree-Fock potential" W_{HF}, similar to an external potential acting in the space of particle 1; this potential is defined as the operator having the matrix elements:

$$\langle\theta_i|\, W_{HF}(1)\,|\theta_k\rangle = \sum_{j=1}^{N} \langle 1:\theta_i; 2:\theta_j|\, W_2(1,2)\,[1-P_{\rm ex}(1,2)]\,|1:\theta_k; 2:\theta_j\rangle \tag{58}$$

This operator is Hermitian, since, as the two operators $P_{\rm ex}$ and W_2 are Hermitian and commute, we can write:

$$\begin{aligned}\langle\theta_k|\, W_{HF}(1)\,|\theta_i\rangle &= \sum_{j=1}^{N} \langle 1:\theta_k; 2:\theta_j|\, W_2(1,2)\,[1-P_{\rm ex}(1,2)]\,|1:\theta_i; 2:\theta_j\rangle \\ &= \sum_{j=1}^{N} \langle 1:\theta_i; 2:\theta_j|\,[1-P_{\rm ex}(1,2)]\, W_2(1,2)\,|1:\theta_k; 2:\theta_j\rangle^* \\ &= \langle\theta_i|\, W_{HF}(1)\,|\theta_k\rangle^* \end{aligned} \tag{59}$$

Furthermore, we recognize in (58) the matrix element of a partial trace on particle 2 (Complement $E_{\rm III}$, § 5-b):

$$\boxed{W_{HF}(1) = {\rm Tr}_2\left\{P_N(2)\; W_2(1,2)\,[1-P_{\rm ex}(1,2)]\right\}} \tag{60}$$

where the projector P_N has been introduced inside the trace to limit the sum over j to its first N terms, as in (57). The one-particle operator $W_{HF}(1)$ is thus the partial trace over a second particle (with the arbitrary label 2) of a product of operators acting on both particles. As the summation over j is now taken into account, we are left in (57) with a summation over i, which introduces a trace over the remaining particle 1, and we get:

$$\left\langle \widehat{W}_{\rm int} \right\rangle = \frac{1}{2}{\rm Tr}_1\left\{P_N(1)\; W_{HF}(1)\right\} \tag{61}$$

This average value depends on the subspace chosen with the variational ket $\widetilde{|\Psi\rangle}$ in two ways: explicitly as above, via the projector $P_N(1)$ that shows up in the average value (61), but also implicitly via the definition of the Hartree-Fock potential in (60).

γ. Role of the one-particle reduced density operator

All the average values can be expressed in terms of the projector P_N onto the subspace $\mathcal{F}_N$ of the space the individual states spanned by the N individual states $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$, which means, according to (50), in terms of the one-particle reduced density operator $\widehat{\rho}_1 = P_N$. Hence it is this operator that is the pertinent variable to optimize rather than the set of individual states: certain variations of those states do not change P_N, and are meaningless for our purpose.

Furthermore, the choice of the trial ket $\widetilde{|\Psi\rangle}$ is equivalent to that of P_N. In other words, the variational ket $\widetilde{|\Psi\rangle}$ built in (1) does not depend on the basis chosen in the subspace $\mathcal{F}_N$: if we choose in this subspace any orthonormal basis $\{|u_j\rangle\}$ other than the $\{|\theta_j\rangle\}$ basis, and if we replace in (1) the $a^\dagger_{u_j}$ by the $a^\dagger_{\theta_i}$, the ket will remain the same

(to within a non-relevant phase factor) as we now show. As seen in § A-6 of Chapter XV, each operator $a^{\dagger}_{u_j}$ is a linear combination of the $a^{\dagger}_{\theta_i}$, so that in the product of all the $a^{\dagger}_{u_j}$ ($j = 1, 2, ..N$) we will find products of N operators $a^{\dagger}_{\theta_i}$. Relation (A-43) of Chapter XV however indicates that the squares of any creation operators are zero, which means that the only non-zero products are those including once and only once each of the N different operators $a^{\dagger}_{\theta_i}$. Each term is then proportional to the ket $\widetilde{|\Psi\rangle}$ built from the $a^{\dagger}_{\theta_i}$. Consequently, the two variational kets built from the two bases are necessarily proportional. As definition (1) ensures they are also normalized, they can only differ by a phase factor, which means they are equivalent from a physical point of view. It is thus the operator $P_N = \widehat{\rho}_1$ that best embodies the trial ket $\widetilde{|\Psi\rangle}$.

2-b. Optimization of the one-particle density operator

We now vary $P_N = \widehat{\rho}_1$ to look for the stationary conditions for the total energy:

$$\widetilde{E} = \langle H_0 \rangle + \langle V_1 \rangle + \left\langle \widehat{W}_{\text{int}} \right\rangle = \text{Tr}_1 \left\{ P_N \left[\frac{\mathbf{P}^2}{2m} + V_1 + \frac{1}{2} W_{HF} \right] \right\} \tag{62}$$

We therefore consider the variation:

$$P_N \Rightarrow P_N + \delta P_N \tag{63}$$

which leads to the following variations for the average values of the one-particle operators:

$$\delta \langle H_0 \rangle + \delta \langle V_1 \rangle = \text{Tr}_1 \left\{ \delta P_N \left[\frac{\mathbf{P}^2}{2m} + V_1 \right] \right\} \tag{64}$$

As for the interaction energy, we get two terms:

$$\delta \left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \text{Tr}_1 \left\{ \delta P_N \; W_{HF} \right\} + \frac{1}{2} \text{Tr}_1 \left\{ P_N \; \delta W_{HF} \right\} \tag{65}$$

which are actually equal since:

$$\text{Tr}_1 \left\{ P_N(1) \; \delta W_{HF}(1) \right\} = \text{Tr}_{1,2} \left\{ P_N(1) \; \delta P_N(2) \; \widehat{W}_2(1,2) \left[1 - P_{\text{ex}}(1,2) \right] \right\} \tag{66}$$

and we recognize in the right-hand side of this expression the trace:

$$\text{Tr}_2 \left\{ \delta P_N(2) \; W_{HF}(2) \right\} \tag{67}$$

As we can change the label of the particle from 2 to 1 without changing the trace, the two terms of the interaction energy are equal. As a result, we end up with the energy variation:

$$\delta \widetilde{E} = \text{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] \delta P_N \right\} \tag{68}$$

To vary the projector P_N, we choose a value j_0 of j and make the change:

$$|\theta_{j_0}\rangle \Rightarrow |\theta_{j_0}\rangle + e^{i\chi} \, |\delta\theta\rangle \tag{69}$$

where $|\delta\theta\rangle$ is any ket from the space of individual states, and χ any real number; no other individual state vector varies except for $|\theta_{j_0}\rangle$. The variation of P_N is then written as:

$$\delta P_N = e^{i\chi} |\delta\theta\rangle \langle\theta_{j_0}| + e^{-i\chi} |\theta_{j_0}\rangle \langle\delta\theta| \tag{70}$$

We assume $|\delta\theta\rangle$ has no components on any $|\theta_i\rangle$, that is no components in $\mathcal{F}_N$, since this would change neither $\mathcal{E}_N$, nor the corresponding projector P_N. We therefore impose:

$$P_N |\delta\theta\rangle = 0 \tag{71}$$

which also implies that the norm of $|\theta_{j_0}\rangle$ remains constant[6] to first order in $|\delta\theta\rangle$. Inserting (70) into (68), we obtain:

$$\delta\widetilde{E} = e^{i\chi}\mathrm{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\delta\theta\rangle \langle\theta_{j_0}| \right\} + e^{-i\chi}\mathrm{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\theta_{j_0}\rangle \langle\delta\theta| \right\} \tag{72}$$

For the energy to be stationary, this variation must remain zero whatever the choice of the arbitrary number χ. Now the linear combination of two exponentials $e^{i\chi}$ and $e^{-i\chi}$ will remain zero for any value of χ only if the two factors in front of the exponentials are zero themselves. As each term can be made equal to zero separately, we obtain:

$$0 = \mathrm{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\theta_{j_0}\rangle \langle\delta\theta| \right\} = \langle\delta\theta| \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\theta_{j_0}\rangle \tag{73}$$

This relation must be satisfied for any ket $|\delta\theta\rangle$ orthogonal to the subspace $\mathcal{F}_N$. This means that if we define the one-particle Hartree-Fock operator as:

$$\boxed{H_{HF} = \frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}} \tag{74}$$

the stationary condition for the total energy is simply that the ket $H_{HF} |\theta_{j_0}\rangle$ must belong to $\mathcal{F}_N$:

$$H_{HF} |\theta_{j_0}\rangle \in \mathcal{F}_N \tag{75}$$

As this relation must hold for any $|\theta_{j_0}\rangle$ chosen among the $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$, it follows that the subspace $\mathcal{F}_N$ is stable under the action of the operator (74).

2-c. Mean field operator

We can then restrict the operator H_{HF} to that subspace:

$$\bar{H}_{HF} = P_N(1) \left[\frac{\mathbf{P}^2}{2m} + V_1(1) + W_{HF}(1) \right] P_N(1) \tag{76}$$

[6]Since (71) shows that $|\delta\theta\rangle$ is orthogonal to any linear combinations of the $|\theta_i\rangle$, we can write $(\langle\theta_{j_0}| + \langle\delta\theta|)\,(|\theta_{j_0}\rangle + |\delta\theta\rangle) = \langle\theta_{j_0}|\,\theta_{j_0}\rangle + \langle\delta\theta|\,\delta\theta\rangle = 1+$ second order terms.

This operator, acting in the subspace $\mathcal{F}_N$ spanned by the N kets $|\theta_j\rangle$, is a Hermitian linear operator, hence it can be diagonalized. We call $|\varphi_n\rangle$ its eigenvectors (n = 1, 2, ..N), which are linear combinations of the kets $|\theta_j\rangle$. The stationary condition for the energy (75) amounts to imposing the $|\varphi_n\rangle$ to be not only eigenvectors of $\bar{H}_{HF}$, but also of the operator H_{HF} defined by (74) in the entire one-particle state space (without the restriction to $\mathcal{F}_N$); consequently, the $|\varphi_n\rangle$ must obey:

$$\boxed{H_{HF}\,|\varphi_n\rangle = \widetilde{e}_n\ |\varphi_n\rangle} \tag{77}$$

Operator H_{HF} is defined in (74), where the operator W_{HF} is given by (60) and depends on the projector P_N. This last operator may be expressed as a function of the $|\varphi_n\rangle$ in the same way as with the $|\theta_i\rangle$, and relation (48) may be replaced by:

$$P_N = \sum_{n=1}^{N} |\varphi_n\rangle\,\langle\varphi_n| \tag{78}$$

Relations (77), together with definition (60) where (78) has been inserted, are a set of equations allowing the self-consistent determination of the $|\varphi_n\rangle$; they are called the Hartree-Fock equations. This operator form (77) is simpler than the one obtained in § 1-c; it emphasizes the similarity with the usual eigenvalue equation for a single particle moving in an external potential, illustrating the concept of a self-consistent mean field. One must keep in mind, however, that via the projector (78) included in W_{HF}, this particle moves in a potential depending on the whole set of states occupied by all the particles. Remember also that we did not carry out an exact computation, but merely presented an approximate theory (variational method).

The discussion in § 1-f is still relevant. As the operator W_{HF} depends on the $|\varphi_n\rangle$, the Hartree-Fock equations have an intrinsic nonlinear character, which generally requires a resolution by successive approximations. We start from a set of N individual states $\left|\varphi_n^0\right\rangle$ to build a first value of P_N and the operator W_{HF}, which are used to compute the Hamiltonian (74). Considering this Hamiltonian now fixed, the Hartree-Fock equations (77) become linear, and can be solved as usual eigenvalue equations. This leads to new values $\left|\varphi_n^1\right\rangle$ for the $|\varphi_n\rangle$, and finishes the first iteration. In the second iteration, we use the $\left|\varphi_n^1\right\rangle$ in (78) to compute a new value of the mean field operator W_{HF}; considering again this operator as fixed, we solve the eigenvalue equation and obtain the second iteration values $\left|\varphi_n^2\right\rangle$ for the $|\varphi_n\rangle$, and so on. If the initial values $\left|\varphi_n^0\right\rangle$ are physically reasonable, one can hope for a rapid convergence towards the expected solution of the nonlinear Hartree-Fock equations.

The variational energy can be computed in the same way as in § 1-e. Multiplying on the left equation (77) by the bra $\langle\varphi_n|$, we get:

$$\widetilde{e}_n = \langle\varphi_n|\left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}\right]|\varphi_n\rangle \tag{79}$$

After summing over the subscript n, we obtain:

$$\sum_{n=1}^{N}\widetilde{e}_n = \sum_{n=1}^{N}\langle\varphi_n|\left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}\right]|\varphi_n\rangle = \mathrm{Tr}_1\left\{P_N(1)\left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}\right]\right\} \tag{80}$$

Taking into account (51), (53), and (61), we get:

$$\sum_{n=1}^{N} \widetilde{e}_n = \langle H_0 \rangle_{HF} + \langle V_1 \rangle_{HF} + 2 \left\langle \widehat{W}_{\rm int} \right\rangle_{HF} \tag{81}$$

where the particle interaction energy is counted twice. To compute the energy $\widetilde{E}$, we can eliminate $\left\langle \widehat{W}_{\rm int} \right\rangle$ between (26) and this relation and we finally obtain:

$$\widetilde{E} = \frac{1}{2} \left[\sum_{n=1}^{N} \widetilde{e}_n + \langle H_0 \rangle_{HF} + \langle V_1 \rangle_{HF} \right] \tag{82}$$

2-d. Hartree-Fock equations for electrons

Assume the fermions we are studying are particles with spin 1/2, electrons for example. The basis $\{|\mathbf{r}\rangle\}$ of the individual states used in § 1 must be replaced by the basis formed with the kets $\{|\mathbf{r}, \nu\rangle\}$, where ν is the spin index, which can take 2 distinct values noted $\pm 1/2$, or more simply $\pm$. To the summation over $\mathrm{d}^3 r$ we must now add a summation over the 2 values of the index spin ν. A vector $|\varphi\rangle$ in the individual state space is now written:

$$|\varphi\rangle = \sum_{\nu = \pm 1/2} \int \mathrm{d}^3 r \; \varphi(\mathbf{r}, \nu) \, |\mathbf{r}, \nu\rangle \tag{83}$$

with:

$$\varphi(\mathbf{r}, \nu) = \langle \mathbf{r}, \nu \, | \varphi \rangle \tag{84}$$

The variables $\mathbf{r}$ and ν play a similar role but the first one is continuous whereas the second is discrete. Writing them in the same parenthesis might hide this difference, and we often prefer noting the discrete index as a superscript of the function φ, and write:

$$\varphi^{\nu}(\mathbf{r}) = \langle \mathbf{r}, \nu \, | \varphi \rangle \tag{85}$$

Let us build an N particle variational state $\widetilde{|\Psi\rangle}$ from N orthonormal states $|\varphi_n^{\nu_n}\rangle$, with $n = 1, 2, .. , N$. Each of the $|\varphi_n^{\nu_n}\rangle$ describes an individual state including the spin and position variables; the first N_+ values of ν_n ($n = 1, 2, ..N_+$) are equal to $+1/2$, the last N_- are equal to $-1/2$, with $N_+ + N_- = N$ (we assume N_+ and N_- are fixed for the moment but we may allow them to vary later to enlarge the variational family). In the space of the individual states, we introduce a complete basis $\{|\varphi_k^{\nu_k}\rangle\}$ whose first N kets are the $|\varphi_n^{\nu_n}\rangle$, but where the subscript k varies from 1 to infinity[7].

We assume the matrix elements of the external potential V_1 to be diagonal for ν; these two diagonal matrix elements can however take different values $V_1^{\pm}(\mathbf{r})$, which allows including the eventual presence of a magnetic field coupled with the spins. We also assume the particle interaction $W_2(1,2)$ to be independent of the spins, and diagonal in the position representation of the two particles, as is the case, for example, for the Coulomb

[7] The subscript k determines both the orbital and the spin state of the particle; the index ν is not independent since it is fixed for each value of k.

interaction between electrons. With these assumptions, the Hamiltonian cannot couple states having different particle numbers N_+ and N_-.

Let us see what the general Hartree-Fock equations become in the $\{|\mathbf{r},\nu\rangle\}$ representation. In this representation, the effect of the kinetic and potential operators are well known. We just have to compute the effect of the Hartree-Fock potential W_{HF}. To obtain its matrix elements, we use the basis $\{|1:\mathbf{r},\nu;\ 2:\varphi_k^{\nu_k}\rangle\}$ to write the trace in (60):

$$\langle \mathbf{r},\nu|\, W_{HF}(1)\,|\mathbf{r}',\nu'\rangle = \sum_{k=1}^{\infty}\ \langle 1:\mathbf{r},\nu;2:\varphi_k^{\nu_k}|\left[\sum_{p=1}^{N}\left|2:\varphi_p^{\nu_p}\right\rangle\left\langle 2:\varphi_p^{\nu_p}\right|\right] W_2(1,2)\,[1-P_{\rm ex}(1,2)]\,|1:\mathbf{r}',\nu';2:\varphi_k^{\nu_k}\rangle \tag{86}$$

As the right-hand side includes the scalar product $\langle 2:\varphi_k^{\nu_k}\ |2:\varphi_p^{\nu_p}\rangle$ which is equal to δ_{kp}, the sum over k disappears and we get:

$$\langle \mathbf{r},\nu|W_{HF}(1)\,|\mathbf{r}',\nu'\rangle = \sum_{p=1}^{N}\left\langle 1:\mathbf{r},\nu;2:\varphi_p^{\nu_p}\right| W_2(1,2)\,[1-P_{\rm ex}(1,2)]\left|1:\mathbf{r}',\nu';2:\varphi_p^{\nu_p}\right\rangle \tag{87}$$

(i) We first deal with the direct term contribution, hence ignoring in the bracket the term in $P_{\rm ex}(1,2)$. We can replace the ket $\left|2:\varphi_p^{\nu_p}\right\rangle$ by its expression:

$$\left|2:\varphi_p^{\nu_p}\right\rangle = \int \mathrm{d}^3 r_2\ \varphi_p^{\nu_p}(\mathbf{r}_2)\,|2:\mathbf{r}_2,\nu_p\rangle \tag{88}$$

As the operator is diagonal in the position representation, we can write:

$$\begin{aligned} W_2(1,2)\left|1:\mathbf{r}',\nu';2:\varphi_p^{\nu_p}\right\rangle &= W_2(1,2)\int \mathrm{d}^3 r_2\ \varphi_p^{\nu_p}(\mathbf{r}_2)\,|1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\rangle \\ &= \int \mathrm{d}^3 r_2\ \varphi_p^{\nu_p}(\mathbf{r}_2)\ W_2(\mathbf{r}',\mathbf{r}_2)\,|1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\rangle \end{aligned} \tag{89}$$

The direct term of (87) is then written:

$$\int \mathrm{d}^3 r_2 \sum_{p=1}^{N} W_2(\mathbf{r}',\mathbf{r}_2)\ \varphi_p^{\nu_p}(\mathbf{r}_2)\ \left\langle 1:\mathbf{r},\nu;2:\varphi_p^{\nu_p}\ |1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\right\rangle \tag{90}$$

where the scalar product of the bra and the ket is equal to $\delta_{\nu\nu'}\delta(\mathbf{r}-\mathbf{r}')\left[\varphi_p^{\nu_p}(\mathbf{r}_2)\right]^*$. We finally obtain:

$$\delta_{\nu\nu'}\ \delta(\mathbf{r}-\mathbf{r}')\int \mathrm{d}^3 r_2\ W_2(\mathbf{r},\mathbf{r}_2)\sum_{p=1}^{N}\left|\varphi_p^{\nu_p}(\mathbf{r}_2)\right|^2 = \delta_{\nu\nu'}\ \delta(\mathbf{r}-\mathbf{r}')V_{\rm dir}(\mathbf{r}) \tag{91}$$

with:

$$V_{\rm dir}(\mathbf{r}) = \sum_{p=1}^{N}\int \mathrm{d}^3 r'\ W_2(\mathbf{r},\mathbf{r}')\ \left|\varphi_p^{\nu_p}(\mathbf{r}')\right|^2 \tag{92}$$

This component of the mean field (Hartree term) contains a sum over all occupied states, whatever their spin is; it is spin independent.

(ii) We now turn to the exchange term, which contains the operator $P_{\rm ex}(1,2)$ in the bracket of (87). To deal with it, we can for example commute in (87) the two operators $W_2(1,2)$ and $P_{\rm ex}(1,2)$; this last operator will then permute the two particles in the bra. Performing this operation in (90), we get, with the minus sign of the exchange term:

$$-\int d^3r_2 \sum_{p=1}^{N} W_2(\mathbf{r}',\mathbf{r}_2)\ \varphi_p^{\nu_p}(\mathbf{r}_2)\ \left\langle 1:\varphi_p^{\nu_p};2:\mathbf{r},\nu\ |1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\right\rangle \tag{93}$$

The scalar product will yield the products of $\delta_{\nu\nu_p}\delta_{\nu_p\nu'}\delta(\mathbf{r}-\mathbf{r}_2)$, making the integral over d^3r_2 disappear; this term is zero if $\nu\neq\nu'$, hence the factor $\delta_{\nu\nu'}$. Since $W_2(\mathbf{r}',\mathbf{r}) = W_2(\mathbf{r},\mathbf{r}')$, we are left with:

$$-\delta_{\nu\nu'}\sum_{\nu_p=\nu} W_2(\mathbf{r},\mathbf{r}')\left[\varphi_p^{\nu_p}(\mathbf{r})\right]\left[\varphi_p^{\nu_p}(\mathbf{r}')\right]^* = -\delta_{\nu\nu'}V_{ex}^{\nu}(\mathbf{r},\mathbf{r}') \tag{94}$$

where the sum is over the values of p for which $\nu_p=\nu=\nu'$ (hence, limited to the first N_+ values of p, or the last N_-, depending on the case); the exchange potential $V_{\rm ex}^{\nu}$ has been defined as:

$$V_{\rm ex}^{\nu}(\mathbf{r},\mathbf{r}') = \sum_{\nu_p=\nu} W_2(\mathbf{r},\mathbf{r}')\ \left[\varphi_p^{\nu_p}(\mathbf{r}')\right]^*\left[\varphi_p^{\nu_p}(\mathbf{r})\right] \tag{95}$$

As is the case for the direct term, the exchange term does not act on the spin. There are however two differences. To begin with, the summation over p is limited to the states having the same spin ν; second, it introduces a contribution which is non-diagonal in the positions (but without an integral), and which cannot be reduced to an ordinary potential (the term "non-local potential" is sometimes used to emphasize this property).

We have shown that the scalar product of equation (77) with $\langle\mathbf{r},\nu|$ introduces three potentials (in addition to the the one-body potential $V_1^{\nu_n}$), a direct potential $V_{\rm dir}(\mathbf{r})$ and two exchange potentials $V_{\rm ex}^{\nu}(\mathbf{r},\mathbf{r}')$ with $\nu=\pm1/2$. Equation (77) then becomes, in the $\{|\mathbf{r},\nu\rangle\}$ representation, a pair of equations:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1^{\nu_n}(\mathbf{r}) + V_{\rm dir.}(\mathbf{r})\right]\varphi_n^{\nu_n}(\mathbf{r}) - \int \mathrm{d}^3r'\ V_{\rm ex}^{\nu_n}(\mathbf{r},\mathbf{r}')\varphi_n^{\nu_n}(\mathbf{r}') = \widetilde{e}_n\ \varphi_n^{\nu_n}(\mathbf{r})} \tag{96}$$

These are the Hartree-Fock equations with spin and in the position representation, widely used in quantum physics and chemistry. It is not necessary to worry, in these equations, about the term in which the subscript p in the summation appearing in (92) and (95) is the same as the subscript n (of the wave function we are looking for); the contributions $n=p$ cancel each other exactly in the direct and exchange potentials.

Both the "Hartree term" giving the direct potential contribution, and the "Fock term " giving the exchange potential, can be interpreted in the same way as above (§ 1-f). The Hartree term contains the contributions of all the other electrons to the mean potential felt by one electron. The exchange potential, on the other hand, only involves electrons in the same spin state, and this can be simply interpreted: the exchange effect only occurs for two totally indistinguishable particles. Now if these particles are in

orthogonal spin states, and as the interactions do not act on the spins, one can in principle determine which is which and the particles become distinguishable: the quantum exchange effects cancel out. As we already pointed out, the exchange potential is not a potential stricto sensu. It is not diagonal in the position representation, even though it basically comes from a particle interaction that is diagonal in position. It is the antisymmetrization of the fermions, together with the chosen variational approximation, which led to this peculiar non-diagonal form. It is however a Hermitian operator, as can be shown using the fact that the initial potential $W_2(\mathbf{r}, \mathbf{r}')$ is real and symmetric with respect to $\mathbf{r}$ and $\mathbf{r}'$.

2-e. Discussion

The resolution of the nonlinear Hartree-Fock equations is generally done by the successive iteration approximate method discussed in § 1-f. There is no particular reason for the solution of the Hartree-Fock equations to be unique[8]; on the contrary, they can yield solutions that depend on the states chosen to begin the nonlinear iterations. They can actually lead to a whole spectrum of possible energies for the system. This is how the ground state and excited state energies of the atom are generally computed. The atomic orbitals discussed in Complement E_{VII}, the central field approximation and the electronic "configurations" discussed in Complement B_{XIV} can now be discussed in a more precise and quantitative way. We note that the exchange energy, introduced in this complement for a two-electron system, is a particular case of the exchange energy term of the Hartree-Fock potential. There exist however many other physical systems where the same ideas can be applied: nuclei (the Coulomb force is then replaced by the nuclear interaction force between the nucleons), atomic aggregates (with an interatomic potential having both repulsive and attractive components, see Complements C_{XI} and G_{XI}), and many others.

Once a Hartree-Fock solution for a complex problem has been found, we can go further. One can use the basis of the eigenfunctions just obtained as a starting point for more precise perturbation calculations, including for example correlations between particles (Chapter XI). In atomic spectra, we sometimes find cases where two configurations yield very close mean field energies. The effects of the interaction terms beyond the mean field approximation will then be more important. Perturbation calculations limited to the space of the configurations in question permits obtaining better approximations for the energy levels and their wave functions; one then speaks of "mixtures", or of "interactions between configurations".

Comment:

The variational method based on the Fock states is not the only one that leads to the Hartree-Fock equations. One could also start from an approximation of the two-particle density operator ρ_{II} by a function of the one-particle density operator ρ_I and write:

$$\rho_{II}(1,2) \propto \frac{1}{2}\left[1 - P_{\text{ex}}(1,2)\right] \rho_I(1) \otimes \rho_I(2) \tag{97}$$

Expressing the energy of the N-particle system as a function of ρ_I, we minimize it by varying this operator, and find the same results as above. This method amounts to a closure of the hierarchy of the N-body equations (§ C-4 of Chapter XVI). We have in

[8] They all yield, however, an upper limit for the ground state energy

fact already seen with equation (21) and in § 2-a that the Hartree-Fock approximation amounts to expressing the two-particle correlation functions as a function of the one-particle correlation functions. In terms of correlation functions (Complement A_{XVI}), this amounts to replacing the two-particle function (four-point function) by a product of one-particle functions (two-point function), including an exchange term. Finally, another method is to use the diagram perturbation theory; the Hartree-Fock approximation corresponds to retaining only a certain class of diagrams (class of connected diagrams).

Finally note that the Hartree-Fock method is not the only one yielding approximate solutions of Schrödinger's equation for a system of interacting fermions; in particular, one can use the "electronic density functional" theory (a functional is a function of another function, as for instance the action S in classical lagrangian mechanics). The method is used to obtain the electronic structure of molecules or condensed phases in physics, chemistry, and materials science. Its study nevertheless lies outside the scope of this book, and the reader is referred to [6], which summarizes the method and gives a number of references.

Complement F_{XV}

Fermions, time-dependent Hartree-Fock approximation

1 Variational ket and notation 1701
2 Variational method 1702
2-a Definition of a functional 1702
2-b Stationarity 1703
2-c Particular case of a time-independent Hamiltonian 1705
3 Computing the optimizer 1705
3-a Average energy 1705
3-b Hartree-Fock potential 1706
3-c Time derivative 1707
3-d Functional value 1707
4 Equations of motion 1707
4-a Time-dependent Hartree-Fock equations 1708
4-b Particles in a single spin state 1709
4-c Discussion 1709

The Hartree-Fock mean field method was introduced in Complement E_{XV} for a time-independent problem: the search for the stationary states of a system of interacting fermions (the search for its thermal equilibrium will be discussed in Complement G_{XV}. In this complement, we show how this method can be used for time-dependent problems. We start, in § 1, by including a time dependence in the Hartree-Fock variational ket (time-dependent Fock state). We then introduce in § 2 a general variational principle that can be used for solving the time-dependent Schrödinger equation. We then compute, in § 3, the function to be optimized for a Fock state; the same mean field operator as the one introduced in Complement E_{XV} will here again play a very useful role. Finally, the time-dependent Hartree-Fock equations will be obtained and discussed in § 4. More details on the Hartree-Fock methods in general can be found, for example, in Chapter 7 of reference [5], and especially in its Chapter 9 for time-dependent problems.

1. Variational ket and notation

We assume the N-particle state vector $\widetilde{|\Psi(t)\rangle}$ to be of the form:

$$\widetilde{|\Psi(t)\rangle} = a^\dagger_{\theta_1(t)} a^\dagger_{\theta_2(t)} \ ... \ a^\dagger_{\theta_N(t)} |0\rangle \tag{1}$$

where the $a^\dagger_{\theta_1(t)}$, $a^\dagger_{\theta_2(t)}$, ..., $a^\dagger_{\theta_N(t)}$ are the creation operators associated with an arbitrary series of orthonormal individual states $|\theta_1(t)\rangle$, $|\theta_2(t)\rangle$, ..., $|\theta_N(t)\rangle$ which depend on time t. This series is, for the moment, arbitrary, but the aim of the following variational calculation is to determine its time dependence.

As in the previous complements, we assume that the Hamiltonian $\widehat{H}$ is the sum of three terms: a kinetic energy Hamiltonian, an external potential Hamiltonian, and a particle interaction term:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\text{ext}}(t) + \widehat{W}_{\text{int}} \tag{2}$$

with:

$$\widehat{H}_0 = \sum_{q=1}^{N} \frac{(\mathbf{P}_q)^2}{2m} \tag{3}$$

(m is the particles' mass, $\mathbf{P}_q$ the momentum operator of particle q), and:

$$\widehat{V}_{\text{ext}}(t) = \sum_{q=1}^{N} V_1(\mathbf{R}_q, t) \tag{4}$$

and finally:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{q \neq q'} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{5}$$

2. Variational method

Let us introduce a general variational principle; using the stationarity of a functional S of the state vector $|\Psi(t)\rangle$, it will yield the time-dependent Schrödinger equation.

2-a. Definition of a functional

Consider an arbitrarily given Hamiltonian $H(t)$. We assume the state vector $|\Psi(t)\rangle$ to have any time dependence, and we note $\left|\overline{\Psi}(t)\right\rangle$ the ket physically equivalent to $|\Psi(t)\rangle$, but with a constant norm:

$$\left|\overline{\Psi}(t)\right\rangle = \frac{|\Psi(t)\rangle}{\sqrt{\langle\Psi(t)\ |\Psi(t)\rangle}} \tag{6}$$

The functional S of $\left|\overline{\Psi}(t)\right\rangle$ is defined as[1]:

$$\begin{aligned} S\left[\left|\overline{\Psi}(t)\right\rangle\right] &= \int_{t_0}^{t_1} \mathrm{d}t\ \ \mathrm{Re}\ \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t)\right] \left|\overline{\Psi}(t)\right\rangle \\ &= \int_{t_0}^{t_1} \mathrm{d}t \left\{ \frac{i\hbar}{2} \left[\left\langle\overline{\Psi}(t)\right| \frac{\mathrm{d}}{\mathrm{d}t}\ \overline{\Psi}(t)\right\rangle - \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\overline{\Psi}(t)\ \left|\overline{\Psi}(t)\right\rangle\right] - \left\langle\overline{\Psi}(t)\right| H(t) \left|\overline{\Psi}(t)\right\rangle \right\} \end{aligned} \tag{7}$$

where t_0 and t_1 are two arbitrary times such that $t_0 < t_1$. In the particular case where the chosen $\left|\overline{\Psi}(t)\right\rangle$ is equal to a solution $\left|\overline{\Psi}_S(t)\right\rangle$ of the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\overline{\Psi}_S(t)\right\rangle = H(t)\left|\overline{\Psi}_S(t)\right\rangle \tag{8}$$

[1]The notation where the differential operator $\mathrm{d}/\mathrm{d}t$ is written between a bra and a ket means that the operator takes the derivative of the ket that follows (and not of the bra just before).

the bracket on the first line of (7) obviously cancels out and we have:

$$S\left[\left|\overline{\Psi}_S(t)\right\rangle\right] = 0 \tag{9}$$

Integrating by parts the second term[2] of the bracket in the second line of (7), we get the same form as the first term in the bracket, plus an already integrated term. The final result is then:

$$\begin{aligned} S\left[\left|\overline{\Psi}(t)\right\rangle\right] &= \left\{\int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t)\right] \left|\overline{\Psi}(t)\right\rangle\right\} \\ &\qquad + \frac{i\hbar}{2}\left[\left\langle\overline{\Psi}(t_0)\right| \left.\overline{\Psi}(t_0)\right\rangle - \left\langle\overline{\Psi}(t_1)\right| \left.\overline{\Psi}(t_1)\right\rangle\right] \\ &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t)\right] \left|\overline{\Psi}(t)\right\rangle \end{aligned} \tag{10}$$

where we have used in the second line the fact that the norm of $\left|\overline{\Psi}(t)\right\rangle$ always remains equal to unity. This expression for S is similar to the initial form (7), but without the real part.

2-b. Stationarity

Suppose now $\left|\overline{\Psi}(t)\right\rangle$ has an arbitrary time dependence between t_0 and t_1, while keeping its norm constant, as imposed by (6); the functional then takes a certain value S, a priori different from zero. Let us see under which conditions S will be stationary when $\left|\overline{\Psi}(t)\right\rangle$ changes by an infinitely small amount $\left|\delta\overline{\Psi}(t)\right\rangle$:

$$\left|\overline{\Psi}(t)\right\rangle \Rightarrow \left|\overline{\Psi}(t)\right\rangle + \left|\delta\overline{\Psi}(t)\right\rangle \tag{11}$$

For what follows, it will be convenient to assume that the variation $\left|\delta\overline{\Psi}(t)\right\rangle$ is free; we therefore have to ensure that the norm of $\left|\overline{\Psi}(t)\right\rangle$ remains constant, equal to unity[3]. We introduce Lagrange multipliers (Appendix V) $\lambda(t)$ to control the square of the norm at every time between t_0 and t_1, and we look for the stationarity of a function where the sum of constraints has been added. This sum introduces an integral, and we the function in question is:

$$\begin{aligned} \overline{S}\left[\left|\overline{\Psi}(t)\right\rangle\right] &= S\left[\left|\overline{\Psi}(t)\right\rangle\right] - \int_{t_0}^{t_1} \mathrm{d}t \ \lambda(t) \left\langle\overline{\Psi}(t) \ \left|\overline{\Psi}(t)\right\rangle\right. \\ &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\overline{\Psi}(t)\right\rangle \end{aligned} \tag{12}$$

where $\lambda(t)$ is a real function of the time t.

[2] If we integrate by parts the first term rather than the second, we get the complex conjugate of equation (10), which brings no new information.

[3] For the normalization of $\left|\overline{\Psi}(t)\right\rangle$ to be conserved to first order, it is necessary (and sufficient) for the scalar product $\left\langle\overline{\Psi}(t) \ \left|\delta\overline{\Psi}(t)\right\rangle\right.$ to be zero or purely imaginary. If this is the case, the Lagrangian multiplier $\lambda(t)$ is not needed

The variation $\delta\overline{S}$ of $\overline{S}$ to first order is obtained by inserting (11) in (10). It yields the sum of a first term $\delta\overline{S}_1$ containing the ket $\left|\delta\overline{\Psi}(t)\right\rangle$ and of another $\delta\overline{S}_2$ containing the bra $\left\langle\delta\overline{\Psi}(t)\right|$:

$$\begin{aligned}\delta\overline{S}_1 &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\delta\overline{\Psi}(t)\right\rangle \\ \delta\overline{S}_2 &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\delta\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\overline{\Psi}(t)\right\rangle \end{aligned} \tag{13}$$

We now imagine another variation for the ket:

$$\left|\overline{\Psi}(t)\right\rangle \Rightarrow \left|\overline{\Psi}(t)\right\rangle + i\left|\delta\overline{\Psi}(t)\right\rangle \tag{14}$$

which yields a variation $\delta'\overline{S}$ of $\overline{S}$; in this second variation, the term in $\left|\delta\overline{\Psi}(t)\right\rangle$ becomes $\delta'\overline{S}_1 = i\delta\overline{S}_1$, whereas the term in $\left\langle\delta\overline{\Psi}(t)\right|$ becomes $\delta'\overline{S}_2 = -i\delta\overline{S}_2$. Now, if the functional is stationary in the vicinity of $\left|\overline{\Psi}(t)\right\rangle$, the two variations $\delta\overline{S}$ and $\delta'\overline{S}$ are necessarily zero, as are also $\delta\overline{S} - i\delta'\overline{S}$ and $\delta\overline{S} + i\delta'\overline{S}$. In those combinations, only terms in $\delta\overline{S}_1$ appear for the first one, and in $\delta\overline{S}_2$ for the second; consequently they must both be zero. As a result, we can write the stationarity conditions with respect to variations of the bra and the ket separately.

Let us write for example that $\delta\overline{S}_2 = 0$, which means the right-hand side of the second line in (13) must be zero. As the time evolution between t_0 and t_1 of the bra $\left\langle\delta\overline{\Psi}(t)\right|$ is arbitrary, this condition imposes this bra multiplies a zero-value ket, at all times. Consequently, the ket $\left|\overline{\Psi}(t)\right\rangle$ must obey the equation:

$$\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\overline{\Psi}(t)\right\rangle = 0 \tag{15}$$

which is none other than the Schrödinger equation associated with the Hamiltonian $H(t) + \lambda(t)$.

Actually, $\lambda(t)$ simply introduces a change in the origin of the energies and this only modifies the total phase[4] of the state vector $\left|\overline{\Psi}(t)\right\rangle$, which has no physical effect. Without loss of generality, this Lagrange factor may therefore be ignored, and we can set:

$$\lambda(t) = 0 \tag{16}$$

A necessary condition[5] for the stationarity of S is that $\left|\overline{\Psi}(t)\right\rangle$ obey the Schrödinger equation (8) – or be physically equivalent (i.e. equal to within a global time-dependent phase factor) to a solution of this equation. Conversely, assume $\left|\overline{\Psi}(t)\right\rangle$ is a solution of the Schrödinger equation, and give this ket a variation as in (11). It is then obvious from the second line of (13) that $\delta\overline{S}_2$ is zero. As for $\delta\overline{S}_1$, an integration by parts over time shows that it is the complex conjugate of $\delta\overline{S}_2$, and therefore also equal to zero. The

[4] If in (15) we set $\left|\overline{\Psi}(t)\right\rangle = e^{-i\alpha(t)}\left|\overline{\Theta}(t)\right\rangle$, we see that $\left|\overline{\Theta}(t)\right\rangle$ obeys the differential equation obtained by replacing $\lambda(t)$ by $\lambda(t) - \hbar\frac{\mathrm{d}\alpha}{\mathrm{d}t}$ in (15). If we simply choose for $\alpha(t)$ the integral over time of the function $\lambda(t)$, this constant will disappear from the differential equation.

[5] The same argument as above, but starting from the variation $\delta S - i\delta' S$, would lead to the complex conjugate of (8), and hence to the same equation.

functional S is thus stationary in the vicinity of any exact solution of the Schrödinger equation.

Suppose we choose any variational family $\mathcal{F}$ of normalized kets $\widetilde{|\Psi_{\mathcal{F}}(t)\rangle}$, but which now includes a ket $\widetilde{|\Psi^{\text{stat}}_{\mathcal{F}}(t)\rangle}$ for which S is stationary. A simple example is the case where $\mathcal{F}$ is a family $\mathcal{F}_0$ that contains the exact solution of the Schrödinger equation; according to what we just saw, this exact solution will make S stationary, and conversely, the ket that makes S stationary is necessarily $\widetilde{\left|\Psi^{\text{stat}}_{\mathcal{F}_0}(t)\right\rangle}$. In this case, imposing the variation of S to be zero allows identifying, inside the family $\mathcal{F}_0$, the exact solution we are looking for. If we now change the family continuously from $\mathcal{F}_0$ to $\mathcal{F}$, in general $\mathcal{F}$ will no longer contain the exact solution of the Schrödinger equation. We can however follow the modifications at all times of the values of the ket $\widetilde{|\Psi^{\text{stat}}_{\mathcal{F}}(t)\rangle}$. Starting from an exact solution of the equation, this ket progressively changes, but, by continuity, will stay in the vicinity of this exact solution if $\mathcal{F}$ stays close to $\mathcal{F}_0$. This is why annulling the variation of S in the family $\mathcal{F}$ is a way of identifying a member of that family whose evolution remains close to that of a solution of the Schrödinger equation. This is the method we will follow, using the Fock states as a particular variational family.

2-c. Particular case of a time-independent Hamiltonian

If the Hamiltonian H is time-independent, one can look for time-independent kets $\left|\overline{\Psi}\right\rangle$ to make the functional S stationary. The function to be integrated in the definition of the functional S also becomes time-independent, and we can write S as:

$$S = (t_1 - t_0) \left\langle \overline{\Psi} \right| H \left| \overline{\Psi} \right\rangle \tag{17}$$

Since the two times t_0 and t_1 are fixed, the stationarity of S is equivalent to that of the diagonal matrix element of the Hamiltonian $\left\langle \overline{\Psi} \right| H \left| \overline{\Psi} \right\rangle$. We find again the stationarity condition of the time-independent variational method (Complement E_{XI}), which appears as a particular case of the more general method of the time-dependent variations. Consequently, it is not surprising that the Hartree-Fock methods, time-dependent or not, lead to the same Hartree-Fock potential, as we now show.

3. Computing the optimizer

The family of the state vectors we consider is the set of Fock kets $\widetilde{|\Psi(t)\rangle}$ defined in (1). We first compute the function to be integrated in the functional (10) when $|\Psi(t)\rangle$ takes the value $\widetilde{|\Psi(t)\rangle}$.

3-a. Average energy

For the term in $H(t)$, the calculation is identical to the one we already did in § 1-b of Complement E_{XV}. We first add to the series of orthonormal states $|\theta_i(t)\rangle$ with $i = 1, 2, ..., N$ other orthonormal states $|\theta_i(t)\rangle$ with $i = N+1, N+2, ...$, to obtain a complete orthonormal basis in the space of individual states. Using this basis, we can express the one-particle and two-particle operators according to relations (B-12) and (C-16) of Chapter XV. This presents no difficulty since the average values of creation and annihilation operator products are easily obtained in a Fock state (they only differ

from zero if the product of operators leaves the populations of the individual states unchanged). Relations (52), (53) and (57) of Complement E_{XV} are still valid when the $|\theta_i\rangle$ become time-dependent. We thus get for the average kinetic energy:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{i=1}^{N} \langle \theta_i(t)| \frac{\mathbf{P}^2}{2m} |\theta_i(t)\rangle \tag{18}$$

for the external potential energy:

$$\left\langle \widehat{V}_{\rm ext}(t) \right\rangle = \sum_{i=1}^{N} \langle \theta_i(t)| V_1(\mathbf{R},t) |\theta_i(t)\rangle \tag{19}$$

and for the interaction energy:

$$\left\langle \widehat{W}_{\rm int} \right\rangle = \frac{1}{2} \sum_{i,j}^{N} \langle 1:\theta_i(t); 2:\theta_j(t)| W_2(1,2)\,[1 - P_{\rm ex}(1,2)]\,|1:\theta_i(t); 2:\theta_j(t)\rangle \tag{20}$$

3-b. Hartree-Fock potential

We recognize in (20) the diagonal element ($i = k$) of the Hartree-Fock potential operator $W_{HF}(1,t)$ whose matrix elements have been defined in a general way by relation (58) of Complement E_{XV}:

$$\begin{aligned} &\langle \theta_i(t)|W_{HF}(1,t)\,|\theta_k(t)\rangle \\ &\qquad = \sum_{j}^{N} \langle 1:\theta_i(t); 2:\theta_j(t)| W_2(1,2)\,[1 - P_{\rm ex}(1,2)]\,|1:\theta_k(t); 2:\theta_j(t)\rangle \end{aligned} \tag{21}$$

We also noted in that complement E_{XV} that $W_{HF}(1,t)$ is a Hermitian operator.

It is often handy to express the Hartree-Fock potential using a partial trace:

$$\boxed{W_{HF}(1,t) = {\rm Tr}_2\Big\{ P_N(2,t)\; W_2(1,2)\,[1 - P_{\rm ex}(1,2)] \Big\}} \tag{22}$$

where P_N is the projector onto the subspace spanned by the N kets $|\theta_i(t)\rangle$:

$$P_N(2,t) = \sum_{i=1}^{N} |2:\theta_i(t)\rangle\,\langle 2:\theta_i(t)| \tag{23}$$

As we have seen before, this projector is actually nothing bu the one-particle reduced density operator $\widehat{\rho}_1$ normalized by imposing its trace to be equal to the total particle number N:

$$P_N(2,t) = \widehat{\rho}_1(2,t) \tag{24}$$

The average value of the interaction energy can then be written as:

$$\boxed{\left\langle \widehat{W}_{\rm int} \right\rangle = \frac{1}{2} {\rm Tr}_1\Big\{ \widehat{\rho}_1(1,t)\; W_{HF}(1,t) \Big\}} \tag{25}$$

3-c. Time derivative

As for the time derivative term, the function it contains can be written as:

$$\sum_{i=1}^{N} \langle 0|\, a_{\theta_N}(t)...\; a_{\theta_i}(t)...\; a_{\theta_1}(t)\; a^{\dagger}_{\theta_1}(t)...\left[\frac{\mathrm{d}}{\mathrm{d}t}a^{\dagger}_{\theta_i}(t)\right]...\; a^{\dagger}_{\theta_N}(t)\,|0\rangle \tag{26}$$

In this summation, all terms involving the individual states j other than the state i (which is undergoing the derivation) lead to an expression of the type:

$$\langle 0|\, a_{\theta_j}(t)a^{\dagger}_{\theta_j}(t)\,|0\rangle \tag{27}$$

which equals 1 since this expression is the square of the norm of the state $a^{\dagger}_{\theta_j}(t)\,|0\rangle$, which is simply the Fock state $|n_j = 1\rangle$. As for the state i, it leads to a factor written in the form of a scalar product in the one-particle state space:

$$\langle 0|\, a_{\theta_i}(t)\frac{\mathrm{d}}{\mathrm{d}t}a^{\dagger}_{\theta_i}(t)\,|0\rangle = \langle \theta_i(t)|\,\frac{\mathrm{d}}{\mathrm{d}t}\,\theta_i(t)\rangle \tag{28}$$

3-d. Functional value

Regrouping all these results, we can write the value of the functional S in the form:

$$\begin{aligned} S\left[\widetilde{|\Psi(t)\rangle}\right] = \sum_{i=1}^{N}\int_{t_0}^{t_1}\mathrm{d}t\;\Bigg\{ & i\hbar\;\langle\theta_i(t)|\,\frac{\mathrm{d}}{\mathrm{d}t}\,\theta_i(t)\rangle - \langle\theta_i(t)|\left[\frac{\mathbf{P}^2}{2m}+V_1(t)\right]|\theta_i(t)\rangle \\ & - \frac{1}{2}\sum_{j=1}^{N}\langle 1:\theta_i(t);2:\theta_j(t)|\,W_2(1,2)\,[1-P_{\mathrm{ex}}(1,2)]\,|1:\theta_i(t);2:\theta_j(t)\rangle\Bigg\} \end{aligned} \tag{29}$$

4. Equations of motion

We now vary the ket $|\theta_k(t)\rangle$ according to:

$$|\theta_k(t)\rangle \Rightarrow |\theta_k(t)\rangle + |\delta\theta_k(t)\rangle \qquad \text{with} \qquad k \leq N \tag{30}$$

As in complement $\mathrm{E_{XV}}$, we will only consider variations $|\delta\theta_k(t)\rangle$ that lead to an actual variation of the ket $\widetilde{|\Psi(t)\rangle}$; those where $|\delta\theta_k(t)\rangle$ is proportional to one of the occupied states $|\theta_l(t)\rangle$ with $l \leq N$ yield no change for $\widetilde{|\Psi(t)\rangle}$ (or at the most to a phase change) and are thus irrelevant for the value of S. As we did in relations (32) or (69) of Complement $\mathrm{E_{XV}}$, we assume that:

$$|\delta\theta_k(t)\rangle = \delta f(t)e^{i\chi}\,|\theta_l(t)\rangle \qquad \text{with} \qquad l > N \tag{31}$$

where $\delta f(t)$ is an infinitesimal time-dependent function.

The computation is then almost identical to that of § 2-b in Complement $\mathrm{E_{XV}}$. When $|\theta_k(t)\rangle$ varies according to (31), all the other occupied states remaining constant, the only changes in the first line of (29) come from the terms $i = k$. In the second line, the changes come from either the $i = k$ terms, or the $j = k$ terms. As the $W_2(1,2)$

operator is symmetric with respect to the two particles, these variations are the same and their sum cancels the 1/2 factor. All these variations involve terms containing either the ket $e^{i\chi}\,|\delta\theta_k(t)\rangle$, or the bra $\langle\delta\theta_k(t)|\,e^{-i\chi}$. Now their sum must be zero for any value of χ, and this is only possible if each of the terms is zero. Inserting the variation (31) of $|\theta_k(t)\rangle$, and canceling the term in $e^{-i\chi}$ leads to the following equality:

$$\int_{t_0}^{t_1} \mathrm{d}t\ \delta f(t) \left\{ i\hbar \langle\theta_l(t)| \frac{\mathrm{d}}{\mathrm{d}t}\ \theta_k(t)\rangle - \langle\theta_l(t)| \left[\frac{\mathbf{P}^2}{2m} + V_1(t)\right] |\theta_k(t)\rangle \right.$$
$$\left. - \sum_{j=1}^{N} \langle 1:\theta_l(t); 2:\theta_j(t)|\, W_2(1,2)\,[1 - P_{\mathrm{ex}}(1,2)]\, |1:\theta_k(t); 2:\theta_j(t)\rangle \right\} = 0 \tag{32}$$

As we recognize in the function to be integrated the Hartree-Fock potential operator $W_{HF}(1,t)$ defined in (21), we can write:

$$\int_{t_0}^{t_1} \mathrm{d}t\ \delta f(t) \left\{ \langle\theta_l(t)| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - \frac{\mathbf{P}^2}{2m} - V_1(t) - W_{HF}(t)\right] |\theta_k(t)\rangle \right\} = 0 \tag{33}$$

with $l > N$.

4-a. Time-dependent Hartree-Fock equations

As the choice of the function $\delta f(t)$ is arbitrary, for expression (33) to be zero for any $\delta f(t)$ requires the function inside the curly brackets to be zero at all times t. Stationarity therefore requires the ket:

$$\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - \frac{\mathbf{P}^2}{2m} - V_1(t) - W_{HF}(t)\right] |\theta_k(t)\rangle \tag{34}$$

to have no components on any of the non-occupied states $|\theta_l(t)\rangle$ with $(l > N)$. In other words, stationarity will be obtained if, for all values of k between 1 and N, we have:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\theta_k(t)\rangle = \left[\frac{\mathbf{P}^2}{2m} + V_1(t) + W_{HF}(t)\right] |\theta_k(t)\rangle + |\xi_k(t)\rangle \tag{35}$$

where $|\xi_k(t)\rangle$ is any linear combination of the occupied states $|\theta_l(t)\rangle$ $(l \leq N)$. As we pointed out at the beginning of § 4, adding to one of the $|\theta_k(t)\rangle$ a component on the already occupied individual states has no effect on the N-particle state (aside from an eventual change of phase), and therefore does not change the value of S; consequently, the stationarity of this functional does not depend on the value of the ket $|\xi_k(t)\rangle$, which can be any ket, for example the zero ket.

Finally, if the $|\theta_n(t)\rangle$ are equal to the solutions $|\varphi_n(t)\rangle$ of the N equations:

$$\boxed{i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\varphi_n(t)\rangle = \left[\frac{\mathbf{P}^2}{2m} + V_1(t) + W_{HF}(t)\right] |\varphi_n(t)\rangle} \tag{36}$$

the functional S is indeed stationary for all times. Furthermore, as we saw in Complement E_{XV} that $W_{HF}(t)$ is Hermitian, so is the operator on the right-hand side of (36).

Consequently, the N kets $|\varphi_n(t)\rangle$ follow an evolution similar to the usual Schrödinger evolution, described by a unitary evolution operator (Complement $\mathrm{F_{III}}$). Such an operator does not change either the norm nor the scalar products of the kets: if the kets $|\varphi_n(t)\rangle$ initially formed an orthonormal set, this remains true at any later time. The whole calculation just presented is thus consistent; in particular, the norm of the N-particle state vector $|\widetilde{\Psi(t)}\rangle$ is constant over time.

Relations (36) are the time-dependent Hartree-Fock equations. Introducing the one-particle mean field operator allowed us not only to compute the stationary energy levels, but also to treat time-dependent problems.

4-b. Particles in a single spin state

Let us return to the particular case of fermions all having the same spin state, as in § 1 of Complement $\mathrm{E_{XV}}$. We can then write the Hartree-Fock equations in terms of the wave functions as:

$$i\hbar\frac{\partial}{\partial t}\,\varphi_n(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + V_{\mathrm{dir}}(\mathbf{r};t)\right]\varphi_n(\mathbf{r},t) - \int \mathrm{d}^3r'\; V_{\mathrm{ex}}(\mathbf{r},\mathbf{r}';t)\varphi_n(\mathbf{r}',t) \tag{37}$$

using definitions of (46) of that complement for the direct and exchange potentials, which are now time-dependent. There is obviously a close relation between the Hartree-Fock equations, whether they are time-dependent or not.

4-c. Discussion

As encountered in the search for a ground state with the time-independent Hartree-Fock equations, there is a strong similarity between equations (36) and an ordinary Schrödinger equation for a single particle. Here again, an exact solution of these equations is generally not possible, and we must use successive approximations. Assume for example that the external time-dependent potential $V_1(t)$ is zero until time t_0 and that for $t < t_0$, the physical system is in a stationary state. With the time-independent Hartree-Fock method we can compute an approximate value for this state and hence a series of initial values for the individual states $|\varphi_n(t_0)\rangle$. This determines the initial Hartree-Fock potential. Between time t_0 and a slightly later time $t_0 + \Delta t$, the evolution equation (36) describes the effect of the external potential $V_1(t)$ on the individual kets, and allows obtaining the $|\varphi_n(t_0 + \Delta t)\rangle$. We can then compute a new value for the Hartree-Fock potential, and use it to extend the computation of the evolution of the $|\varphi_n(t)\rangle$ until a later time $t_0 + 2\Delta t$. Proceeding step by step, we can obtain this evolution until the final time t_1. For the approach to be precise, Δt must be small enough for the Hartree-Fock potential to change only slightly from one time step to another.

Another possibility is to proceed as in the search for the stationary states. We start from a first family of orthonormal kets, now time-dependent, and which are not too far from the expected solution over the entire time interval; we then try to improve it by successive iterations. Inserting in (21) the first series of orthonormal trial functions, we get a first approximation of the Hartree-Fock potential and its associated dynamics. We then solve the corresponding equation of motion, with the same initial conditions at

$t = t_0$, which yields a new series of orthonormal functions. Using again (21), we get a value for the Hartree-Fock potential, a priori different from the previous one. We start the same procedure anew until an acceptable convergence is obtained.

Applications of this method are quite numerous, in particular in atomic, molecular, and nuclear physics. They allow, for example, the study of the electronic cloud oscillations in an atom, a molecule or a solid, placed in an external time-dependent electric field (dynamic polarisability), or the oscillations of nucleons in their nucleus. We mentioned in the conclusion of Complement E_{XV} that the time-independent Hartree-Fock method is sometimes replaced by the functional density method; this is also the case when dealing with time-dependent problems.

In concluding this complement we underline the close analogy between the Hartree-Fock theory and a time-independent or a time-dependent mean field theory. In both cases the same Hartree-Fock potential operators come into play. Even though they are the result of an approximation, these operators have a very large range of applicability.

Complement G_{XV}

Fermions or Bosons: Mean field thermal equilibrium

1 Variational principle 1712
- 1-a Notation, statement of the problem 1712
- 1-b A useful inequality 1713
- 1-c Minimization of the thermodynamic potential 1715

2 Approximation for the equilibrium density operator 1716
- 2-a Trial density operators 1716
- 2-b Partition function, distributions 1717
- 2-c Variational grand potential 1721
- 2-d Optimization 1722

3 Temperature dependent mean field equations 1725
- 3-a Form of the equations 1726
- 3-b Properties and limits of the equations 1727
- 3-c Differences with the zero-temperature Hartree-Fock equations (fermions) 1728
- 3-d Zero-temperature limit (fermions) 1729
- 3-e Wave function equations 1729

Understanding the thermal equilibrium of a system of interacting identical particles is important for many physical problems: conductor or semiconductor electronic properties, liquid Helium or ultra-cold gas properties, etc. It is also essential for studying phase transitions, various and multiple examples of which occur in solid and liquid physics: spontaneous magnetism appearing below a certain temperature, changes in electrical conduction, and many others. However, even if the Hamiltonian of a system of identical particles is known, calculation of the equilibrium properties cannot, in general, be carried to completion: these calculations present real difficulties in the handling of state vectors and interaction operators, where non-trivial combinations of creation and annihilation operators occur. One must therefore use one or several approximations. The most common one is probably the mean field approximation, which, as we saw in Complement E_{XV}, is the base of the Hartree-Fock method. In that complement, we showed, in terms of state vectors, how this method could be used to obtain approximate values for the energy levels of a system of interacting particles. As we consider here the more complex problem of thermal equilibrium, which must be treated in terms of density operators, we show how the Hartree-Fock method can be extended to this more general case.

We are going to see that, thanks to this approach, one can obtain compact formulas for an approximate value of the density operator at thermal equilibrium, in the framework of the grand canonical ensemble. The equations to be solved are fairly similar[1] to those of Complement E_{XV}. The Hartree-Fock method also gives a value of the

[1] They are not simply the juxtaposition of that complement's equations: one could imagine writing

thermodynamic grand potential, which leads directly to the pressure of the system. The other thermodynamic quantities can then be obtained via partial derivatives with respect to the equilibrium parameters (volume, temperature, chemical potential, eventually external applied field, etc. – see Appendix VI). It is clearly a powerful method even though it still is an approximation as the particles interactions are treated via a mean field approach where certain correlations are not taken into account. Furthermore, for bosons, it can only be applied to physical systems far from Bose-Einstein condensation; the reasons for this limitation will be discussed in detail in § 4-a of Complement H_{XV}.

Once we have recalled the notation and a few generalities, we shall establish (§ 1) a variational principle that applies to any density operator. It will allow us to search in any family of operators for the one closest to the density operator at thermal equilibrium. We will then introduce (§ 2) a family of trial density operators whose form reflects the mean field approximation; the variational principle will help us determine the optimal operator. We shall obtain Hartree-Fock equations for a non-zero temperature, and study some of their properties in the last section (§ 3). Several applications of these equations will be presented in Complement H_{XV}.

The general idea and the structure of the computations will be the same as in Complement E_{XV}, and we keep the same notation: we establish a variational condition, choose a trial family, and then optimize the system description within this family. This is why, although the present complement is self-contained, it might be useful to first read Complement E_{XV}.

1. Variational principle

In order to use a certain number of general results of quantum statistical mechanics (see Appendix refappend-6 for a more detailed review), we first introduce the notation.

1-a. Notation, statement of the problem

We assume the Hamiltonian is of the form:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\text{ext}} + \widehat{W}_{\text{int}} \tag{1}$$

which is the sum of the particles' kinetic energy $\widehat{H}_0$, their coupling energy $\widehat{V}_{\text{ext}}$ with an external potential:

$$\widehat{V}_{\text{ext}} = \sum_q V_1(q) \tag{2}$$

and their mutual interaction $\widehat{W}_{\text{int}}$, which can be expressed as:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{q \neq q'} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{3}$$

We are going to use the "grand canonical" ensemble (Appendix VI, § 1-c), where the particle number is not fixed, but takes on an average value determined by the chemical

those equations independently for each energy level, and then performing a thermal average. We are going to see (for example in § 2-d-β) that the determination of each level's position already implies thermal averages, meaning that the levels are coupled.

potential μ. In this case, the density operator ρ is an operator acting in the entire Fock space $\mathcal{E}_F$ (where N can take on all the possible values), and not only in the state space $\mathcal{E}_N$ for N particles (which is is more restricted since it corresponds to a fixed value of N). We set, as usual:

$$\beta = \frac{1}{k_B T} \tag{4}$$

where k_B is the Boltzmann constant and T the absolute temperature. At the grand canonical equilibrium, the system density operator depends on two parameters, β and the chemical potential μ, and can be written as:

$$\rho_{\text{eq}} = \frac{1}{Z} \exp\left[-\beta\left(\widehat{H} - \mu\widehat{N}\right)\right] \tag{5}$$

with the relation that comes from normalizing to 1 the trace of ρ_{eq}:

$$Z = \text{Tr}\left\{\exp\left[-\beta\left(\widehat{H} - \mu\widehat{N}\right)\right]\right\} \tag{6}$$

The function Z is called the "grand canonical partition function" (see Appendix VI, § 1-c). The operator $\widehat{N}$ associated with the total particle number is defined in (B-17) of Chapter XV. The temperature T and the chemical potential μ are two intensive quantities, respectively conjugate to the energy and the particle number.

Because of the particle interactions, these formulas generally lead to calculations too complex to be carried to completion. We therefore look, in this complement, for approximate expressions of ρ_{eq} and Z that are easier to use and are based on the mean field approximation.

1-b. A useful inequality

Consider two density operators ρ and ρ', both having a trace equal to 1:

$$\text{Tr}\{\rho\} = \text{Tr}\{\rho'\} = 1 \tag{7}$$

As we now show, the following relation is always true:

$$\text{Tr}\{\rho \ln\rho\} \geq \text{Tr}\{\rho \ln\rho'\} \tag{8}$$

We first note that the function $x \ln x$, defined for $x \geq 0$, is always larger than the function $x - 1$, which is the equation of its tangent at $x = 1$ (Fig. 1). For positive values of x and y we therefore always have:

$$\frac{x}{y} \ln\frac{x}{y} \geq \frac{x}{y} - 1 \tag{9}$$

or, after multiplying by y:

$$x \ln x - x \ln y \geq x - y \tag{10}$$

the equality occurring only if $x = y$.

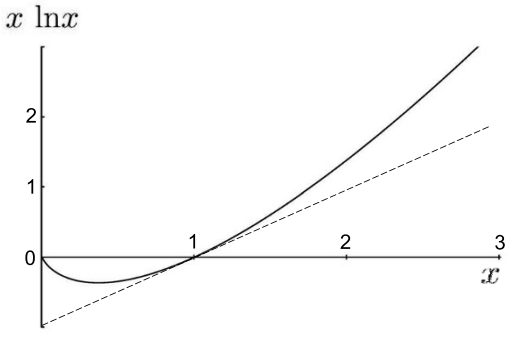


Figure 1: Plot of the function x lnx. At $x = 1$, this curve is tangent to the line $y = x - 1$ (dashed line) but always remains above it; the function value is thus always larger than $x - 1$.

Let us call p_n the eigenvalues of ρ corresponding to the normalized eigenvectors $|u_n\rangle$, and p'_m the eigenvalues of ρ' corresponding to the normalized eigenvectors $|v_m\rangle$. Used for the positive numbers p_n and p'_m, relation (10) yields:

$$p_n \ln p_n - p_n \ln p'_m \geq p_n - p'_m \tag{11}$$

We now multiply this relation by the square of the modulus of the scalar product:

$$|\langle u_n|\, v_m\rangle|^2 = \langle u_n|\, v_m\rangle\, \langle v_m|\, u_n\rangle \tag{12}$$

and sum over m and n. For the term in $p_n \ln p_n$ of (11), the summation over m yields in (12) the identity operator expanded on the basis $\{|v_m\rangle\}$; we then get $\langle u_n|\, u_n\rangle = 1$, and are left with the sum over n of $p_n \ln p_n$, that is the trace $\mathrm{Tr}\{\rho \ln \rho\}$. As for the term in $p_n \ln p'_m$, the summation over m introduces:

$$\sum_m \ln p'_m \;\; |v_m\rangle\, \langle v_m| = \ln \rho' \tag{13}$$

and we get:

$$\sum_n \langle u_n|\, p_n \ln \rho' \,|u_n\rangle = \mathrm{Tr}\,\{\rho \ln \rho'\} \tag{14}$$

As for the terms on the right-hand side of inequality (11), the term in p_n yields:

$$\sum_{m,n} \langle v_m|\, u_n\rangle\, p_n\, \langle u_n|\, v_m\rangle = \sum_m \langle v_m|\, \rho\, |v_m\rangle = \mathrm{Tr}\,\{\rho\} = 1 \tag{15}$$

and the one in p'_m also yields 1 for the same reasons, and both terms cancel out. We finally obtain the inequality:

$$\mathrm{Tr}\,\{\rho\ \ln\rho\} - \mathrm{Tr}\,\{\rho\ \ln\rho'\} \geq 0 \tag{16}$$

which proves (8).

Comment:

One may wonder under which conditions the above relation becomes an equality. This requires the inequality (11) to become an equality, which means $p_n = p'_m$ whenever the scalar product (12) is non-zero; consequently all the eigenvalues of the two operators ρ and ρ' must be equal. In addition, the eigenvectors of each operator corresponding to different eigenvalues must be orthogonal (their scalar product must be zero). In other words, the eigenvalues and the subspaces spanned by their eigenvectors are identical, which amounts to saying that $\rho = \rho'$.

1-c. Minimization of the thermodynamic potential

The entropy S associated with any density operator ρ having a trace equal to 1 is defined by relation (6) of Appendix VI:

$$S = -k_B\ \mathrm{Tr}\,\{\rho\ \ln\rho\} \tag{17}$$

The thermodynamic potential of the grand canonical ensemble is defined by the "grand potential" Φ, which can be expressed as a function of ρ by relation (Appendix VI, § 1-c-β):

$$\Phi = \left\langle \widehat{H} \right\rangle - TS - \mu \left\langle \widehat{N} \right\rangle = \mathrm{Tr}\left\{ \left[\widehat{H} + k_B T\ \ln\rho - \mu\widehat{N} \right] \rho \right\} \tag{18}$$

Inserting (5) into (18), we see that the value of Φ at equilibrium, Φ_{eq}, can be directly obtained from the partition function Z:

$$\begin{aligned} \Phi_{\mathrm{eq}} &= \mathrm{Tr}\left\{ \left[\widehat{H} + \beta^{-1} \left(-\beta\widehat{H} + \beta\mu\widehat{N} - \ln Z \right) - \mu\widehat{N} \right] \rho_{\mathrm{eq}} \right\} \\ &= -\beta^{-1} \ln Z\ \mathrm{Tr}\,\{\rho_{\mathrm{eq}}\} = -k_B T\ \ln Z \end{aligned} \tag{19}$$

We therefore have:

$$Z = e^{-\beta\Phi_{eq}} \qquad \text{or} \qquad \Phi_{eq} = -k_B T\ \ln Z \tag{20}$$

Consider now any density operator ρ and its associated function Φ obtained from (18). According to (5) and (20), we can write:

$$-\beta \left(\widehat{H} - \mu\widehat{N} \right) = \ln\rho_{eq} + \ln Z = \ln\rho_{eq} - \beta\Phi_{eq} \tag{21}$$

Inserting this result in (18) yields:

$$\Phi = \mathrm{Tr}\left\{ \beta^{-1} \left[-\ln\rho_{eq} + \beta\Phi_{eq} + \ln\rho \right] \rho \right\} = \beta^{-1} \mathrm{Tr}\left\{ \left[-\ln\rho_{eq} + \ln\rho \right] \rho \right\} + \Phi_{eq} \tag{22}$$

Now relation (16), used with $\rho' = \rho_{\mathrm{eq}}$, is written as:

$$\mathrm{Tr}\left\{ \left[\ln\rho - \ln\rho_{\mathrm{eq}} \right] \rho \right\} \geq 0 \tag{23}$$

Relation (22) thus implies that for any density operator ρ having a trace equal to 1, we have:

$$\Phi \geq \Phi_{\text{eq}} \tag{24}$$

the equality occurring if, and only if, $\rho = \rho_{\text{eq}}$.

Relation (24) can be used to fix a variational principle: choosing a family of density operators ρ having a trace equal to 1, we try to identify in this family the operator that yields the lowest value for Φ. This operator will then be the optimal operator within this family. Furthermore, this operator yields an upper value for the grand potential, with an error of second order with respect to the error made on ρ.

2. Approximation for the equilibrium density operator

We now use this variational principle with a family of density operators that leads to manageable calculations.

2-a. Trial density operators

The Hartree-Fock method is based on the assumption that a good approximation is to consider that each particle is independent of the others, but moving in the mean potential they create. We therefore compute an approximate value of the density operator by replacing the Hamiltonian $\widehat{H}$ by a sum of independent particles' Hamiltonians $\widetilde{h}(q)$:

$$\widetilde{H} = \sum_{q=1}^{N} \widetilde{h}(q) \tag{25}$$

We now introduce the basis of the creation and annihilation operators, associated with the eigenvectors of the one-particle operator $\widetilde{h}$:

$$a_k^\dagger |0\rangle = \widetilde{|\theta_k\rangle} \qquad \text{with} \quad \widetilde{h}\widetilde{|\theta_k\rangle} = \widetilde{e}_k \widetilde{|\theta_k\rangle} \tag{26}$$

The symmetric one-particle operator $\widetilde{H}$ can then be written, according to relation (B-14) of Chapter XV:

$$\widetilde{H} = \sum_k \widetilde{e}_k \; a_k^\dagger a_k \tag{27}$$

where the real constants $\widetilde{e}_k$ are the eigenvalues of the operator $\widetilde{h}$.

We choose as trial operators acting in the Fock space the set of operators $\widetilde{\rho}$ that can be written in the form corresponding to an equilibrium in the grand canonical ensemble – see relation (42) of Appendix VI. We then set:

$$\widetilde{\rho} = \frac{1}{\widetilde{Z}} \exp\left[-\beta\left(\widetilde{H} - \mu\widehat{N}\right)\right] \tag{28}$$

where $\widetilde{H}$ is any symmetric one-particle operator, the constant β the inverse of the temperature defined in (4), μ a real constant playing the role of a chemical potential, and $\widetilde{Z}$ the trace of $\widetilde{\rho}$:

$$\widetilde{Z} = \text{Tr}\left\{\exp\left[-\beta\left(\widetilde{H} - \mu\widehat{N}\right)\right]\right\} \tag{29}$$

Consequently, the relevant variables in our problem are the states $\widetilde{|\theta_k\rangle}$, which form an arbitrary orthonormal basis in the individual state space, and the energies $\widetilde{e}_k$. These variables determine the a_k as well as $\widetilde{H}$, and we have to find which of their values minimizes the function:

$$\widetilde{\Phi} = \mathrm{Tr}\left\{\left[\widehat{H} + k_B T\, \ln\widetilde{\rho} - \mu\widehat{N}\right]\widetilde{\rho}\right\} \tag{30}$$

Taking (27) and (28) into account, we can write:

$$\widetilde{\rho} = \frac{1}{\widetilde{Z}} \exp\left[-\sum_k \beta\left(\widetilde{e}_k - \mu\right) a_k^\dagger a_k\right] \tag{31}$$

The following computations are simplified since the Fock space can be considered to be the tensor product of independent spaces associated with the individual states $\widetilde{|\theta_k\rangle}$; consequently, the trial density operator (28) can be written as a tensor product of operators each acting on a single mode k:

$$\widetilde{\rho} = \frac{1}{\widetilde{Z}} \prod_k \exp\left[-\beta\left(\widetilde{e}_k\ - \mu\right) a_k^\dagger a_k\right] \tag{32}$$

2-b. Partition function, distributions

Equality (32) has the same form as relation (5) of Complement $\mathrm{B_{XV}}$, with a simple change: the replacement of the free particle energies $e_k = \hbar^2 k^2/2m$ by the energies $\widetilde{e}_k$, which are as yet unknown. As this change does not impact the mathematical structure of the density operator, we can directly use the results of Complement $\mathrm{B_{XV}}$.

α. Variational partition function

The function $\widetilde{Z}$ only depends on the variational energies $\widetilde{e}_k$, since the trace of (32) may be computed in the basis $\left\{\widetilde{|\theta_k\rangle}\right\}$, which yields:

$$\widetilde{Z} = \prod_k \sum_{n_k} \exp\left[-n_k \beta\left(\widetilde{e}_k - \mu\right)\right] \tag{33}$$

We simply get an expression similar to relation (7) of Complement $\mathrm{B_{XV}}$, obtained for an ideal gas. Since for fermions n_k can only take the values 0 and 1, we get:

$$\widetilde{Z}_{fermions} = \prod_k \left[1 + e^{-\beta\left(\widetilde{e}_k - \mu\right)}\right] \tag{34}$$

whereas for bosons n_k varies from 0 to infinity, so that:

$$\widetilde{Z}_{bosons} = \prod_k \frac{1}{1 - e^{-\beta\left(\widetilde{e}_k - \mu\right)}} \tag{35}$$

In both cases we can write:

$$\ln\widetilde{Z} = -\eta \sum_k \ln\left[1 - \eta e^{-\beta\left(\widetilde{e}_k - \mu\right)}\right] \tag{36}$$

with $\eta = +1$ for bosons, and $\eta = -1$ for fermions.

Computing the entropy can be done in a similar way. As the density operator $\widetilde{\rho}$ has the same form as the one describing the thermal equilibrium of an ideal gas, we can use for a system described by $\widetilde{\rho}$ the formulas obtained for the entropy of a system without interactions.

β. One particle, reduced density operator

Let us compute the average value of $a_i^\dagger a_j$ with the density operator $\widetilde{\rho}$:

$$\left\langle a_i^\dagger a_j \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}\left\{ a_i^\dagger a_j \widetilde{\rho} \right\} \tag{37}$$

We saw in § 2-c of Complement B_{XV} that:

$$\mathrm{Tr}\left\{ a_i^\dagger a_j \widetilde{\rho} \right\} = \delta_{ij}\ f_\beta(\widetilde{e}_i - \mu) \tag{38}$$

where the distribution function f_β is noted f_β^{FD} for fermions, and f_β^{BE} for bosons:

$$f_\beta(e-\mu) = \begin{cases} f_\beta^{FD}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}+1} & \text{for fermions} \\ f_\beta^{BE}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}-1} & \text{for bosons} \end{cases} \tag{39}$$

When the system is described by the density operator $\widetilde{\rho}$, the average populations of the individual states $\widetilde{|\theta_k\rangle}$ are therefore determined by the usual Fermi-Dirac or Bose-Einstein distributions. From now on, and to simplify the notation, we shall write simply $|\theta_k\rangle$ for the kets $\widetilde{|\theta_k\rangle}$.

We can introduce a "one-particle reduced density operator" $\widetilde{\rho}_1(1)$ by[2] :

$$\widetilde{\rho}_1(1) = \sum_k f_\beta(\widetilde{e}_k - \mu)\ \ |1:\theta_k\rangle\langle 1:\theta_k| \tag{40}$$

where the 1 enclosed in parentheses and the subscript 1 on the left-hand side emphasize we are dealing with an operator acting in the one-particle state space (as opposed to $\widetilde{\rho}$ that acts in the Fock space); needless to say, this subscript has nothing to do with the initial numbering of the particles, but simply refers to any single particle among all the system particles. The diagonal elements of $\widetilde{\rho}_1(1)$ are the individual state populations. With this operator, we can compute the average value over $\widetilde{\rho}$ of any one-particle operator $\widehat{M}$:

$$\left\langle \widehat{M} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}\left\{ \widehat{M}\ \widetilde{\rho} \right\} \tag{41}$$

as we now show. Using the expression (B-12) of Chapter XV for any one-particle opera-

[2] Contrary to what is usually the case for a density operator, the trace of this reduced operator is not equal to 1, but to the average particle number – see relation (44). This different normalization is often more useful when studying systems composed of a large number of particles.

tor[3], as well as (38), we can write:

$$\left\langle \widehat{M} \right\rangle_{\widetilde{\rho}} = \sum_{k,l} \langle \theta_k | \, \widehat{m} \, | \theta_l \rangle \, \mathrm{Tr} \left\{ a_k^\dagger a_l \widetilde{\rho} \right\} = \sum_k \langle \theta_k | \, \widehat{m} \, | \theta_k \rangle \;\; f_\beta(\widetilde{e}_k - \mu)$$
$$= \sum_k \langle \theta_k | \, \widehat{m} \;\, \widetilde{\rho}_1(1) \, | \theta_k \rangle \tag{42}$$

that is:

$$\left\langle \widehat{M} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}_1 \left\{ \widehat{m} \;\, \widetilde{\rho}_1(1) \right\} \tag{43}$$

As we shall see, the density operator $\widetilde{\rho}_1(1)$ is quite useful since it allows obtaining in a simple way all the average values that come into play in the Hartree-Fock computations. Our variational calculations will simply amount to varying $\widetilde{\rho}_1(1)$. This operator presents, in a certain sense, all the properties of the variational density operator $\widetilde{\rho}$ chosen in (28) in the Fock space. It plays the same role[4] as the projector P_N (which also represents the essence of the variational N-particle ket) played in Complement $\mathrm{E_{XV}}$. In a general way, one can say that the basic principle of the Hartree-Fock method is to reduce the binary correlation functions of the system to products of single-particle correlation functions (more details on this point will be given in § 2-b of Complement $\mathrm{C_{XVI}}$).

The average value of the operator $\widehat{N}$ for the total particle number is written:

$$\left\langle \widehat{N} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr} \left\{ \widehat{N} \;\, \widetilde{\rho} \right\} = \mathrm{Tr}_1 \left\{ \widetilde{\rho}_1(1) \right\} = \sum_{i=1}^{\infty} f_\beta(\widetilde{e}_i - \mu) \tag{44}$$

Both functions f_β^{FD} and f_β^{BE} increase as a function of μ and, for any given temperature, the total particle number is controlled by the chemical potential. For a large physical system whose energy levels are very close, the orbital part of the discrete sum in (44) can be replaced by an integral. Figure 1 of Complement $\mathrm{B_{XV}}$ shows the variations of the Fermi-Dirac and Bose-Einstein distributions. We also mentioned that for a boson system, the chemical potential could not exceed the lowest value e_0 of the energies e_l; when it approaches that value, the population of the corresponding level diverges, which is the Bose-Einstein condensation phenomenon we will come back to in the next complement. For fermions, on the other hand, the chemical potential has no upper boundary, as, whatever its value, the population of states having an energy lower than μ cannot exceed 1.

γ. *Two particles, distribution functions*

We now consider an arbitrary two-particle operator $\widehat{G}$ and compute its average value with the density operator $\widetilde{\rho}$. The general expression of a symmetric two-particle

[3] We have changed the notation $\widehat{F}$ and $\widehat{f}$ of Chapter XV into $\widehat{M}$ and $\widehat{m}$ to avoid any confusion with the distribution functions f_β.

[4] For fermions, and when the temperature approaches zero, the distribution function f_β^{FD} included in the definition of $\rho_I(1)$ becomes a step function and $\rho_I(1)$ does indeed coincide with $P_N(1)$.

operator is given by relation (C-16) of Chapter XV, and we can write:

$$\begin{aligned}\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} &= \mathrm{Tr}\left\{ \widehat{G}\ \widetilde{\rho} \right\} \\ &= \frac{1}{2} \sum_{i,j,k,l} \langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2) \,|1:\theta_k; 2:\theta_l\rangle\, \mathrm{Tr}\left\{ a_i^\dagger a_j^\dagger a_l a_k\ \widetilde{\rho} \right\}\end{aligned} \tag{45}$$

We follow the same steps as in § 2-a of Complement E_{XV}: we use the mean field approximation to replace the computation of the average value of a two-particle operator by that of average values for one-particle operators. We can, for example, use relation (43) of Complement B_{XV}, which shows that:

$$\mathrm{Tr}\left\{ a_i^\dagger a_j^\dagger a_l a_k\ \widetilde{\rho} \right\} = [\delta_{ik}\delta_{jl} + \eta\ \delta_{il}\delta_{jk}]\ \ f_\beta(\widetilde{e}_i - \mu) f_\beta(\widetilde{e}_j - \mu) \tag{46}$$

We then get:

$$\begin{aligned}&\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} = \tfrac{1}{2} \textstyle\sum_{i,j} f_\beta(\widetilde{e}_i - \mu)\ f_\beta(\widetilde{e}_j - \mu) \times \\ &\quad \Big[\langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2) \,|1:\theta_i; 2:\theta_j\rangle + \eta\ \langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2) \,|1:\theta_j; 2:\theta_i\rangle \Big]\end{aligned} \tag{47}$$

Which, according to (40), can also be written as:

$$\begin{aligned}\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} = \tfrac{1}{2} \textstyle\sum_{i,j} &\langle 1:\theta_i |\, \widetilde{\rho}_1(1) \,|1:\theta_i\rangle\ \ \langle 2:\theta_j |\, \widetilde{\rho}_1(2) \,|2:\theta_j\rangle \times \\ &\langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2)\, [1 + \eta P_{\mathrm{ex}}(1,2)] \,|1:\theta_i; 2:\theta_j\rangle\end{aligned} \tag{48}$$

where P_{ex} is the exchange operator between particles 1 and 2. Since:

$$\begin{aligned}&\langle 1:\theta_i |\, \widetilde{\rho}_1(1) \,|1:\theta_i\rangle\ \ \langle 2:\theta_j |\, \widetilde{\rho}_1(2) \,|2:\theta_j\rangle \\ &\qquad = \langle 1:\theta_i; 2:\theta_j |\, \widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2) \,|1:\theta_i; 2:\theta_j\rangle\end{aligned} \tag{49}$$

and as the operators $\widetilde{\rho}_1(1)$ and $\widetilde{\rho}_1(2)$ are diagonal in the basis $|\theta_i\rangle$, we can write the right-hand side of (48) as:

$$\frac{1}{2} \sum_{i,j} \langle 1:\theta_i; 2:\theta_j |\, [\widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2)]\, \widehat{g}(1,2)\, [1 + \eta P_{\mathrm{ex}}(1,2)] \,|1:\theta_i; 2:\theta_j\rangle \tag{50}$$

which is simply a (double) trace on two particles 1 and 2. This leads to:

$$\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} = \frac{1}{2} \mathrm{Tr}_{1,2} \left\{ [\widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2)]\, \widehat{g}(1,2)\, [1 + \eta P_{\mathrm{ex}}(1,2)] \right\} \tag{51}$$

As announced above, the average value of the two-particle operator $\widehat{G}$ can be expressed, within the Hartree-Fock approximation, in terms of the one-particle reduced density operator $\widetilde{\rho}_1(1)$; this relation is not linear.

Comment:

The analogy with the computations of Complement $\mathrm{E_{XV}}$ becomes obvious if we regroup its equations (57) and (58) and write:

$$\left\langle \hat{W}_{\rm int} \right\rangle = \frac{1}{2} \mathrm{Tr}_{1,2} \left\{ [P_N(1) \otimes P_N(2)] \, W_2(1,2) \, [1 + \eta P_{\rm ex}(1,2)] \right\} \tag{52}$$

Replacing $W_2(1,2)$ by G, we get a relation very similar to (51), except for the fact that the projectors P_N must be replaced by the one-particle operators $\widetilde{\rho}_1$. In § 3-d, we shall come back to the correspondence between the zero and non-zero temperature results.

2-c. Variational grand potential

We now have to compute the grand potential $\widetilde{\Phi}$ written in (30). As the exponential form (28) for the trial operator makes it easy to compute $\ln\widetilde{\rho}$, we see that the terms in $\mu\widehat{N}$ cancel out, and we get:

$$\widetilde{\Phi} = \mathrm{Tr} \left\{ \left[\widehat{H} - \widetilde{H} \right] \widetilde{\rho} \right\} - k_B T \, \ln \widetilde{Z} \tag{53}$$

We now have to compute the average energy, with the density operator $\widetilde{\rho}$, of the difference between the Hamiltonians $\widehat{H}$ and $\widetilde{H}$ respectively defined by (1) and (25).

We first compute the trace:

$$\mathrm{Tr} \left\{ \widehat{H} \; \widetilde{\rho} \right\} = \left\langle \widehat{H} \right\rangle_{\widetilde{\rho}} \tag{54}$$

starting with the kinetic energy contribution $\widehat{H}_0$ in (1). We call K_0 the individual kinetic energy operator:

$$K_0 = \frac{\mathbf{P}^2}{2m} \tag{55}$$

(m is the particle mass). Equality (43) applied to $\widehat{H}_0$ yields the average kinetic energy when the system is described by $\widetilde{\rho}$:

$$\left\langle \widehat{H_0} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}_1 \left\{ K_0 \; \widetilde{\rho}_1(1) \right\} = \sum_i \langle \theta_i | \, K_0 \, | \theta_i \rangle \times f_\beta(\widetilde{e}_i - \mu) \tag{56}$$

This result is easily interpreted; each individual state contributes its average kinetic energy, multiplied by its population.

The computation of the average value $\left\langle \widehat{V_{\rm ext}} \right\rangle_{\widetilde{\rho}}$ follows the same steps:

$$\left\langle \widehat{V_{\rm ext}} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}_1 \left\{ V_1 \; \widetilde{\rho}_1(1) \right\} = \sum_i \langle \theta_i | \, V_1 \, | \theta_i \rangle \times f_\beta(\widetilde{e}_i - \mu) \tag{57}$$

(as in Complement $\mathrm{E_{XV}}$, operator V_1 is the one-particle external potential operator).

To complete the calculation of the average value of $\widehat{H}$, we now have to compute the trace $\mathrm{Tr}\{W_{\rm int}\widetilde{\rho}\}$, the average value of the interaction energy when the system is described by $\widetilde{\rho}$. Using relation (51) we can write this average value as a double trace:

$$\langle W_{\rm int} \rangle_{\widetilde{\rho}} = \frac{1}{2} \mathrm{Tr}_{1,2} \left\{ \widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2) \;\; [W_2(1,2)] \, [1 + \eta P_{\rm ex}(1,2)] \right\} \tag{58}$$

We now turn to the average value of $\widetilde{H}$. The calculation is simplified since $\widetilde{H}$ is, like $\widehat{H_0}$, a one-particle operator; furthermore, the $|\theta_i\rangle$ have been chosen to be the eigenvectors of $\widetilde{h}$ with eigenvalues $\widetilde{e}_i$ – see relation (26). We just replace in (56), $\widehat{H_0}$ by $\widetilde{H}$, and obtain:

$$\left\langle \widetilde{H} \right\rangle_{\widetilde{\rho}} = \sum_i \langle\theta_i| \left(\widetilde{e}_i\right) |\theta_i\rangle \times f_\beta(\widetilde{e}_i - \mu) = \sum_i \widetilde{e}_i \; f_\beta(\widetilde{e}_i - \mu) \tag{59}$$

Regrouping all these results and using relation (36), we can write the variational grand potential as the sum of three terms:

$$\widetilde{\Phi} = \widetilde{\Phi}_1 + \widetilde{\Phi}_2 + \widetilde{\Phi}_3 \tag{60}$$

with:

$$\begin{aligned}
\widetilde{\Phi}_1 &= \mathrm{Tr}_1 \Big\{ \left[K_0 + V_1\right] \; \widetilde{\rho}_1(1) \Big\} \\
\widetilde{\Phi}_2 &= \frac{1}{2} \mathrm{Tr}_{1,2} \Big\{ \left[\widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2)\right] W_2(1,2) \left[1 + \eta P_{\mathrm{ex}}(1,2)\right] \Big\} \\
\widetilde{\Phi}_3 &= \sum_i \left\{ -\widetilde{e}_i \; f_\beta(\widetilde{e}_i - \mu) + \eta k_B T \, \ln \left[1 - \eta e^{-\beta\left(\widetilde{e}_i - \mu\right)}\right] \right\}
\end{aligned} \tag{61}$$

2-d. Optimization

We now vary the eigenenergies $\widetilde{e}_k$ and eigenstates $|\theta_k\rangle$ of $\widetilde{H}$ to find the value of the density operator $\widetilde{\rho}$ that minimizes the average value $\widetilde{\Phi}$ of the potential. We start with the variations of the eigenstates, which induce no variation of $\widetilde{\Phi}_3$. The computation is actually very similar to that of Complement $\mathrm{E_{XV}}$, with the same steps: variation of the eigenvectors, followed by the demonstration that the stationarity condition is equivalent to a series of eigenvalue equations for a Hartree-Fock operator (a one-particle operator). Nevertheless, we will carry out this computation in detail, as there are some differences. In particular, and contrary to what happened in Complement $\mathrm{E_{XV}}$, the number of states $|\theta_i\rangle$ to be varied is no longer fixed by the particle number N; these states form a complete basis of the individual state space, and their number can go to infinity. This means that we can no longer give to one (or several) state(s) a variation orthogonal to all the other $|\theta_j\rangle$; this variation will necessarily be a linear combination of these states. In a second step, we shall vary the energies $\widetilde{e}_k$.

α. Variations of the eigenstates

As the eigenstates $|\theta_i\rangle$ vary, they must still obey the orthogonality relations:

$$\langle\theta_i \, |\theta_j\rangle = \delta_{ij} \tag{62}$$

The simplest idea would be to vary only one of them, $|\theta_l\rangle$ for example, and make the change:

$$|\theta_l\rangle \Rightarrow |\theta_l\rangle + |\mathrm{d}\theta_l\rangle \tag{63}$$

The orthogonality conditions would then require:

$$\langle\theta_i \, |\mathrm{d}\theta_l\rangle = 0 \qquad \text{for all } i \neq l \tag{64}$$

preventing $|\mathrm{d}\theta_l\rangle$ from having a component on any ket $|\theta_i\rangle$ other than $|\theta_l\rangle$: in other words, $|\mathrm{d}\theta_l\rangle$ and $|\theta_l\rangle$ would be colinear. As $|\theta_l\rangle$ must remain normalized, the only possible variation would thus be a phase change, which does not affect either the density operator $\widetilde{\rho}_1(1)$ or any average values computed with $\widetilde{\rho}$. This variation does not change anything and is therefore irrelevant.

It is actually more interesting to vary simultaneously two eigenvectors, which will be called $|\theta_l\rangle$ and $|\theta_m\rangle$, as it is now possible to give $|\theta_l\rangle$ a component on $|\theta_m\rangle$, and the reverse. This does not change the two-dimensional subspace spanned by these two states; hence their orthogonality with all the other basis vectors is automatically preserved. Let us give the two vectors the following infinitesimal variations (without changing their energies $\widetilde{e}_l$ and $\widetilde{e}_m$):

$$\left\{ \begin{array}{rcl} |\mathrm{d}\theta_l\rangle & = & \mathrm{d}\alpha\ e^{i\chi}\ |\theta_m\rangle \\ |\mathrm{d}\theta_m\rangle & = & -\mathrm{d}\alpha\ e^{-i\chi}\ |\theta_l\rangle \end{array} \right. \quad \text{with } l \neq m \tag{65}$$

where $\mathrm{d}\alpha$ is an infinitesimal real number and χ an arbitrary but fixed real number. For any value of χ, we can check that the variation of $\langle\theta_l\ |\theta_l\rangle$ is indeed zero (it contains the scalar products $\langle\theta_l\ |\theta_m\rangle$ or $\langle\theta_m\ |\theta_l\rangle$ which are zero), as is the symmetrical variation of $\langle\theta_m\ |\theta_m\rangle$, and that we have:

$$\mathrm{d}\,\langle\theta_l\ |\theta_m\rangle = \mathrm{d}\alpha\ e^{-i\chi}\,\langle\theta_m\ |\theta_m\rangle - \mathrm{d}\alpha\ e^{-i\chi}\,\langle\theta_l\ |\theta_l\rangle = 0 \tag{66}$$

The variations (65) are therefore acceptable, for any real value of χ.

We now compute how they change the operator $\widetilde{\rho}_1(1)$ defined in (40). In the sum over k, only the $k = l$ and $k = m$ terms will change. The $k = l$ term yields a variation:

$$\mathrm{d}\alpha\ f_\beta(\widetilde{e}_l - \mu)\left[e^{i\chi}\,|\theta_m\rangle\,\langle\theta_l| + e^{-i\chi}\,|\theta_l\rangle\,\langle\theta_m|\right] \tag{67}$$

whereas the $k = m$ term yields a similar variation but where $f_\beta(\widetilde{e}_l - \mu)$ is replaced by $-f_\beta(\widetilde{e}_m - \mu)$. This leads to:

$$\mathrm{d}\widetilde{\rho}_1(1) = \mathrm{d}\alpha\ \left[f_\beta(\widetilde{e}_l - \mu) - f_\beta(\widetilde{e}_m - \mu)\right]\left[e^{i\chi}\,|\theta_m\rangle\,\langle\theta_l| + e^{-i\chi}\,|\theta_l\rangle\,\langle\theta_m|\right] \tag{68}$$

We now include these variations in the three terms of (61); as the distributions f are unchanged, only the terms $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$ will vary. The infinitesimal variation of $\widetilde{\Phi}_1$ is written as:

$$\mathrm{d}\widetilde{\Phi}_1 = \mathrm{Tr}_1\left\{[K_0 + V_1]\ \ \mathrm{d}\widetilde{\rho}_1(1)\right\} \tag{69}$$

As for $\mathrm{d}\widetilde{\Phi}_2$, it contains two contributions, one from $\mathrm{d}\widetilde{\rho}_1(1)$ and one from $\mathrm{d}\widetilde{\rho}_1(2)$. These two contributions are equal since the operator $W_2(1,2)$ is symmetric (particles 1 and 2 play an equivalent role). The factor $1/2$ in $\widetilde{\Phi}_2$ disappears and we get:

$$\mathrm{d}\widetilde{\Phi}_2 = \mathrm{Tr}_{1,2}\Big\{\mathrm{d}\widetilde{\rho}_1(1)\otimes\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1 + \eta P_{\mathrm{ex}}(1,2)]\Big\} \tag{70}$$

We can regroup these two contributions, using the fact that for any operator $O(1,2)$, it can be shown that:

$$\mathrm{Tr}_{1,2}\left\{\mathrm{d}\widetilde{\rho}_1(1)\ O(1,2)\right\} = \mathrm{Tr}_1\Big\{\mathrm{d}\widetilde{\rho}_1(1)\ \mathrm{Tr}_2\left\{O(1,2)\right\}\Big\} \tag{71}$$

This equality is simply demonstrated[5] by using the definition of the partial trace $\mathrm{Tr}_2\{O(1,2)\}$ of operator $O(1,2)$ with respect to particle 2. We then get:

$$\begin{aligned} \mathrm{d}\widetilde{\Phi} &= \mathrm{d}\widetilde{\Phi}_1 + d\widetilde{\Phi}_2 \\ &= \mathrm{Tr}_1\Big\{ \mathrm{d}\widetilde{\rho}_1(1) \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] \Big\} \end{aligned} \tag{72}$$

Inserting now the expression (68) for $d\widetilde{\rho}_1(1)$, we get two terms, one proportional to $e^{i\chi}$, another one to $e^{-i\chi}$, whose value is:

$$\begin{aligned} &\mathrm{d}\alpha\ e^{-i\chi}\ \ [f_\beta(\widetilde{e}_l-\mu) - f_\beta(\widetilde{e}_m-\mu)] \\ &\quad \times \mathrm{Tr}_1\Big\{ |\theta_l\rangle\langle\theta_m| \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] \Big\} \end{aligned} \tag{73}$$

Now, for any operator $O(1)$, we can write:

$$\begin{aligned} \mathrm{Tr}_1\{|\theta_l\rangle\langle\theta_m| O(1)\} &= \sum_i \langle\theta_i\ |\theta_l\rangle\langle\theta_m| O(1)|\theta_i\rangle = \sum_i \langle\theta_m| O(1)|\theta_i\rangle\langle\theta_i\ |\theta_l\rangle \\ &= \langle\theta_m| O(1)|\theta_l\rangle \end{aligned} \tag{74}$$

so that the variation (73) can be expressed as:

$$\begin{aligned} &\mathrm{d}\alpha\ e^{-i\chi}\ \ [f_\beta(\widetilde{e}_l-\mu) - f_\beta(\widetilde{e}_m-\mu)] \\ &\qquad \langle\theta_m| \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] |\theta_l\rangle \end{aligned} \tag{75}$$

The term in $e^{i\chi}$ has a similar form, but it does not have to be computed for the following reason. The variation $\mathrm{d}\widetilde{\Phi}$ is the sum of a term in $e^{i\chi}$ and another in $e^{-i\chi}$:

$$\mathrm{d}\widetilde{\Phi} = \mathrm{d}\alpha \left[c_1 e^{-i\chi} + c_2 e^{i\chi} \right] \tag{76}$$

and the stationarity condition requires $\mathrm{d}\widetilde{\Phi}$ to be zero for any choice of χ. Choosing $\chi = 0$, yields $c_1 + c_2 = 0$; choosing $\chi = \pi/2$, and multiplying by $-i$, we get $c_1 - c_2 = 0$. Adding and subtracting those two relations shows that both coefficients c_1 and c_2 must be zero. Consequently, it suffices to impose the terms in $e^{\pm i\chi}$, and hence expression (75), to be zero. When $\widetilde{e}_l \neq \widetilde{e}_m$, the distribution functions f_β are not equal, and we get:

$$\langle\theta_m| \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] |\theta_l\rangle = 0 \tag{77}$$

(if $\widetilde{e}_l = \widetilde{e}_m$, however, we have not yet obtained any particular condition to be satisfied[6]).

[5] The definition of partial traces is given in § 5-b of Complement $\mathrm{E}_{\mathrm{III}}$. The left hand side of (71) can be written as $\sum_{i,j} \langle 1:\theta_i; 2:\theta_j| d\rho_I(1) O(1,2) |1:\theta_i; 2:\theta_j\rangle$. We then insert, after $d\rho_I(1)$, a closure relation on the kets $|1:\theta_k; 2:\theta_{k'}\rangle$, with $k' = j$ since $d\rho_I(1)$ does not act on particle 2. This yields: $\sum_{i,k} \langle 1:\theta_i| d\rho_I(1) |1:\theta_k\rangle \sum_j \langle 1:\theta_k; 2:\theta_j| O(1,2) |1:\theta_i; 2:\theta_j\rangle$, where the sum over j is the definition of the matrix element between $\langle 1:\theta_k|$ and $|1:\theta_i\rangle$ of the partial trace over particle 2 of the operator $O(1,2)$. We then get the right-hand side of (71).

[6] This was expected, since this choice does not lead to any variation of the trial density operator.

β. Variation of the energies

Let us now see what happens if the energy $\widetilde{e}_l$ varies by $\mathrm{d}\widetilde{e}_l$. The function $f_\beta(\widetilde{e}_l-\mu)$ then varies by $\mathrm{d}f^l_\beta$ which, according to relation (40), induces a variation of $\widetilde{\rho}_1$:

$$\mathrm{d}\widetilde{\rho}_1 = |\theta_l\rangle\,\langle\theta_l|\,\mathrm{d}f^l_\beta \tag{78}$$

and thus leads to variations of expressions (61) of $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$. Their sum is:

$$\mathrm{d}\widetilde{\Phi}_1 + \mathrm{d}\widetilde{\Phi}_2 = \mathrm{Tr}_1\Big\{\mathrm{d}\widetilde{\rho}_1(1)\Big[K_0 + V_1 + \mathrm{Tr}_2\big\{\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1+\eta P_{\mathrm{ex}}(1,2)]\ \big\}\Big]\Big\} \tag{79}$$

where the factor $1/2$ in $\widetilde{\Phi}_2$ has been canceled since the variations induced by $\widetilde{\rho}_1(1)$ and $\widetilde{\rho}_1(2)$ double each other. Inserting (78) in this relation and using again (74), we get:

$$\mathrm{d}\widetilde{\Phi}_1 + \mathrm{d}\widetilde{\Phi}_2 = \mathrm{d}f^l_\beta\,\langle\theta_l|\,\Big[K_0 + V_1 + \mathrm{Tr}_2\big\{\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1+\eta P_{\mathrm{ex}}(1,2)]\ \big\}\Big]\,|\theta_l\rangle \tag{80}$$

As for $\widetilde{\Phi}_3$, its variation is the sum of a term in $\mathrm{d}\widetilde{e}_l$ coming from the explicit presence of the energies $\widetilde{e}_i$ in its definition (61), and a term in $\mathrm{d}f^l_\beta$. If we let only the energy $\widetilde{e}_l$ vary (not taking into account the variations of the distribution function), we get a zero result, since:

$$\begin{aligned}\Bigg[-f_\beta(\widetilde{e}_l-\mu) + \eta k_B T\ \ \Big[(-\eta)\,(-\beta)\ \ e^{-\beta\left(\widetilde{e}_l-\mu\right)}\Big]\,\frac{1}{1-\eta e^{-\beta\left(\widetilde{e}_l-\mu\right)}}\Bigg]\,\mathrm{d}\widetilde{e}_l \\ = \big[-f_\beta(\widetilde{e}_l-\mu) + f_\beta(\widetilde{e}_l-\mu)\big]\mathrm{d}\widetilde{e}_l = 0\end{aligned} \tag{81}$$

Consequently, we just have to vary by $\mathrm{d}f^l_\beta$ the distribution function, and we get:

$$\mathrm{d}\widetilde{\Phi}_3 = -\widetilde{e}_l\ \mathrm{d}f^l_\beta \tag{82}$$

Finally, after simplification by $\mathrm{d}f^l_\beta$ (which, by hypothesis, is different from zero), imposing the variation $\mathrm{d}\widetilde{\Phi}$ to be zero leads to the condition:

$$\begin{aligned}&\mathrm{d}\widetilde{\Phi}_1 + \mathrm{d}\widetilde{\Phi}_2 + \mathrm{d}\widetilde{\Phi}_3 \\ &\quad = \langle\theta_l|\,[K_0 + V_1 - \widetilde{e}_l + \mathrm{Tr}_2\,\{\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1+\eta P_{\mathrm{ex}}(1,2)]\}]\,|\theta_l\rangle = 0\end{aligned} \tag{83}$$

This expression does look like the stationarity condition at constant energy (77), but now the subscripts l and m are the same, and a term in $-\widetilde{e}_l$ is present in the operator.

3. Temperature dependent mean field equations

Introducing a Hartree-Fock operator acting in the single particle state space allows writing the stationarity relations just obtained in a more concise and manageable form, as we now show.

3-a. Form of the equations

Let us define a temperature dependent Hartree-Fock operator as the partial trace that appears in the previous equations:

$$\boxed{W_{HF}(\beta) = \mathrm{Tr}_2\left\{\widetilde{\rho}_1(2) W_2(1,2)\left[1+\eta P_{\mathrm{ex}}(1,2)\right]\right\}} \tag{84}$$

It is thus an operator acting on the single particle 1. It can be defined just as well by its matrix elements between the individual states:

$$\begin{aligned}&\langle\theta_k|W_{HF}(\beta)\,|\theta_l\rangle \\ &\quad = \sum_j f_\beta(\widetilde{e}_j-\mu)\,\langle 1:\theta_k; 2:\theta_j|\,W_2(1,2)\left[1+\eta P_{\mathrm{ex}}(1,2)\right]|1:\theta_l; 2:\theta_j\rangle\end{aligned} \tag{85}$$

Equation (77) is valid for any two chosen values l and m, as long as $\widetilde{e}_l \neq \widetilde{e}_m$. When l is fixed and m varies, it simply means that the ket:

$$\left[K_0+V_1+W_{HF}(\beta)\right]|\theta_l\rangle \tag{86}$$

is orthogonal to all the eigenvectors $|\theta_m\rangle$ having an eigenvalue $\widetilde{e}_m$ different from $\widetilde{e}_l$; it has a zero component on each of these vectors. As for equation (83), it yields the component of this ket on $|\theta_l\rangle$, which is equal to $\widetilde{e}_l\,|\theta_l\rangle$. The set of $|\theta_m\rangle$ (including those having the same eigenvalue as $|\theta_l\rangle$) form a basis of the individual state space, defined by (26) as the basis of eigenvectors of the individual operator $\widetilde{h}$. Two cases must be distinguished:

(i) If $\widetilde{e}_l$ is a non-degenerate eigenvalue of $\widetilde{h}$, the set of equations (77) and (83) determine all the components of the ket $\left[K_0+V_1+W_{HF}(\beta)\right]|\theta_l\rangle$. This shows that $|\theta_l\rangle$ is an eigenvector of the operator $K_0+V_1+W_{HF}$ with the eigenvalue $\widetilde{e}_l$.

(ii) If this eigenvalue of $\widetilde{h}$ is degenerate, relation (77) only proves that the eigen-subspace of $\widetilde{h}$, with eigenvalue $\widetilde{e}_l$, is stable under the action of the operator $K_0+V_1+W_{HF}$; it does not yield any information on the components of the ket (86) inside that subspace. It is possible though to diagonalize $K_0+V_1+W_{HF}$ inside each of the eigen-subspace of $\widetilde{h}$, which leads to a new eigenvectors basis $|\varphi_n\rangle$, now common to $\widetilde{h}$ and $K_0+V_1+W_{HF}$.

We now reason in this new basis where all the $\left[K_0+V_1+W_{HF}(\beta)\right]|\varphi_n\rangle$ are proportional to $|\varphi_n\rangle$. Taking (83) into account, we get:

$$\boxed{\left[K_0+V_1+W_{HF}(\beta)\right]|\varphi_n\rangle = \widetilde{e}_n\,|\varphi_n\rangle} \tag{87}$$

As we just saw, the basis change from the $|\theta_l\rangle$ to the $|\varphi_n\rangle$ only occurs within the eigen-subspaces of $\widetilde{h}$ corresponding to given eigenvalues $\widetilde{e}$; one can then replace the $|\theta_l\rangle$ by the $|\varphi_n\rangle$ in the definition (40) of $\widetilde{\rho}_1(1)$ and write:

$$\boxed{\widetilde{\rho}_1(1) = \sum_n f_\beta(\widetilde{e}_n-\mu)\;\;|1:\varphi_n\rangle\,\langle 1:\varphi_n|} \tag{88}$$

Inserting this relation in the definition (84) of $W_{HF}(\beta)$ leads to a set of equations only involving the eigenvectors $|\varphi_n\rangle$.

For all the values of n we get a set of equations (87), which, associated with (84) and (88) defining the potential $W_{HF}(\beta)$ as a function of the $|\varphi_n\rangle$, are called the temperature dependent Hartree-Fock equations.

3-b. Properties and limits of the equations

We now discuss how to apply the mean field equations we have obtained, and their limit of validity, which are more stringent for bosons than for fermions.

α. *Using the equations*

Hartree-Fock equations concern a self-consistent and nonlinear system: the eigenvectors $|\varphi_n\rangle$ and eigenvalues of the density operator $\widetilde{\rho}_1(1)$ are solutions of an eigenvalue equation (87) which itself depends on $\widetilde{\rho}_1(1)$. This situation is reminiscent of the one encountered with the zero-temperature Hartree-Fock equations, and, a priori, no exact solutions can be found.

As for the zero-temperature case, we proceed by iteration: starting from a physically reasonable density operator $\widetilde{\rho}_1(1)$, we use it in (84) to compute a first value of the Hartree-Fock potential operator. We then diagonalize this operator to get its eigenkets and eigenvalues $\bar{e}_n$. Next, we build the operator $\overline{\rho}_1$ that has the same eigenkets, but whose eigenvalues are the $f_\beta(\bar{e}_n - \mu)$. Inserting this new operator $\overline{\rho}_1$ in (84), we get a second iteration of the Hartree-Fock operator. We again diagonalize this operator to compute new eigenvalues and eigenvectors, on which we build the next approximation $\overline{\overline{\rho}}_1$ of $\widetilde{\rho}_1$, and so on. After a few iterations, we may expect convergence towards a self-consistent solution.

β. *Validity limit*

For a fermion system, there is no fundamental general limit for using the Hartree-Fock approximation. The pertinence of the final result obviously depends on the nature of the interactions, and whether a mean field treatment of these interactions is a good approximation. One can easily understand that the larger the interaction range, the more each particle will be submitted to the action of many others. This will lead to an averaging effect improving the mean field approximation. If, on the other hand, each particle only interacts with a single partner, strong binary correlations may appear, which cannot be correctly treated by a mean field acting on independent particles.

For bosons, the same general remarks apply, but the populations are no longer limited to 1. When, for example, Bose-Einstein condensation occurs, one population becomes much larger than the others, and presents a singularity that is not accounted for in the calculations presented above. The Hartree-Fock approximation has therefore more severe limitations than for the fermions, and we now discuss this problem.

For a boson system in which many individual states have comparable populations, taking into account the interactions by the Hartree-Fock mean field yields as good an approximation as for a fermion system. If the system however is close to condensation, or already condensed, the mean field equations we have written are no longer valid. This is because the trial density operator in relation (31) contains a distribution function associated with each individual quantum state and varies as for an ideal gas, i.e. as an exponentially decreasing function of the occupation numbers. Now we saw in § 3-b-β of Complement $\mathrm{B_{XV}}$ that, in an ideal gas, the fluctuations of the particle numbers in each of the individual states are as large as the average values of those particle numbers. If the individual state has a large population, these fluctuations can become very important, which is physically impossible in the presence of repulsive interactions. Any population fluctuation increases the average value of the square of the occupation number (equal

to the sum of the squared average value and the squared fluctuation), and hence of the interaction energy (proportional to the average value of the square). A large fluctuation in the populations would lead to an important increase of the interaction repulsive energy, in contradiction with the minimization of the thermodynamic potential. In other words, the finite compressibility of the physical system, introduced by the interactions, prevents any large fluctuation in the density. Consequently, the fluctuations in the number of condensed particles predicted by the trial Hartree-Fock density operator are not physically acceptable, in the presence of condensation.

It is worth analyzing more precisely the origin of this Hartree-Fock approximation limit, in terms of correlations between the particles. Relation (51) concerns any two-particle operator $\widehat{G}$. It shows that, using the trial density operator (31), the two-particle reduced density operator can be written as:

$$\widetilde{\rho}_2(1,2) = \widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(1)\,[1 + P_{\rm ex}(1,2)] \tag{89}$$

Its diagonal matrix elements are then written:

$$\begin{aligned} &\langle 1:\theta_i; 2:\theta_j|\,\widetilde{\rho}_2(1,2)\,|1:\theta_i; 2:\theta_j\rangle \\ &= \langle 1:\theta_i|\,\widetilde{\rho}_1\,|1:\theta_i\rangle\,\langle 2:\theta_j|\,\widetilde{\rho}_1\,|2:\theta_j\rangle + \langle 1:\theta_i|\,\widetilde{\rho}_1\,|1:\theta_j\rangle\,\langle 2:\theta_j|\,\widetilde{\rho}_1\,|2:\theta_i\rangle \end{aligned} \tag{90}$$

and are the sum of a direct term, and an exchange term. When $i \neq j$, the presence of an exchange term is not surprising, and corresponds to the general discussion of § C-5 in Chapter XV. It is similar to the expression of the spatial correlation function written in (C-34) of that chapter, which is also the sum of two contributions, a direct one (C-32) and an exchange one (C-33). Since this last contribution is positive when $r_1 \simeq r_2$, the physical consequence of the exchange is a spatial bunching of the bosons. What is surprising though is that the exchange term still exists in (90) when $i = j$, even though the notion of exchange is meaningless: when dealing with a single individual state, the four expressions (C-21) of Chapter XV reduce to a single one, the direct term. We can also check that the exchange term (C-34) of Chapter XV includes the explicit condition $i \neq j$, which means it receives no contribution from $i = j$. We shall furthermore confirm in § 3 of Complement A_{XVI} that bosons all placed in the same individual quantum state are not spatially correlated, and therefore present neither bunching nor exchange effects. The mathematical expression of the trial two-particle Hartree-Fock density operator thus contains too many exchange terms. This does not really matter as long as the boson system remains far from Bose-Einstein condensation: the error involved is small since the $i = j$ terms play a negligible role compared to the $i \neq j$ terms in the summations over i and j appearing in the interaction energy. However, as soon as an individual state becomes highly populated, significant errors can occur and the Hartree-Fock method must be abandoned. There exist, however, more elaborate theoretical treatments better adapted to this case.

3-c. Differences with the zero-temperature Hartree-Fock equations (fermions)

The main difference between the approach we just used and that of Complements C_{XV} and E_{XV} is that these complements were only looking for a single eigenstate of the Hamiltonian $\hat{H}$, generally its ground state. If we are now interested in several of these states, we have to redo the computation separately for each of them. To study the

properties of thermal equilibrium, one could imagine doing the calculations a great many times, and then weigh the results with occupation probabilities. This method obviously leads to heavy computations, which become impossible for a macroscopic system having an extremely large number of levels. In the present complement, the Hartree-Fock equations yield immediately thermal averages, as well as eigenvectors of a one-particle density operator with their energies.

Another important difference is that the Hartree-Fock operator now depends on the temperature, because of the presence in (85) of a temperature dependent distribution function – or, which amounts to the same thing, of the presence in (84) of an operator dependent on β, and which replaces the projector $P_N(2)$ onto all the populated individual states. The equations obtained remind us of those governing independent particles, each finding its thermodynamic equilibrium while moving in the self-consistent mean field created by all the others, also including the exchange contribution (which can be ignored in the simplified "Hartree" version).

We must keep in mind, however, that the Hartree-Fock potential associated with each individual state now depends on the populations of an infinity of other individual states, and these populations are function of their energy as well as of the temperature. In other words, because of the nonlinear character of the Hartree-Fock equations, the computation is not merely a juxtaposition of separate mean field calculations for stationary individual states.

3-d. Zero-temperature limit (fermions)

Let us check that the Hartree-Fock method for non-zero temperature yields the same results as the zero temperature method explained in Complement $\mathrm{E_{XV}}$ for fermions.

In § 2-d of Complement $\mathrm{B_{XV}}$, we introduced for an ideal gas the concept of a degenerate quantum gas. It can be generalized to a gas with interactions: in a fermion system, when $\beta\mu \gg 1$, the system is said to be strongly degenerate. As the temperature goes to zero, a fermion system becomes more and more degenerate. Can we be certain that the results of this complement are in agreement with those of Complement $\mathrm{E_{XV}}$, valid at zero temperature?

We saw that the temperature comes into play in the definition (85) of the mean Hartree-Fock potential, W_{HF}. In the limit of a very strong degeneracy, the Fermi-Dirac distribution function appearing in the definition (40) of $\widetilde{\rho}_1(1)$ becomes practically a step function, equal to 1 for energies e_j less than the chemical potential μ, and zero otherwise (Figure 1 of Complement $\mathrm{B_{XV}}$. In other words, the only populated states (and by a single fermion) are the states having energies less than μ, i.e. less than the Fermi level. Under such conditions, the $\widetilde{\rho}_1(2)$ of (84) becomes practically equal to the projector $P_N(2)$ which, in Complement $\mathrm{E_{XV}}$, appears in the definition (52) of the zero-temperature Hartree-Fock potential; in other words, the partial trace appearing in this relation (85) is then strictly limited to the individual states having the lowest energies. We thus obtain the same Hartree Fock equations as for zero temperature, leading to the determination of a set of individual eigenstates on which we can build a unique N-particle state.

3-e. Wave function equations

Let us write the Hartree-Fock equations (87) in terms of wave functions: these equations are strictly equivalent to (87), written in terms of operators and kets, but their

form is sometimes easier to use, in particular for numerical calculations.

Assuming the particles have a spin, we shall note the wave functions $\varphi^{\nu}(\mathbf{r})$, with:

$$\varphi^{\nu}(\mathbf{r}) = \langle \mathbf{r}, \nu \, | \varphi \rangle \tag{91}$$

where the spin quantum number ν can take $(2S+1)$ values; according to the nature of the particles, the possible spins S are $S = 0$, $S = 1/2$, $S = 1$ etc. As in Complement E_{XV} (§ 2-d), we introduce a complete basis $\{|\varphi_k^{\nu_k}\rangle\}$ for the individual state space, built from kets that are all eigenvectors of the spin component along the quantization axis, with eigenvalue ν_k. For each value of n, the spin index ν takes on a given value ν_n and is not, therefore, an independent index. As for the potentials, we assume here again that V_1 is diagonal in ν, but that its diagonal elements $V_1^{\nu}(\mathbf{r})$ may depend on ν. The interaction potential, however, is described by a function $W_2(\mathbf{r}, \mathbf{r}')$ that only depends on $\mathbf{r} - \mathbf{r}'$, but does not act on the spins.

To obtain the matrix elements of $\widehat{W}_{HF}(\beta)$ in the representation $\{|\mathbf{r}, \nu\rangle\}$, we use (85) after replacing the $|\theta\rangle$ by the $|\varphi\rangle$(we showed in § 3 that this was possible). We now multiply both sides by $\langle \mathbf{r}, \nu \, | \varphi_k^{\nu_k} \rangle$ and $\langle \varphi_l^{\nu_l} \, | \mathbf{r}', \nu \rangle$, and sum over the subscripts k and l; we recognize in both sides the closure relations:

$$\sum_k \langle \mathbf{r}, \nu \, | \varphi_k^{\nu_k} \rangle \langle \varphi_k^{\nu_k} | = \langle \mathbf{r}, \nu | \quad \text{and} \quad \sum_l | \varphi_l^{\nu_l} \rangle \langle \varphi_l^{\nu_l} \, | \mathbf{r}', \nu' \rangle = | \mathbf{r}', \nu' \rangle \tag{92}$$

This leads to:

$$\begin{aligned} &\langle \mathbf{r}, \nu | \, \widehat{W}_{HF}(\beta) \, | \mathbf{r}', \nu' \rangle = \\ &\qquad \sum_j f_\beta(\widetilde{e}_j - \mu) \; \left\langle 1 : \mathbf{r}, \nu; 2 : \varphi_j^{\nu_j} \right| W_2(1,2) \left[1 + \eta P_{ex}(1,2) \right] \left| 1 : \mathbf{r}', \nu'; 2 : \varphi_j^{\nu_j} \right\rangle \end{aligned} \tag{93}$$

As in § C-5 of Chapter XV, we get the sum of a direct term (the term 1 in the central bracket) and an exchange term (the term in ηP_{ex}). This expression contains the same matrix element as relation (87) of Complement E_{XV}, the only difference being the presence of a coefficient $f_\beta(\widetilde{e}_j - \mu)$ in each term of the sum (plus the fact that the summation index goes to infinity).

(i) For the direct term, as we did in that complement, we insert a closure relation on the particle 2 position:

$$\left| 2 : \varphi_j^{\nu_j} \right\rangle = \int d^3 r_2 \; \varphi_j^{\nu_j}(\mathbf{r}_2) \, | 2 : \mathbf{r}_2, \nu_j \rangle \tag{94}$$

Since the interaction operator is diagonal in the position representation, the part of the matrix element of (93) that does not contain the exchange operator becomes:

$$\int d^3 r_2 \; \left| \varphi_j^{\nu_j}(\mathbf{r}_2) \right|^2 \; \langle 1 : \mathbf{r}, \nu; 2 : \mathbf{r}_2 | \, W_2(1,2) \, | 1 : \mathbf{r}', \nu'; 2 : \mathbf{r}_2 \rangle \tag{95}$$

The direct term of (93) is then written:

$$\delta_{\nu\nu'} \; \delta(\mathbf{r} - \mathbf{r}') \int \mathrm{d}^3 r_2 \; W_2(\mathbf{r}, \mathbf{r}_2) \sum_j f_\beta(\widetilde{e}_j - \mu) \left| \varphi_j^{\nu_j}(\mathbf{r}_2) \right|^2 \tag{96}$$

which is equivalent to relation (91) of Complement E_{XV}.

(ii) The exchange term is obtained by permutation of the two particles in the ket appearing on the right-hand side of (93); the diagonal character of $W_2(1,2)$ in the position representation leads to the expression:

$$\left\langle 1:\mathbf{r},\nu \left| 1:\varphi_j^{\nu_j} \right\rangle \left\langle 2:\varphi_j^{\nu_j} \right| 2:\mathbf{r}',\nu' \right\rangle W_2(\mathbf{r},\mathbf{r}') \tag{97}$$

For the first scalar product to be non-zero, the subscript j must be such that $\nu_j = \nu$; in the same way, for the second product to be non-zero, we must have $\nu_j = \nu'$. For both conditions to be satisfied, we must impose $\nu = \nu'$, and the exchange term (93) is equal to:

$$\eta\ \delta_{\nu\nu'} \sum_{\nu_j=\nu} f_\beta(\widetilde{e}_j - \mu)\ \left[\varphi_j^{\nu_j}(\mathbf{r})\right] \left[\varphi_j^{\nu_j}(\mathbf{r}')\right]^* W_2(\mathbf{r},\mathbf{r}') \tag{98}$$

where the summation is on all the values of j such that $\nu_j = \nu$: this term only exists if the two interacting particles are totally indistinguishable, which requires that they be in the same spin state (see the discussion in Complement $\mathrm{E_{XV}}$).

We now define the direct and exchange potentials by:

$$\begin{aligned} V_{\mathrm{dir}}(\mathbf{r}) &= \int d^3r''\ W_2(\mathbf{r},\mathbf{r}'') \sum_j f_\beta(\widetilde{e}_j - \mu) \left|\varphi_j^{\nu_j}(\mathbf{r}'')\right|^2 \\ V_{\mathrm{ex}}^{\nu}(\mathbf{r},\mathbf{r}') &= W_2(\mathbf{r},\mathbf{r}') \sum_{\nu_j=\nu} f_\beta(\widetilde{e}_j - \mu)\ \left[\varphi_j^{\nu_j}(\mathbf{r})\right] \left[\varphi_j^{\nu_j}(\mathbf{r}')\right]^* \end{aligned} \tag{99}$$

The equalities (87) then lead to the Hartree-Fock equations in the position representation:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1^{\nu_n}(\mathbf{r}) + V_{\mathrm{dir}}(\mathbf{r})\right] \varphi_n^{\nu_n}(\mathbf{r}) + \eta \int \mathrm{d}^3r'\ V_{\mathrm{ex}}^{\nu_n}(\mathbf{r},\mathbf{r}')\varphi_n^{\nu_n}(\mathbf{r}') = \widetilde{e}_n\ \varphi_n^{\nu_n}(\mathbf{r})} \tag{100}$$

The general discussion of § 3-b can be applied here without any changes. These equations are both nonlinear and self-consistent, as the direct and exchange potentials are themselves functions of the solutions $\varphi_j^{\nu}(\mathbf{r})$ of the eigenvalue equations (100). This situation is reminiscent of the zero-temperature case, and we can, once again, look for solutions using iterative methods. The number of equations to be solved, however, is infinite and no longer equal to the finite number N, as already pointed out in § 3-c. The set of solutions must span the entire individual state space. Along the same line, in the definitions (99) of the direct and exchange potentials, the summations over j are not limited to N states, but go to infinity. However, even though the number of these wave functions is in principle infinite, it is limited in practice (for numerical calculations) to a high but finite number. As for the initial conditions to start the iteration process, one can choose for example the states and energies of a free fermion gas, but any other conjecture is equally possible.

Conclusion

There are many applications of the previous calculations, and more generally of the mean field theory. We give a few examples in the next complement, which are far from showing the richness of the possible application range. The main physical idea is to

reduce, whenever possible, the calculation of the various physical quantities to a problem similar to that of an ideal gas, where the particles have independent dynamics. We have indeed shown that the individual level populations, as well as the total particle number, are given by the same distribution functions f_β as for an ideal gas – see relations (38) and (44). The same goes for the system entropy S, as already mentioned at the end of § 2-b-α. If we replace the free particle energies by the modified energies $\widetilde{e}_{\mathbf{k},\nu}$, the analogy with independent particles is quite strong.

If we now want to compute other thermodynamic quantities, as for example the average energy, we can no longer use the ideal gas formulas; we must go back to the equations of § 2-c. The grand potential may be calculated by inserting in (61) the $|\theta_i\rangle$ and the $\widetilde{e}_i$ obtained from the Hartree-Fock equations. Another method uses the fact that $\left\langle \widehat{N} \right\rangle$ is given by ideal gas formulas that contain the distribution f_β, and hence do not require any further calculations. As:

$$\left\langle \widehat{N} \right\rangle = \beta^{-1} \frac{\partial}{\partial \mu} \ln Z \tag{101}$$

we can integrate $\left\langle \widehat{N} \right\rangle$ over μ (between $-\infty$ and the current value μ, for a fixed value of β) to obtain $\ln Z$, and hence the grand potential. From this grand potential, all the other thermodynamic quantities can be calculated, taking the proper derivatives (for example a derivative with respect to β to get the average energy). We shall see an example of this method in § 4-a of the next complement.

We must however keep in mind that all these calculations derive from the mean field approximation, in which we replaced the exact equilibrium density operator by an operator of the form (32). In many cases this approximation is good, even excellent, as is the case, in particular, for a long-range interaction potential: each particle will interact with several others, therefore enhancing the averaging effect of the interaction potential. It remains, however, an approximation: if, for example, the particles interact via a "hard core" potential (infinite potential when the mutual distance becomes less than a certain microscopic distance), the particles, in the real world, can never be found at a distance from each other smaller than the hard core diameter; now this impossibility is not taken into account in (32). Consequently, there is no guarantee of the quality of a mean field approximation in all situations, and there are cases for which it is not sufficient.

Complement H_{XV}

Applications of the mean field method for non-zero temperature (fermions and bosons)

1 Hartree-Fock for non-zero temperature, a brief review . . . 1733
2 Homogeneous system . . . 1734
2-a Calculation of the energies . . . 1735
2-b Quasi-particules . . . 1736
3 Spontaneous magnetism of repulsive fermions . . . 1737
3-a A simple model . . . 1737
3-b Resolution of the equations by graphical iteration . . . 1739
3-c Physical discussion . . . 1743
4 Bosons: equation of state, attractive instability . . . 1745
4-a Repulsive bosons . . . 1746
4-b Attractive bosons . . . 1747

In the previous complement, we presented the Hartree-Fock (mean field) method for non-zero temperatures, which has numerous applications – a few of them will be discussed in this complement. We start in § 1 with a brief review of the results obtained with this method in the previous complement, and which will be used in the present complement. The general properties of a homogeneous system are then studied in § 2, as this particular case is often encountered, hence giving it a special importance. The last two sections are concerned with the study of phase transitions in homogeneous systems. Section § 3 studies fermions; we show how the mean field theory predicts the existence of a transition where a fermion system becomes spontaneously magnetic because of the repulsion between particles (even though this repulsion is supposed to be completely independent of the spins). Finally, the last section deals with bosons and the study of their equation of state. This will allow us to show, in particular, the appearance of an instability when the bosons are attractive and close to Bose-Einstein condensation.

1. Hartree-Fock for non-zero temperature, a brief review

We start with a brief review of the results obtained previously (§ 2 of Complement B_{XV} and § 3 of Complement G_{XV}, which will be useful for what follows.

For an ideal gas, the distribution function f_β^{FD} for fermions, or f_β^{BE} for bosons, is given by:

$$f_\beta(e-\mu) = \begin{cases} f_\beta^{FD}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}+1} & \text{for fermions} \\ f_\beta^{BE}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}-1} & \text{for bosons} \end{cases} \tag{1}$$

where $\beta = 1/k_B T$ and μ is the chemical potential. The average total particle number $\left\langle \widehat{N} \right\rangle$ is then obtained by a sum over all the individual accessible states, labeled by the

subscript i:

$$\left\langle \widehat{N} \right\rangle = \sum_{i=1}^{\infty} f_\beta(e_i - \mu) \tag{2}$$

The temperature dependent Hartree-Fock equations (mean field equations) in the position representation are given by relation (100) of Complement G_{XV}:

$$\left[-\frac{\hbar^2}{2m}\Delta + V_1^{\nu_n}(\mathbf{r}) + V_{\rm dir}(\mathbf{r}) \right] \varphi_n^{\nu_n}(\mathbf{r}) + \eta \int \mathrm{d}^3 r' \; V_{\rm ex}^{\nu_n}(\mathbf{r},\mathbf{r}') \varphi_n^{\nu_n}(\mathbf{r}') = \widetilde{e}_n \; \varphi_n^{\nu_n}(\mathbf{r}) \tag{3}$$

where $\eta = +1$ for bosons, $\eta = -1$ for fermions, and where $V_{\rm dir}(\mathbf{r})$ and $V_{\rm ex}^{\nu}(\mathbf{r})$ are given by relation (99) of that same complement (we assume the interaction potential does not act on the spin quantum numbers ν):

$$\begin{aligned} V_{\rm dir}(\mathbf{r}) &= \int d^3 r'' \; W_2(\mathbf{r},\mathbf{r}'') \sum_j f_\beta(\widetilde{e}_j - \mu) \left| \varphi_j^{\nu_j}(\mathbf{r}'') \right|^2 \\ V_{\rm ex}^{\nu}(\mathbf{r},\mathbf{r}') &= W_2(\mathbf{r},\mathbf{r}') \sum_{\nu_j = \nu} f_\beta(\widetilde{e}_j - \mu) \; \left[\varphi_j^{\nu_j}(\mathbf{r}) \right] \left[\varphi_j^{\nu_j}(\mathbf{r}') \right]^* \end{aligned} \tag{4}$$

2. Homogeneous system

We assume from now on that the physical system is subjected to boundary conditions created by a one-body potential, which confines the particles inside a cubic box of edge length L; this potential is zero ($\widehat{V}_1 = 0$) inside the box, and takes on an infinite value outside. To take this confinement into account, we shall use the periodic boundary conditions (Complement C_{XIV}, § 1-c), for which the normalized eigenfunctions of the kinetic energy are written as:

$$\frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{5}$$

where the possible wave vectors $\mathbf{k}$ are those whose three components are integer multiples of $2\pi/L$. Because of the spin, the eigenvectors of the kinetic energy are labeled by the values of both $\mathbf{k}$ and ν, and are written $|\mathbf{k},\nu\rangle$, with:

$$\langle \mathbf{r}, \nu \, | \mathbf{k}, \nu \rangle = \varphi_{\mathbf{k}}^{\nu}(\mathbf{r}) = \frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{6}$$

The index n (or j) that labeled the basis vectors in the previous complements is now replaced by two indices, $\mathbf{k}$ and ν (which are independent, as opposed to the indices n and ν_n used in § 3-e of Complement G_{XV}). We shall finally assume that the particle interaction is invariant under translation: $W_2(\mathbf{r}_1,\mathbf{r}_2)$ only depends on $\mathbf{r}_1 - \mathbf{r}_2$.

We are going to see that, in such a case, solutions of the Hartree-Fock equations can be found without having to search for the eigenfunctions of the (Hartree-Fock) operator written on the left-hand side of (3); these solutions are simply the plane waves written in (5). Only the operator's eigenvalues $\widetilde{e}_{\mathbf{k},\nu}$ remain to be calculated, and can be interpreted as the energies of independent objects called "quasi-particles" (§ 2-b).

Comment:

We shall verify that these plane waves are solutions of the Hartree-Fock equations, while neither being necessarily the only ones, nor even those leading to the lowest energy of the total system. A phenomenon called "symmetry breaking" (translation symmetry in this case) could occur and introduce solutions whose moduli vary in space and correspond to lower energies. The Wigner crystal of electrons is such an example, where the particle density spontaneously shows a periodic spatial modulation. Another example of spontaneous symmetry breaking will be discussed in § 3-c of this complement. There are many other cases (in nuclear physics in particular) where the Hartree-Fock method can be used to study symmetry breaking phenomena.

2-a. Calculation of the energies

As the plane waves are obviously eigenfunctions of the kinetic energy, and since the potential is zero inside the box, we just have to demonstrate that they are also eigenfunctions of the direct and exchange potentials. Inserting (5) in (4), we get:

$$\begin{aligned} V_{\text{dir}}(\mathbf{r}) &= \frac{1}{L^3} \sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'} - \mu) \int \mathrm{d}^3 r'' \;\; W_2(\mathbf{r} - \mathbf{r}'') \\ &= \frac{\overline{V}_0}{L^3} \sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'} - \mu) \end{aligned} \tag{7}$$

where $\overline{V}_0$ is defined as (with a change of variable $\mathbf{r} - \mathbf{r}'' = \mathbf{s}$):

$$\overline{V}_0 = \int \mathrm{d}^3 r'' \;\; W_2(\mathbf{r} - \mathbf{r}'') = \int \mathrm{d}^3 s \; W_2(\mathbf{s}) \tag{8}$$

The direct potential is therefore a constant, independent of the position $\mathbf{r}$; multiplying an exponential $e^{i\mathbf{k}\cdot\mathbf{r}}$, it yields a function proportional to it. This means that $e^{i\mathbf{k}\cdot\mathbf{r}}$ is an eigenfunction of the direct potential. As for the exchange potential, using the second relation of (4), we get:

$$V_{\text{ex}}^{\nu}(\mathbf{r}, \mathbf{r}') = \frac{1}{L^3} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu} - \mu) \; e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')} \; W_2(\mathbf{r} - \mathbf{r}') \tag{9}$$

The exchange potential is thus also translation-invariant (it only depends on the difference $\mathbf{r} - \mathbf{r}'$). Consequently, the last term on the left-hand side of (3) can be written as:

$$\begin{aligned} &\frac{\eta}{L^3} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu} - \mu) \int \mathrm{d}^3 r' e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')} \; W_2(\mathbf{r} - \mathbf{r}') \frac{e^{i\mathbf{k}\cdot\mathbf{r}'}}{L^{3/2}} \\ &\qquad = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{L^{3/2}} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu} - \mu) \; \frac{\eta}{L^3} \overline{V}_{(\mathbf{k}'-\mathbf{k})} \end{aligned} \tag{10}$$

where (with the change of variable $\mathbf{r} - \mathbf{r}' = \mathbf{s}$):

$$\overline{V}_{(\mathbf{k}'-\mathbf{k})} = \int \mathrm{d}^3 s \; e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{s}} \; W_2(\mathbf{s}) \tag{11}$$

Consequently, the exchange term simply multiplies the plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}/L^{3/2}$ by:

$$\frac{\eta}{L^3}\sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k},\nu}-\mu)\ \overline{V}_{(\mathbf{k}'-\mathbf{k})} \tag{12}$$

To sum up, we showed that, for a uniform system, the plane waves are indeed solutions of the Hartree-Fock equations (3). It is no longer necessary to solve the eigenvector equations, but we simply have to replace in (3) the $\varphi^\nu_{\mathbf{k}}(\mathbf{r})$ by plane waves. This leads to:

$$\boxed{\widetilde{e}_{\mathbf{k},\nu} = e_k + \frac{\overline{V}_0}{L^3}\sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'}-\mu) + \frac{\eta}{L^3}\sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)\overline{V}_{(\mathbf{k}'-\mathbf{k})}} \tag{13}$$

where e_k is the kinetic energy of a free particle:

$$e_k = \frac{\hbar^2\mathbf{k}^2}{2m} \tag{14}$$

We have obtained self-consistent conditions for the eigenvalues, which are a set of coupled nonlinear equations because of the $f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)$ dependence on the energies $\widetilde{e}_{\mathbf{k}',\nu}$.

Comment:

The exchange term contains the Fourier transform at the spatial frequency $\mathbf{k}'-\mathbf{k}$ of the particle interaction potential; the direct term, however, contains the Fourier component at zero spatial frequency. This property is easy to understand from a physical point of view. Consider two particles, having respectively an initial momentum $\hbar\mathbf{k}$ and $\hbar\mathbf{k}'$. We saw in Chapter VIII (§ B-4-a) that the effect, to first order (Born approximation), of an interaction potential is proportional to the Fourier transform of that potential, calculated at the value of the variation of the relative momentum between the two particles (Chapter VII, § B-2-a); this variation is none other than the momentum transfer between the particle as they interact. Consequently, it is normal that the system energy is the sum of two terms: a direct term where no particle changes its momentum (no momentum transfer, hence a Fourier variable equal to zero); and another one where the two particles exchange their momenta, so that the relative momentum changes sign and the Fourier variable is proportional to the difference $\mathbf{k}'-\mathbf{k}$.

2-b. Quasi-particules

Equations (13) yield the individual energies $\widetilde{e}_{\mathbf{k},\nu}$, which are the sum of the free particle energy $\hbar^2k^2/2m$ and a contribution from the interactions. One can look at them as energies of individual objects[1], often referred to as "quasi-particles". The populations of the corresponding levels, as well as the total number of quasi-particles, are given by the same distribution functions f_β as for an ideal gas – see relations (39) and (44) of Complement G_{XV}. The same is true for the system entropy S, as we already mentioned at the end of § 2-b-α of that complement. Provided we replace the free particle energy by the modified energies $\widetilde{e}_{\mathbf{k},\nu}$, the analogy with independent particles is quite strong.

[1]The concept of quasi-particle is not necessarily limited to systems whose interacting particles are free inside a box; it remains valid for non-zero V_1 potentials (a harmonic potential for example). The first term in (13) must then be replaced by the particle energy in the potential V_1, and the direct and exchange terms will have a different expression.

3. Spontaneous magnetism of repulsive fermions

Consider a system of spin 1/2 fermions, contained in a box. To make the computations easier, we shall make a few simplifying hypotheses. They lead to a simple model, giving a good illustration of the nonlinear character of the Hartree-Fock theory. They involve the resolution of a set of nonlinear equations, containing only two variables – equations we will write in (22).

3-a. A simple model

We assume the mutual interaction potential to be repulsive and to have a very short range r_0. In relation (13) the only vectors $\mathbf{k}$ and $\mathbf{k}'$ that matter are those for which the distributions $f_\beta(\widetilde{e}_{\mathbf{k},\nu}-\mu)$ and $f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)$ are not negligible. If, for all these vectors, the products kr_0 and $k'r_0$ are very small compared to 1, the product $(\mathbf{k}'-\mathbf{k})\cdot\mathbf{s}$ may be replaced by zero in (11), and we get:

$$\overline{V}_{(\mathbf{k}'-\mathbf{k})} \simeq \overline{V}_0 \tag{15}$$

α. Energy of the quasi-particles; spin state populations

As $\eta = -1$ for fermions, equation (13) can be simplified:

$$\widetilde{e}_{\mathbf{k},\nu} = e_k + \frac{\overline{V}_0}{L^3}\left[\sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'}-\mu) - \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)\right] \tag{16}$$

or else:

$$\widetilde{e}_{\mathbf{k},\nu} = e_k + \frac{\overline{V}_0}{L^3}\sum_{\mathbf{k}',\nu'\neq\nu} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'}-\mu) \tag{17}$$

Consequently, the energy of a quasi-particle with a given ν is only modified by the interaction with quasi-particles having a different spin component ν' (opposite spin if the particles have a spin $S = 1/2$). This result was to be expected since if the spins of the two quasi-particles are parallel, they cannot be distinguished; the Pauli principle then forbids them to approach at a distance closer than r_0, and they cannot interact. On the other hand, if their spins are opposite, they can be identified by the direction of their spin (we have assumed the interaction does not act on the spins): they behave as distinguishable particles, the exclusion principle does not apply, and they now interact.

We note $\left\langle \widehat{N}_+ \right\rangle$ and $\left\langle \widehat{N}_- \right\rangle$ the total particle numbers respectively in the spin state $+$ or $-$:

$$\left\langle \widehat{N}_\pm \right\rangle = \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\pm}-\mu) \tag{18}$$

Equation (17) shows that the energies of the $+$ and $-$ spin states are modified according to:

$$\widetilde{e}_{\mathbf{k},+} = e_k + g\left\langle \widehat{N}_- \right\rangle \quad ; \quad \widetilde{e}_{\mathbf{k},-} = e_k + g\left\langle \widehat{N}_+ \right\rangle \tag{19}$$

where the coupling constant g (having the dimension of an energy) is defined by:

$$g = \frac{\overline{V}_0}{L^3} \tag{20}$$

Since the particle numbers only depend on the difference between the energies $\widetilde{e}$ and the chemical potential μ, we can account for the terms in g appearing in (19) by keeping the energies e_k for free particles, but lowering the chemical potentials $\mu_\pm$ by the quantity $g\left\langle \widehat{N}_\pm \right\rangle$. Calling $N_{ig}^{FD}(\beta,\mu)$, as in relation (47) of Complement B_{XV}, the total number of fermions in an ideal gas:

$$N_{ig}^{FD}(\beta,\mu) = \left(\frac{L}{2\pi}\right)^3 \int \mathrm{d}^3k \frac{1}{e^{\beta(e_k-\mu)}+1} \tag{21}$$

we get, for an interacting gas:

$$\begin{cases} \left\langle \widehat{N}_+ \right\rangle = N_{ig}^{FD}\left(\beta, \mu - g\left\langle \widehat{N}_- \right\rangle\right) \\ \left\langle \widehat{N}_- \right\rangle = N_{ig}^{FD}\left(\beta, \mu - g\left\langle \widehat{N}_+ \right\rangle\right) \end{cases} \tag{22}$$

These equations determine the populations of the two spin states as a function of the parameters β (or the temperature), μ and finally the volume $\mathcal{V}$. These are, however, two coupled equations since the population $\left\langle \widehat{N}_+ \right\rangle$ depends on $\left\langle \widehat{N}_- \right\rangle$ and conversely. Finding their solution is not obvious, and we shall use a change of variables and resort to a graphic construction.

β. Change of variables

It is useful to write the previous relations in terms of dimensionless variables. We shall thus introduce the "thermal wavelength" λ_T by:

$$\lambda_T = \hbar\sqrt{\frac{2\pi}{mk_BT}} = \hbar\sqrt{\frac{2\pi\beta}{m}} \tag{23}$$

We can now make the same change of integration variable as in § 4-a of Complement B_{XV}:

$$\kappa = k\,\frac{\lambda_T}{2\sqrt{\pi}} \tag{24}$$

Relation (21) then becomes:

$$N_{ig}^{FD}(\beta,\mu) = \left(\frac{L}{\lambda_T}\right)^3 I_{3/2}(\beta\mu) \tag{25}$$

where:

$$I_{3/2}(\beta\mu) = \pi^{-3/2}\int \mathrm{d}^3\kappa \frac{1}{e^{\kappa^2-\beta\mu}+1} \tag{26}$$

These relations are just the same as those written in (51) and (52) in Complement B_{XV}. The value of $I_{3/2}$ only depends on a dimensionless variable, the product $\beta\mu$. As opposed

to N_{ig}^{FD} which is an "extensive" quantity (proportional[2] to the volume $\mathcal{V} = L^3$ of the system), $I_{3/2}(\beta\mu)$ is an "intensive" quantity (independent of the volume).

We can also replace the two unknowns, which are the populations $\left\langle \widehat{N}_{\pm} \right\rangle$ of the two spin states, by two dimensionless and intensive variables $x_{\pm}$:

$$x_{\pm} = \left(\frac{\lambda_T}{L} \right)^3 \left\langle \widehat{N}_{\pm} \right\rangle \tag{27}$$

To characterize the interactions appearing in relations (22) via the constant g, we introduce the dimensionless parameter $\overline{g}$:

$$\overline{g} = \beta g \left(\frac{L}{\lambda_T} \right)^3 = \frac{m\overline{V}_0}{2\pi\hbar^2\lambda_T} \tag{28}$$

Replacing in equations (22) the $\left\langle \widehat{N}_{\pm} \right\rangle$ by the $x_{\pm}$, and g by $\overline{g}$, we get the simpler form:

$$\begin{cases} x_+ = F^{(1)}(x_-) \\ x_- = F^{(1)}(x_+) \end{cases} \tag{29}$$

where the function $F^{(1)}$ is defined by:

$$F^{(1)}(x) = I_{3/2}\left(\beta\mu - \overline{g}x\right) \tag{30}$$

(this function depends not only on x, but also on the parameters $\beta\mu$ and $\overline{g}$). As (22), the system (29) contains two coupled equations: x_- allows computing directly x_+, and vice versa.

3-b. Resolution of the equations by graphical iteration

We now show how to solve equations (29) by a graphical method. The two variables x_+ and x_- can be uncoupled by noting that:

$$x_+ = F^{(1)}\left[F^{(1)}(x_+)\right] \tag{31}$$

with the same equation for x_-. We now introduce a second iterated function $F^{(2)}$ of the function $F^{(1)}$ (function of the same function) as:

$$F^{(2)}(x) = F\left[F(x)\right] \tag{32}$$

This leads to:

$$x_+ = F^{(2)}(x_+) \tag{33}$$

Applying the function $F^{(2)}$ to the variable x yields the same value x, which is said to be the abscissa of a "fixed point" of this function. Graphically, the fixed points of any function F are at the intersections of the curve representing the function F with the first bisector.

[2] In a more general way, Appendix VI recalls that a quantity is said to be extensive if, in the limit of large volumes (for T and μ constant), its ratio to the volume $\mathcal{V}$ tends toward a constant; this does not prevent the quantity from containing terms in L^2 for example. On the other hand, it is said to be intensive if, in this same limit (and without dividing by the volume), it tends toward a constant.

α. Iterations of a function

Consider, from a general point of view, the equation:

$$x = G(x) \tag{34}$$

whose solutions correspond to the fixed points of the function G. These solutions can be found by iteration: starting from an approximate value x_1 of the solution, we compute $G(x_1)$, then use $x_2 = G(x_1)$ as a new value of the variable to compute $x_3 = G(x_2)$, etc. It can be shown that this iteration process converges toward the solution of equation (34), hence toward the fixed point on the first bisector, if the slope of the function G at that point is included between -1 and $+1$, that is if:

$$-1 < G'(x) < +1 \tag{35}$$

where G' is the derivative of the function G. The fixed point of the application G is then said to be stable. On the other hand, if that slope is outside the interval $[-1, +1]$, the fixed point on the first bisector becomes unstable; the iteration method for G no longer converges.

We can also introduce the "second iterated function" $G^{(2)}(x) = G\,[G(x)]$. Any fixed point of G is necessarily a fixed point of $G^{(2)}$. The inverse is not true, as it is possible to get a "two-order cycle" where two different values of x are swapped under the effect of G:

$$\begin{aligned} x_2 &= G\,(x_1) \\ x_1 &= G\,(x_2) \end{aligned} \tag{36}$$

In such a case, x_1 and x_2 are both fixed points of $G^{(2)}$, but not of G (we shall see below an example of such a situation, illustrated by Figure 3). These fixed points can be stable for $G^{(2)}$, in which case they constitute a "stable cycle of order two" for the initial function G. After a certain number of iterations of G, the solution converges toward a series taking alternatively two distinct values, x_1 and x_2.

The process may repeat itself: it is possible for the fixed point of $G^{(2)}$ to next become unstable, and yield fixed points for an iterated function of a higher order, and hence to a stable cycle of that order.

β. Form of the function $F^{(1)}$

Relation (25) shows that the variations of $I_{3/2}(x)$ as a function of x are very similar to the variations of $N_{ig}^{FD}(\beta, \mu)$ as a function of μ, already studied in Complement B_{XV} (Figure 2). The equality (30) shows that the plot of $F^{(1)}(x)$ can be deduced from that of $I_{3/2}(x)$:

- reversing the variable x (symmetry with respect to the vertical axis)
- multiplying this variable by $\overline{g}$ (scale change of the abscissa)
- and finally shifting to the right the abscissa origin by the value $\beta\mu$.

This leads to the solid line curve in Figure 1, that plots a constantly decreasing function (for fixed values of β, μ and $\overline{g}$).

As the parameter $\overline{g}$ changes (for β and μ constant), we get a set of curves representing $F^{(1)}(x)$ for each value of $\overline{g}$. For $x = 0$, all these curves go through the same point of ordinate $I_{3/2}(\beta\mu)$. If $\overline{g} = 0$, the curve is a simple horizontal line going through this point. When $\overline{g}$ becomes slightly positive, the curve starts decreasing but still extends along the abscissa axis; it has a small negative slope at the origin. As $\overline{g}$ increases more and more, the curve contracts more and more toward the ordinate axis and its slope at the origin is more and more negative. In the limit where $\overline{g}$ goes to infinity, the curve becomes a straight vertical line.

γ. Form of the function $F^{(2)}$

Figure 1 also shows the geometric construction used to go from $F^{(1)}$ to $F^{(2)}$. For a given value of x, we start from point M_1 of ordinate $F^{(1)}(x)$, and draw a horizontal line until it crosses the first bisector in M_2, which transfers the ordinate of M_1 onto the abscissa. From the intersection point M_2 we draw a vertical line that intersects the function $F^{(1)}$ at point M_3 of ordinate $F^{(2)}(x)$, which we simply transfer to the initial abscissa x to get the final point (surrounded by a triangle).

This construction shows that $F^{(2)}(x)$ is an increasing function of x confined between two horizontal asymptotes: the abscissa axis, and a horizontal line of ordinate $I_{3/2}(\beta\mu)$. The larger the value of $\overline{g}$, the faster the increase of $F^{(2)}(x)$.

δ. Influence of the coupling parameter on the fixed points

We now discuss the influence of the parameter $\overline{g} \sim \overline{V}_0$ on the stability of single or multiple fixed points.

(i) The trivial case where $\overline{V}_0 = \overline{g} = 0$ (no interaction between the fermions) is particularly simple: the two curves are now identical horizontal lines, whose zero slope makes their intersection with the first bisector obviously stable. We then get the ideal gas results, with equal x_+ and x_- densities.

(ii) As long as $\overline{g}$ (hence $\overline{V}_0$) is weak enough, the slope of $F^{(1)}$ at the intersection point is small and the corresponding fixed point remains stable, as shown in Figure 1. This same point is obviously a fixed point for $F^{(2)}$ as well; as the derivative of a function $f[f(x)]$ with respect to x is $f'[f(x)]\,f'(x)$, that is $[f'(x)]^2$ at a point where $x = f(x)$, the slope of $F^{(2)}$ is less than 1, and that fixed point is also stable.

In such a case, both functions have only one common fixed point, which determines the only solution of the equations: the numbers x_+ and x_- are necessarily equal since they correspond to a fixed point of $F^{(1)}$. The only effect of the fermion repulsion is to lower in an equal way the densities associated with each of the spin states.

(iii) If now $\overline{g}$ (or $\overline{V}_0$) gets larger, we come to a situation, for a certain critical value g_c of g, where the slope of $F^{(1)}$ at the intersection with the first bisector becomes equal to -1, and that of $F^{(2)}$ now takes the value $+1$. The corresponding critical situation is plotted in Figure 2, where the curve representing the function $F^{(2)}$ is now tangent to the first bisector at their intersection (even osculating[3] to it, as their contact is of order two). For both functions, the fixed point is now right at the border of its stability domain.

[3]The first derivative of the function $f[f(x)]$ is equal to $f'(x) \times f'[f(x)]$, and its second derivative, to $f''(x)f'[f(x)]+[f'(x)]^2 f''[f(x)]$. At a fixed point, that second derivative becomes $f'(x)f''(x)[1 + f'(x)]$, which cancels out when $f'(x) = -1$.

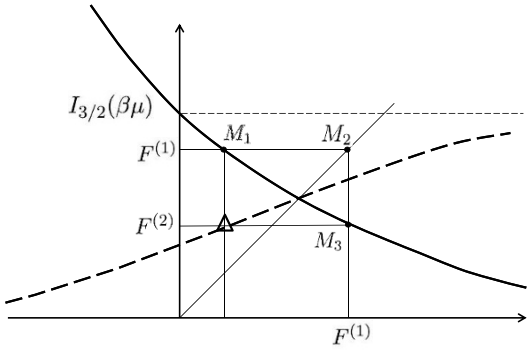


Figure 1: Plots of the functions $F^{(1)}$ (solid line) and $F^{(2)}$ (dashed line) as a function of x. Starting from any initial value of the variable x, we place a point M_1 on the first iterated $F^{(1)}$ curve, whose ordinate we transfer on the x axis by using the first bisector (point M_2); a new vertical intersection with the solid line (point M_3) yields the second iterated value $F^{(2)}$, that must be simply transferred to the initial value of the variable x (final point surrounded by a triangle). The whole dashed line curve can thus be constructed point by point. This method shows that when $x \to -\infty$, that curve is asymptotic to the abscissa axis; when $x \to +\infty$, the curve now has another horizontal asymptote, of ordinate $F^{(1)}(0) = I_{3/2}(\beta\mu)$, represented by a line with smaller dashes. The general form of the second iterated function is plotted in the figure: a uniformly increasing function between those two asymptotic values. In the case represented here, the coupling constant $\overline{g}$ is supposed to be weak enough for the two curves to intersect the first bisector with slopes of moduli less than 1; we then get a unique stable solution where x_- and x_+ are equal, which corresponds to a non-polarized spin system.

(iv) Beyond that situation, as shown in Figure 3, the curve representing $F^{(2)}$ intersects the first bisector in three points; the middle one is unstable as it corresponds to a slope larger than 1, but the two points on each sides are stable since they are associated with slopes between -1 and $+1$. As far as the central point is concerned, the slightest perturbation moves the iteration away from that point. On the other hand, the other two points are fixed stable points for $F^{(2)}$; they correspond to a physically acceptable solution of equations (29). As those two points are not fixed for $F^{(1)}$, two different values, x_+ and x_-, are swapped under the action of the function $F^{(1)}$ (two-cycle fixed points, represented by the arrows in the figure). We get a solution of the equations where the spin state populations are different: the gas develops a spontaneous polarization when the repulsion goes beyond a certain critical value and a phase transition occurs.

Comment:

For convenience, we discussed the emergence of the spontaneous polarization as a function of the parameter $\overline{g} \sim \overline{V}_0$, for fixed values of β and μ: the plot of the curves is then simply

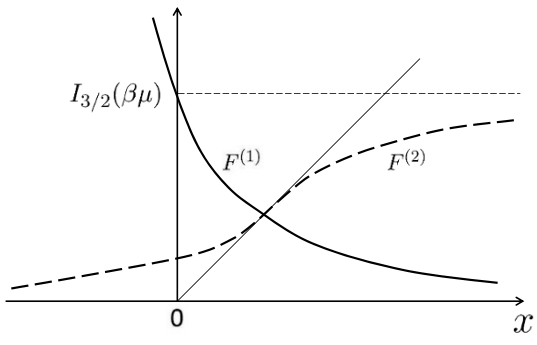


Figure 2: Plots of the functions $F^{(1)}$ (solid line) and $F^{(2)}$ (dashed line) in the critical case $g = g_c$, where the function $F^{(1)}$ intersects the first bisector with a slope equal to -1; the slope of $F^{(2)}$ is then equal to $+1$, and this function is not only tangent but also osculating to the first bisector (in other words it intersects this bisector at three points grouped together).

obtained by a scale change along the x axis. In general, however, the phase transition is observed by changing either the physical system density, or its temperature, while keeping the interactions constant. Our line of reasoning can be applied to this case, keeping the interactions constant while changing either the chemical potential μ that controls the particle density, or β, which is the inverse of the temperature. When either of these two parameters gets larger, the ordinate at the origin of $I_{3/2}(\beta\mu)$ gets higher, which increases the absolute values of the slopes of $F^{(1)}$ and $F^{(2)}$; the same phenomenon as above (instability and phase transition) will thus occur when the temperature is lowered or the density increased. If $I_{3/2}(\beta\mu) \gg 1$, relation (25) shows that the particle number contained in a volume $(\lambda_T)^3$ is large compared to one, which means that the average distance between the particles is smaller than the thermal wavelength; the fermion gas is then degenerate.

3-c. Physical discussion

As the spins carry a magnetic moment, a spontaneous polarization implies a transition towards a ferromagnetic phase. The origin of this phenomenon comes from an equilibrium between two opposite tendencies. On one hand, the "motor" of the transition comes from the fact that, to minimize the repulsion energy, the system tends to put all its particles in the same spin state (polarized system), which prevents them from interacting. This is because the Pauli principle forbids them to be at the same point in space, and as we assumed their interaction potential to be of zero-range, they can no longer interact. On the other hand, the system polarization (for a fixed total density) increases its kinetic energy: the same number of particles must be placed in a single Fermi sphere, instead of two, which results in a sphere with a larger radius, i.e. a higher Fermi level. It also changes the system entropy. The compromise between gain and loss (for

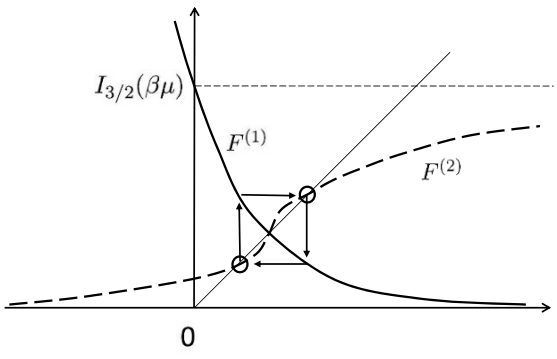


Figure 3: Beyond the critical point, the function $F^{(1)}$ intersects the first bisector with a slope less than -1, and there are now three distinct intersection points of the function $F^{(2)}$ and the bisector. The middle point corresponding to an $F^{(2)}$ slope larger than 1 is unstable, but the other two points (surrounded by circles) are stable. These two points are swapped under the action of $F^{(1)}$ (two-cycle fixed points, represented by the arrows). They yield different values for the spin densities, which leads to the appearance of a spontaneous spin polarization.

the grand potential) varies as a function of the parameters; when those parameters take a value where gain and loss balance each other perfectly, a spontaneous ferromagnetic transition occurs.

A more detailed study is possible; examining the shapes of the curves we plotted, we deduce that the conditions that favor the transition are: strong repulsion, high density, low temperature. It is worth noting that no Hamiltonian acting on the spins comes into play in this phase transition. Even though the interactions are totally independent of the spin, the Fermi-Dirac statistics has an effect on the spins, and can induce a transition polarizing those spins.

At the critical point (Figure 2), the two new stable points appear at the same place, and move away from each other in a continuous way. The phase transition is therefore continuous, which puts it into the category of second order phase transitions. The study of critical transitions is a very large domain of physics that we cannot discuss here in a general way. We can, however, take the analysis a little further, without too much difficulty: we note, from the equations written above, that the distance between the two stable points increases, beyond the critical point where $\mu = \mu_c$ and $\beta = \beta_c$, as the square root of the difference $\mu - \mu_c$ (or $\beta - \beta_c$). In other words, the system spontaneous magnetization varies as the square root of the distance to the critical point, which is typical of the so-called "Hopf bifurcation". In addition, at the critical point, the magnetic susceptibility of the spin system diverges.

Comments:

(i) A very general concept plays a role here: spontaneous symmetry breaking. The first symmetry breaking concerns the two opposed directions along the quantization axis

Oz. Equations (29) are invariant upon a permutation of x_+ and x_-; for any solution of these equations, there exists another one where these two variables are interchanged, and where the spin magnetization points in the opposite direction. This was to be expected since nothing physically distinguishes those two directions. The symmetry is said to be broken if the stable solutions of the set of equations are asymmetric, corresponding to different values of x_+ and x_-; there are then necessarily two (or more) distinct solutions, symmetric to one another.

Furthermore, the quantization axis Oz we used is arbitrary; had we chosen a different direction, we would have found that the spontaneous magnetization could point in any spatial direction. This again was to be expected since our problem is rotation invariant. The ferromagnetic transition phenomenon we have just studied corresponds to a spontaneous breaking of the rotational symmetry of the usual space, often called, in terms of group symmetry, "$SO(3)$ symmetry breaking". There are many other second order transitions that break various symmetry groups, as for example the symmetry $U(1)$ for the superfluid transition, etc.

(ii) A mean field theory like the one we used – i.e. an approximate theory – may identify the existence of a critical transition (second order transition) as explained above, but does not allow an exhaustive study of all its aspects, in particular in the vicinity of the critical point. Several critical phenomena (large wavelength critical fluctuations for example) cannot be accounted for with such an approximation, and require more elaborate theoretical methods.

4. Bosons: equation of state, attractive instability

For bosons, the equations (3) are very similar to those we used for fermions, except for a change of sign of η, and hence of the exchange potential. For a barely degenerate system, this modifies the interaction effects, but does not drastically change their consequences. On the other hand, for a system of degenerate bosons, the situation is radically different since expression (1) presents a singularity when $(e - \mu)$ is zero – whereas none occurs for fermions. As pointed out in Complement B_{XV}, this is the origin of the "Bose-Einstein condensation" phenomenon: as the chemical potential μ increases, the singularity becomes significant when μ gets close (through lower values) to the lowest individual energy among the e_i, that is close to the ground level energy e_0. The population of this level then increases more and more and can become "extensive" (proportional to the system volume $\mathcal{V}$ in the limit of large volumes).

Actually, using the Hartree-Fock equations for condensed boson systems leads to some difficulties, which will be briefly discussed below – see Comment (ii) of § 4-a. We shall limit ourselves to the study of non-condensed systems, not excluding the possibility that they approach condensation. We assume the bosons are without spin, and, as we did for fermions, that the range of the interaction potential $W_2(\mathbf{s})$ appearing in relation (11) is short enough so that:

$$\overline{V}_{(\mathbf{k}'-\mathbf{k})} = V_0 \tag{37}$$

In that case, the direct and exchange contributions in (13) are equal. For a homogeneous system, this equation then becomes:

$$\widetilde{e}_{\mathbf{k}} = e_k + \frac{2\overline{V}_0}{L^3} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}'} - \mu) = e_k + \frac{2\overline{V}_0}{L^3} \langle N \rangle \tag{38}$$

where $\langle N \rangle$ is the average total number of particles:

$$\langle N \rangle = \sum_{\mathbf{k}} f_\beta(\widetilde{e}_{\mathbf{k}} - \mu) = \sum_{\mathbf{k}} f_\beta(e_{\mathbf{k}} - \mu - \Delta\mu) \tag{39}$$

with:

$$\Delta\mu = -\frac{2\overline{V}_0}{L^3} \langle N \rangle \tag{40}$$

We therefore find that the average total number of particles is the same as for a boson gas without interactions, provided the chemical potential μ is replaced by an effective chemical potential $\widetilde{\mu} = \mu + \Delta\mu$. The same holds true for the average population of each individual state $\mathbf{k}$.

As in Complement B_{XV}, we note $N_{ig}^{BE}(\beta, \mu)$ the function yielding the particle number for an ideal gas of bosons:

$$N_{ig}^{BE}(\beta, \mu) = \sum_{\mathbf{k}} f_\beta(e_{\mathbf{k}} - \mu) = \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3k \frac{1}{e^{\beta(e_k - \mu)} - 1} \tag{41}$$

(the second equality is valid for large volumes). Equation (39) then becomes:

$$\langle N \rangle = N_{ig}^{BE}(\beta, \widetilde{\mu}) \equiv N_{ig}^{BE}(\beta, \mu + \Delta\mu) \tag{42}$$

4-a. Repulsive bosons

For repulsive interactions, Figure 4 shows how to graphically obtain, by a geometric construction, the system density predicted by equation (42). For a given chemical potential, the particle number decreases because of the repulsion, which takes the system further away from condensation; consequently, its description by the Hartree-Fock equations is a good approximation.

To first order in $\overline{V}_0$, relation (42) may be approximated by:

$$\begin{aligned} \langle N \rangle &\simeq N_{ig}^{BE}(\beta, \mu) + \Delta\mu \frac{\partial}{\partial\mu} N_{ig}^{BE}(\beta, \mu) \\ &= N_{ig}^{BE}(\beta, \mu) - 2\overline{V}_0 \frac{N_{ig}^{BE}(\beta, \mu)}{L^3} \frac{\partial}{\partial\mu} N_{ig}^{BE}(\beta, \mu) \end{aligned} \tag{43}$$

Noting $\Phi(\beta, \mu)$ the grand potential, relation (62) of Appendix VI shows that:

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial}{\partial\mu} \ln Z = -\frac{\partial}{\partial\mu} \Phi(\beta, \mu) \tag{44}$$

Integrating over μ relation (43) from $-\infty$ to the value μ, we get the grand potential:

$$\Phi(\beta, \mu) = \Phi_{ig}(\beta, \mu) + \frac{\overline{V}_0}{L^3} \left[N_{ig}^{BE}(\beta, \mu) \right]^2 \tag{45}$$

where $\Phi_{ig}(\beta, \mu)$ is the grand potential for the ideal gas, at the same temperature and chemical potential. In addition, relation (62) of Appendix VI shows that the grand potential is equal to minus the product of the volume and the pressure P:

$$\Phi(\beta, \mu) = -P\mathcal{V} \tag{46}$$

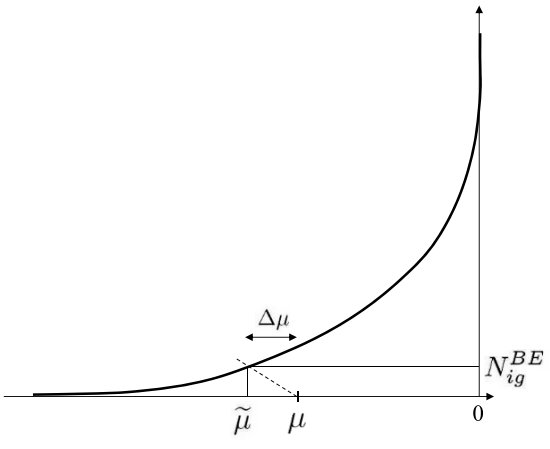


Figure 4: Geometric solutions of equations (40) and (42) for a gas of repulsive bosons. On the graph plotting the total number of particles as a function of the chemical potential, we draw a line starting from the point on the abscissa axis with chemical potential μ, and having the slope $-L^3/2\overline{V}_0$. It intersects the curve at a point whose abscissa is $\widetilde{\mu}$ and ordinate N_{ig}^{BE}. As $\overline{V}_0$ is positive for a repulsive gas, we see that the interactions lower the density (at constant temperature and chemical potential). The text explains how this geometric construction yields the equation of state for the interacting gas.

This means that if, at constant β, we vary the parameter μ in (43) and (45), we obtain in the plane $(\langle N\rangle, P)$ a curve representing the pressure as a function of the particle number in the volume $\mathcal{V}$, which is an isothermal line of the equation of state. Repeating this plot for several values of β, we get a set of curves covering the whole equation of state, taking into account the changes introduced by the interactions.

Comments:

(i) To keep the computations as simple as possible, we limited ourselves to the first order in $\overline{V}_0$. It is however possible to make the graphical construction of Fig. 4 more precise by including the higher order terms.

(ii) We discussed in § 3-b-β of Complement $\mathrm{G_{XV}}$ the limits of the Hartree-Fock approximation for bosons, which can no longer be used when the physical system gets too close to Bose-Einstein condensation. The graphical construction shown in Figure 4 loses its physical meaning if the intersection point on the curve is too close to the contact point of the curve with the vertical axis.

4-b. Attractive bosons

Attractive interactions ($\overline{V}_0 < 0$) result in an increase of the effective chemical potential, and consequently raises the value of $\langle N\rangle$. This in turn increases the effective chemical potential, and this positive feedback may even induce an avalanche effect leading to an instability if μ is too close to zero.

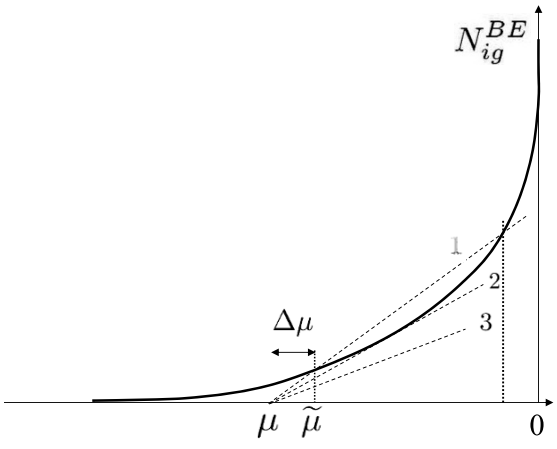


Figure 5: Graphical construction similar to that of Figure 4, but for an attractive boson gas (where $\overline{V}_0$ is negative). When the attractive potential $\overline{V}_0$ is not too large, the line noted 1 in the figure yields two possible solutions, only one of which is close to the solution in the absence of interactions, and hence suitable for our approximation. As $\overline{V}_0$ increases, for a certain critical value we only get one solution (tangent line 2), then none (line 3). In this last case, no solution signals an instability of the gas, which collapses onto itself because of the attractive interactions. Starting from a nearly condensed ideal gas, the closer it is to condensation, the weaker the attractive interactions necessary to trigger the instability.

The geometrical construction that yields $\Delta\mu$ and $\langle N \rangle$ from the intersection of a straight line with a curve is shown on Fig. 5. If $\overline{V}_0$ is weak enough, and for a fixed value of μ, we get two intersection points, corresponding to possible solutions. We only keep the first one, yielding the lowest value of $\Delta\mu$. The other point yields a high value of $\Delta\mu$, which changes radically and increases considerably the system density; in that case, chances are the approximate mean field treatment of the interactions is no longer valid. Beyond the value of $\overline{V}_0$ for which the straight line becomes tangent to the curve, the couple of equations (39) and (40) do not have a solution: there no longer exists any stable solution.

Figure 5 also shows that as the chemical potential gets closer to zero, the effect of the attractive interactions between bosons becomes more and more important; weak interactions are enough to render the system unstable. The reason we did not find any solution to the equations is that we assumed, in the computations, that the system was perfectly homogeneous; now this homogeneity cannot be maintained beyond a certain attraction intensity. We must therefore enlarge the theoretical framework, and include the possibility for the system to become spontaneously inhomogeneous. A more precise study would show that the system may develop local instabilities, hence breaking spontaneously the translation invariance symmetry. In the limit of large systems (thermodynamic limit), condensed bosons tend to collapse onto themselves under the effect

of an attractive interaction, however weak it may be[4].

As a general conclusion, the Hartree-Fock method applied to fermions yields results valid in a very large parameter range. As an example, it allowed computing effects of the interactions on the particle number and the pressure of the system. In addition, this method was able to predict the existence of phase transitions. This is also true for non-degenerate bosons, and the mean field method actually has a very large number of applications that we cannot detail here. We must, however, keep in mind that when Bose-Einstein condensation occurs, certain predictions pertaining to the condensate may not be realistic from a physical point of view, as they depend too closely on the mean field approximation which does not properly account for the correlations between particles.

[4]If the interaction potential is attractive at large distance, but strongly repulsive at short range (hard core for example), the system spontaneously forms a high density liquid or solid.

Chapter XVI

Field operator

A Definition of the field operator 1752
A-1 Definition 1752
A-2 Commutation and anticommutation relations 1754
B Symmetric operators 1755
B-1 General expression 1755
B-2 Simple examples 1756
B-3 Field spatial correlation functions 1758
B-4 Hamiltonian operator 1761
C Time evolution of the field operator (Heisenberg picture) . 1763
C-1 Contribution of the kinetic energy 1763
C-2 Contribution of the potential energy 1764
C-3 Contribution of the interaction energy 1764
C-4 Global evolution 1765
D Relation to field quantization 1765

Introduction

This chapter is a continuation of the previous chapter and uses the same mathematical tools. The main difference is that, until now, we have mainly used discrete bases in the individual state space, $\{|u_i\rangle\}$ or $\{|v_s\rangle\}$. In this chapter, we shall use a continuous basis, which is the basis, for spinless particles, of the position eigenvectors (see Chapter II, § E). As they now depend on the position $\mathbf{r}$, the creation and annihilation operators become field operators depending on a continuous subscript $\mathbf{r}$. They are the operator analog of the classical fields (which are numbers and not operators), and are often called "field operators". They are useful for concisely describing numerous properties of identical particle systems. They have commutation relations for bosons, and anticommutation relations for fermions. This chapter is a preparation for Chapters XIX and XX, where we will introduce the quantization of the electromagnetic field.

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

After defining these operators in § A-1, we discuss some of their properties. Their commutation and anticommutation relations are examined in § A-2. As in Chapter XV, we then study (§ B) symmetric operators and their expression as a function of the field operators; special attention will be given to the operators associated with the field correlation functions. In § C, we shall use the Heisenberg picture to study the time dependence of these operators. As a conclusion, we shall briefly come back (in § D) to the field quantization procedure and its link to the concept of identical particles.

A. Definition of the field operator

The field operator is defined as the annihilation operator a_i, but associated with a basis of individual states labeled by the continuous position index $\mathbf{r}$ instead of a discrete index i. Our starting point will be the basis change relation (A-52) of Chapter XV:

$$a_{v_s} = \sum_i \langle v_s |\ u_i \rangle\ \ a_{u_i} \tag{A-1}$$

where the subscripts i and s label the kets of the two orthonormal bases $\{|u_i\rangle\}$ and $\{|v_s\rangle\}$ in the individual state space. In what follows, and as already done in Chapter XV, we will simplify the notation and often replace the subscript u_i by i, and the subscript v_s by s.

A-1. Definition

We will first define the field operator for spinless particles, and later generalize it to the case with spin.

A-1-a. Spinless particles

We replace, in relation (A-1), the basis $\{|v_s\rangle\}$ by the basis of vectors $\{|\mathbf{r}\rangle\}$, where $\mathbf{r}$ symbolizes three continuous indices (the three vector components). The operator a_s now becomes an operator depending on the continuous index $\mathbf{r}$, that we shall call "*field operator*" for the system of identical particles we consider. We could write it simply as $a_{\mathbf{r}}$, but we shall prefer the commonly used classical notation $\Psi(\mathbf{r})$. Like any annihilation operator, $\Psi(\mathbf{r})$ acts in the Fock space where it lowers by one unit the particle number. In (A-1), the coefficient appearing in the sum is now the wave function $u_i(\mathbf{r})$ associated with the ket $|u_i\rangle$:

$$u_i(\mathbf{r}) = \langle \mathbf{r} |\ u_i \rangle \tag{A-2}$$

and this relation becomes:

$$\boxed{\Psi(\mathbf{r}) = \sum_i u_i(\mathbf{r})\ a_i} \tag{A-3}$$

Formally, definition (A-3) looks like the expansion of a wave function on the basis functions $u_i(\mathbf{r})$; here, however, the "components" a_i are operators and no longer simple complex numbers. In the same way as a_s annihilates a particle in the state $|v_s\rangle$, the operator $\Psi(\mathbf{r})$ now annihilates a particle at point $\mathbf{r}$.

It does not depend on the basis $\{|u_i\rangle\}$, i.e. on the wave functions chosen to define it in (A-3), as we now show. We insert in this equality the closure relation in any arbitrarily chosen basis $\{|v_s\rangle\}$, and use again (A-1):

$$\Psi(\mathbf{r}) = \sum_i \sum_s \langle \mathbf{r}|\ v_s\rangle\ \langle v_s|\ u_i\rangle\ \ a_{u_i} = \sum_s \langle \mathbf{r}|\ v_s\rangle\ \ a_{v_s} \tag{A-4}$$

(we temporarily came back to the explicit notation for the annihilation operators). We can thus write:

$$\Psi(\mathbf{r}) = \sum_s v_s(\mathbf{r})\ a_{v_s} \tag{A-5}$$

which means $\Psi(\mathbf{r})$ satisfies, in the new basis, a relation similar to (A-3).

We now take the Hermitian conjugate of (A-3):

$$\Psi^\dagger(\mathbf{r}) = \sum_i u_i^*(\mathbf{r})\ a_i^\dagger \tag{A-6}$$

The operator $\Psi^\dagger(\mathbf{r})$ creates a particle at point $\mathbf{r}$, as can be shown for example by computing the ket resulting from its action on the vacuum:

$$\Psi^\dagger(\mathbf{r})\,|0\rangle = \sum_i u_i^*(\mathbf{r})\ a_i^\dagger\,|0\rangle = \sum_i u_i^*(\mathbf{r})\ \ |u_i\rangle \tag{A-7}$$

that is:

$$\Psi^\dagger(\mathbf{r})\,|0\rangle = \sum_i\ \ |u_i\rangle\,\langle u_i|\ \mathbf{r}\rangle = |\mathbf{r}\rangle \tag{A-8}$$

which represents, as announced, a particle localized at point $\mathbf{r}$.

One can easily invert formulas (A-3) and (A-6) by writing, for example:

$$\begin{aligned}\int \mathrm{d}^3r\ u_i^*(\mathbf{r})\ \Psi(\mathbf{r}) = &\int \mathrm{d}^3r\ u_i^*(\mathbf{r})\ \sum_j u_j(\mathbf{r})\ a_j \\ &= \sum_j \delta_{ij}\ a_j\ =\ a_i\end{aligned} \tag{A-9}$$

or else, by Hermitian conjugation:

$$\int \mathrm{d}^3r\ u_i(\mathbf{r})\ \Psi^\dagger(\mathbf{r}) = a_i^\dagger \tag{A-10}$$

A-1-b. Particles with spin

When the particles have a spin S, the basis vector $\{|\mathbf{r}\rangle\}$ used above must be replaced by the basis vector $\{|\mathbf{r},\nu\rangle\}$, where ν is the spin index, which can take $2S+1$ discrete values ($\nu = -S,\ -S+1,\ ...,+S$). To all the summations over d^3r, we must now add a summation on the $2S+1$ values of the spin index ν. As an example, a basis vector $|u_i\rangle$ in the individual state space must now be written:

$$|u_i\rangle = \sum_{\nu=-S}^{+S} \int \mathrm{d}^3r\ u_i(\mathbf{r},\nu)\,|\mathbf{r},\nu\rangle \tag{A-11}$$

with:

$$u_i(\mathbf{r},\nu) = \langle \mathbf{r},\nu \,|u_i\rangle \tag{A-12}$$

The variables $\mathbf{r}$ and ν play a similar role. The first one is, however, continuous, whereas the second is discrete. Writing them in the same parenthesis might hide this difference and one often prefers noting the discrete index as a superscript of the function u_i, writing for example:

$$u_i^{\nu}(\mathbf{r}) = \langle \mathbf{r},\nu \,|u_i\rangle \tag{A-13}$$

Let us again use relation (A-1). On the left-hand side, the index v_s now symbolizes both the position $\mathbf{r}$ and the spin quantum number ν, which leads us to define a field operator $\Psi_\nu(\mathbf{r})$ having $2S+1$ spin components. Inserting (A-13) in the right-hand side of (A-1), we get:

$$\boxed{\Psi_\nu(\mathbf{r}) = \sum_i u_i^{\nu}(\mathbf{r})\; a_i} \tag{A-14}$$

The Hermitian conjugate operator $\Psi_\nu^\dagger(\mathbf{r})$ now creates a particle at point $\mathbf{r}$ with a spin ν:

$$\Psi_\nu^\dagger(\mathbf{r})\,|0\rangle = \sum_i \langle u_i|\;\mathbf{r},\nu\rangle \;\; |u_i\rangle = |\mathbf{r},\nu\rangle \tag{A-15}$$

As we did above, we can invert those relations. In relation (A.51) of Chapter XV (basis change), we replace v_s by i, and u_i by $\mathbf{r},\nu$ (which means that the summation over i is replaced by an integral over d^3r and a summation over ν), and use equality (A-11); we therefore get:

$$a_i^\dagger = \sum_{\nu=-S}^{+S} \int \mathrm{d}^3r \; u_i^{\nu}(\mathbf{r}) \; \Psi_\nu^\dagger(\mathbf{r}) \tag{A-16}$$

which is the analog, in the presence of spin, of relation (A-10).

A-2. Commutation and anticommutation relations

Commutation relations for field operators are analogous to those obtained in § A-5 of Chapter XV, but the discrete index i is now replaced by a continuous index.

A-2-a. Spinless particles

The commutator (or anticommutator) of two field operators:

$$[\Psi(\mathbf{r}) \;, \Psi(\mathbf{r}')]_{-\eta} = \sum_{i,j} u_i(\mathbf{r}) u_j(\mathbf{r}') \,[a_i, a_j]_{-\eta} = 0 \tag{A-17}$$

is indeed zero, as expected from expression (A-48) of Chapter XV. In the same way, by Hermitian conjugation:

$$\left[\Psi^\dagger(\mathbf{r}) \;, \Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \sum_{i,j} u_i^*(\mathbf{r}) u_j^*(\mathbf{r}') \left[a_i^\dagger, a_j^\dagger\right]_{-\eta} = 0 \tag{A-18}$$

However, when we (anti)commute the field operator and the adjoint operator, we get:

$$\left[\Psi(\mathbf{r})\ ,\Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \sum_{i,j} u_i(\mathbf{r})u_j^*(\mathbf{r}')\left[a_i, a_j^\dagger\right]_{-\eta} \tag{A-19}$$

which yields, taking into account the commutation relations (A-49) of Chapter XV:

$$\left[\Psi(\mathbf{r})\ ,\Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \sum_{i} u_i(\mathbf{r})u_i^*(\mathbf{r}') = \sum_{i} \langle\mathbf{r}|\ u_i\rangle\langle u_i|\ \mathbf{r}'\rangle = \langle\mathbf{r}|\ \mathbf{r}'\rangle \tag{A-20}$$

Finally, we get:

$$\boxed{\left[\Psi(\mathbf{r})\ ,\Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \delta(\mathbf{r}-\mathbf{r}')} \tag{A-21}$$

which is the equivalent of the relations (A-49) of Chapter XV in the case of a continuous basis.

A-2-b. Particles with spin

Relation (A-17) now becomes:

$$[\Psi_\nu(\mathbf{r})\ ,\Psi_{\nu'}(\mathbf{r}')]_{-\eta} = \sum_{i,j} u_i^\nu(\mathbf{r})u_j^{\nu'}(\mathbf{r}')\,[a_i, a_j]_{-\eta} = 0 \tag{A-22}$$

Relation (A-18) remains valid even if the field operators have spin indices. Finally, relation (A-19) becomes:

$$\begin{aligned}\left[\Psi_\nu(\mathbf{r})\ ,\Psi_{\nu'}^\dagger(\mathbf{r}')\right]_{-\eta} &= \sum_{i,j} u_i(\mathbf{r},\nu)u_j^*(\mathbf{r}',\nu')\left[a_i, a_j^\dagger\right]_{-\eta} \\ &= \langle\mathbf{r},\nu|\ \mathbf{r}',\nu'\rangle\end{aligned} \tag{A-23}$$

that is:

$$\boxed{\left[\Psi_\nu(\mathbf{r})\ ,\Psi_{\nu'}^\dagger(\mathbf{r}')\right]_{-\eta} = \delta_{\nu\nu'}\delta(\mathbf{r}-\mathbf{r}')} \tag{A-24}$$

B. Symmetric operators

In the previous chapter, we wrote one- or two-particle symmetric operators in terms of creation and annihilation operators in the discrete states $|u_i\rangle$. We are now going to express those operators in terms of the field operator (and its Hermitian conjugate).

B-1. General expression

We start with spinless particles. We can either directly transpose expressions (B-12) and (C-16) of Chapter XV to a continuous basis $\{|\mathbf{r}\rangle\}$ (replacing the sums by integrals), or insert in those expressions the form (A-9) for the operators a_i. In both cases, we get:

$$\widehat{F} = \int \mathrm{d}^3r \int \mathrm{d}^3r'\,\langle\mathbf{r}|\,\widehat{f}\,|\mathbf{r}'\rangle\ \ \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}') \tag{B-1}$$

and:

$$\widehat{G} = \frac{1}{2}\int \mathrm{d}^3 r \int \mathrm{d}^3 r' \int \mathrm{d}^3 r'' \int \mathrm{d}^3 r''' \\ \langle 1:\mathbf{r}; 2:\mathbf{r}'|\,\widehat{g}\,|1:\mathbf{r}''; 2:\mathbf{r}'''\rangle \;\; \Psi^\dagger(\mathbf{r})\Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}''')\Psi(\mathbf{r}'') \tag{B-2}$$

where, as in relation (C-16) of Chapter XV, the order of the annihilation operators is the inverse of the order that appears in the ket of the matrix element.

Expression (B-1) reminds us of the average value $\left\langle \widehat{f}(1) \right\rangle$ of the operator $\widehat{f}(1)$ for a single particle (without spin), described by the wave function $\psi(\mathbf{r}_1)$:

$$\left\langle \widehat{f}(1) \right\rangle = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_1' \, \langle \mathbf{r}_1|\, \widehat{f}\,|\mathbf{r}_1'\rangle \;\; \psi^*(\mathbf{r}_1)\psi(\mathbf{r}_1') \tag{B-3}$$

Both expressions are not equivalent since (B-1) concerns any number of identical particles, rather than a single one; furthermore, the Ψ are now operators, and their respective order matters – as opposed to the order of the ψ in (B-3). As for formula (B-2), it can be compared to the average value of an operator $\widehat{g}(1,2)$ acting on two particles, 1 and 2, both described by the same wave function ψ. Here again, the order of the field operators is important, as opposed to the order in a product of wave functions.

For particles with spins, we simply complement each integral over $\mathbf{r}$ with a sum over the spin index ν, we include this index in the matrix elements and add a spin index to the field operator. As an example, relation (B-1) is generalized to:

$$\widehat{F} = \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \sum_{\nu=-S}^{S} \sum_{\nu'=-S}^{S} \langle \mathbf{r}, \nu|\, f\,|\mathbf{r}', \nu'\rangle \;\; \Psi_\nu^\dagger(\mathbf{r})\Psi_{\nu'}(\mathbf{r}') \tag{B-4}$$

As for relation (B-2), we get four summations over the spin indices ν, and the operator matrix elements are taken between bras and kets where an index ν is added to the variable $\mathbf{r}$.

B-2. Simple examples

We start with a few examples concerning operators for one spinless particle. For a single particle, the operator associated with the local density at point $\mathbf{r}_0$ is:

$$|\mathbf{r}_0\rangle \, \langle \mathbf{r}_0| \tag{B-5}$$

and its matrix elements are:

$$\langle \mathbf{r}|\, \mathbf{r}_0\rangle \, \langle \mathbf{r}_0 \,|\mathbf{r}'\rangle = \delta(\mathbf{r} - \mathbf{r}_0) \;\; \delta(\mathbf{r}' - \mathbf{r}_0) \tag{B-6}$$

The corresponding N-particle operator is written as:

$$\widehat{D}^{(N)}(\mathbf{r}_0) = \sum_{q=1}^{N} |q:\mathbf{r}_0\rangle \, \langle q:\mathbf{r}_0| \tag{B-7}$$

Replacing $\widehat{f}$ by $|\mathbf{r}_0\rangle \, \langle \mathbf{r}_0|$ in (B-1) yields the operator acting in the Fock space:

$$\widehat{D}(\mathbf{r}_0) = \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}_0) \tag{B-8}$$

This operator annihilates a particle at point $\mathbf{r}_0$, and immediately recreates it at the same point. The average value:

$$\widehat{D}(\mathbf{r}_0) = \langle\Phi|\,\widehat{D}(\mathbf{r}_0)\,|\Phi\rangle \tag{B-9}$$

in a normalized state $|\Phi\rangle$ of the N-particle system yields the particle density associated with this state at point $\mathbf{r}_0$.

The operator $\widehat{N}$, total number of particles, has been written in (B-15) of Chapter XV. As the discrete summation index i is changed to a continuous index $\mathbf{r}$, the summation becomes an integral over all space:

$$\widehat{N} = \int \mathrm{d}^3r\ \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) \tag{B-10}$$

As expected, it is the integral over d^3r of the operator $\widehat{D}(\mathbf{r})$.

The operator $V_1(\mathbf{r})$ describing the one-particle potential energy is also diagonal in the position representation; in the Fock space, it becomes the operator $\widehat{V}_1$:

$$\widehat{V}_1 = \int \mathrm{d}^3r\ V_1(\mathbf{r})\Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) \tag{B-11}$$

As for the particle current, it can be deduced from the expression of the current $\widehat{\mathbf{j}}(\mathbf{r}_0)$ associated with a single particle of mass m; it is the product of the operators giving the local density $|\mathbf{r}_0\rangle\langle\mathbf{r}_0|$ and the velocity $\widehat{\mathbf{p}}/m$ (product that must obviously be symmetrized):

$$\widehat{\mathbf{j}}(\mathbf{r}_0) = \frac{1}{2}\left[|\mathbf{r}_0\rangle\langle\mathbf{r}_0|\,\frac{\widehat{\mathbf{p}}}{m} + \frac{\widehat{\mathbf{p}}}{m}\,|\mathbf{r}_0\rangle\langle\mathbf{r}_0|\right] \tag{B-12}$$

If the particle is described by a wave function $\psi(\mathbf{r})$, a simple calculation[1] shows that the average value of this operator yields the usual expression for the probability current – see equation (D-17) of Chapter III:

$$\mathbf{J}(\mathbf{r}_0) = \frac{\hbar}{2mi}\left[\psi^*(\mathbf{r}_0)\boldsymbol{\nabla}\psi(\mathbf{r}_0) - \psi(\mathbf{r}_0)\boldsymbol{\nabla}\psi^*(\mathbf{r}_0)\right] \tag{B-14}$$

The current $\mathbf{J}(\mathbf{r}_0)$ of a system of identical particles is obtained by replacing in (B-1) the operator $\widehat{f}$ by $\widehat{\mathbf{j}}(\mathbf{r}_0)$:

$$\widehat{\mathbf{J}}(\mathbf{r}_0) = \frac{\hbar}{2mi}\left[\Psi^\dagger(\mathbf{r}_0)\boldsymbol{\nabla}\Psi(\mathbf{r}_0) - \Psi(\mathbf{r}_0)\boldsymbol{\nabla}\Psi^\dagger(\mathbf{r}_0)\right] \tag{B-15}$$

Another way of arriving at this equality is to use the substitution process mentioned in § B-1. To obtain the operator we are looking for in terms of $\Psi(\mathbf{r})$, we start from the

[1] Let us calculate the average value $\langle\Psi|\,\widehat{\mathbf{j}}(\mathbf{r}_0)\,|\Psi\rangle$. The first term on the right-hand side of (B-12) yields:

$$\frac{1}{2m}\,\langle\Psi\ |\mathbf{r}_0\rangle\langle\mathbf{r}_0|\,\widehat{\mathbf{p}}\,|\boldsymbol{\Psi}\rangle = \frac{\hbar}{2mi}\Psi^*(\mathbf{r}_0)\boldsymbol{\nabla}\psi(\mathbf{r}_0) \tag{B-13}$$

since the action of the operator $\widehat{\mathbf{p}}$ in the position representation is given by $(\hbar/i)\boldsymbol{\nabla}$. The second term is its complex conjugate, and we therefore get (B-14).

expression of the average value for a single particle, described by the wave function $\psi(\mathbf{r})$, which is then replaced by the field operator $\Psi(\mathbf{r})$.

For particles with spin, the local density at point $\mathbf{r}_0$, with spin ν, is written in the same way:

$$\widehat{D}^{(N)}(\mathbf{r}_0,\nu) = \sum_{q=1}^{N} |q:\mathbf{r}_0,\nu\rangle \langle q:\mathbf{r}_0,\nu| \tag{B-16}$$

and yields, in the Fock space, the operator:

$$\widehat{D}(\mathbf{r}_0,\nu) = \Psi_\nu^\dagger(\mathbf{r}_0)\Psi_\nu(\mathbf{r}_0) \tag{B-17}$$

The total density is obtained by a summation over ν:

$$\widehat{D}(\mathbf{r}_0) = \sum_{\nu=-S}^{S} \Psi_\nu^\dagger(\mathbf{r}_0)\Psi_\nu(\mathbf{r}_0) \tag{B-18}$$

For particles with spin, the operator $\widehat{N}$ associated with the total particle number, or the operator probability current $\widehat{\mathbf{J}}_{\mathbf{r}_0}$, can be obtained in a similar way.

B-3. Field spatial correlation functions

Operators associated with spatial correlation functions can also be defined with field operators; their average values are very useful for characterizing the field properties at different points of space. When we reason in terms of fields, we generally characterize each correlation function by the number of points concerned, which is different from the number of particles involved: the two-point functions characterize the properties of a single particle, the four-point ones concern two particles, etc. The reason is simple: a one-particle density operator ρ_I is characterized by non-diagonal elements $\langle \mathbf{r}_0'| \rho_I |\mathbf{r}_0\rangle$ depending on two positions, a two-particle density operator ρ_{II} involves elements depending on four positions, etc.

B-3-a. Two-point correlation functions

Defining a non-diagonal operator depending on two parameters $\mathbf{r}_0$ and $\mathbf{r}_0'$, we can generalize relation (B-5):

$$|\mathbf{r}_0\rangle \langle \mathbf{r}_0'| \tag{B-19}$$

A calculation very similar to the one leading to equation (B-7) – we simply add a "prime" to the second $\mathbf{r}_0$ – gives the N-particle symmetric operator:

$$\widehat{D}^{(N)}(\mathbf{r}_0,\mathbf{r}_0') = \sum_{q=1}^{N} |q:\mathbf{r}_0\rangle \langle q:\mathbf{r}_0'| \tag{B-20}$$

which yields, in the Fock space, the operator:

$$\widehat{D}(\mathbf{r}_0,\mathbf{r}_0') = \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}_0') \tag{B-21}$$

We thus obtain an operator that annihilates a particle at point $\mathbf{r}'_0$ and recreates it at a different point $\mathbf{r}_0$.

When the N-particle system is described by a quantum state $|\Phi\rangle$, we call *two-point field correlation function*, $G_1(\mathbf{r}_0, \mathbf{r}'_0)$, the average value:

$$G_1(\mathbf{r}_0, \mathbf{r}'_0) = \left\langle \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}'_0) \right\rangle = \langle\Phi|\, \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}'_0)\, |\Phi\rangle \tag{B-22}$$

which also yields the matrix element, in the $\{|\mathbf{r}\rangle\}$ representation, of the one-particle operator[2] ρ_I:

$$\left\langle \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}'_0) \right\rangle = \langle \mathbf{r}'_0|\, \rho_I\, |\mathbf{r}_0\rangle \tag{B-23}$$

Demonstration:

The matrix elements of the density operator $\rho_I^{(q)}$ of particle q are:

$$\left\langle \mathbf{r}'_0 \right| \rho_I^{(q)} \left| \mathbf{r}_0 \right\rangle = \mathrm{Tr}\left\{ \rho_I^{(q)} \left| q : \mathbf{r}_0 \right\rangle \left\langle q : \mathbf{r}'_0 \right| \right\} = \left\langle \left| q : \mathbf{r}_0 \right\rangle \left\langle q : \mathbf{r}'_0 \right| \right\rangle \tag{B-24}$$

For an N-particle system, we define the one-particle density operator ρ_I by a sum over all the particles:

$$\rho_I = \sum_{q=1}^{N} \rho_I^{(q)} \tag{B-25}$$

(be careful: the trace of this operator is N, not 1). Its matrix elements are the sum of the average values written in (B-24), i.e. the average value of the one-particle symmetric operator obtained by summing over q the $|q : \mathbf{r}_0\rangle \langle q : \mathbf{r}'_0|$. This result is simply the operator $\widehat{D}^{(N)}(\mathbf{r}_0, \mathbf{r}'_0)$ of (B-20), which, as seen above, yields in the Fock space expression (B-21). We then simply take the average value of each side of this expression to get equality (B-23).

The average value (B-22) at two different points plays an important role in the study of Bose-Einstein condensation. For a system at thermal equilibrium, this average value generally tends to zero rapidly as the distance between $\mathbf{r}$ and $\mathbf{r}'$ increases; it only remains non-zero in a domain of microscopic size. However, for a Bose-Einstein condensed gas, this average value behaves quite differently as it tends toward a non-zero value at large distance. This difference is actually the "Penrose and Onsager condensation criterion "; they have defined the existence of such a condensation as the appearance of a non-zero value of the matrix element of ρ_I at large distance; this definition is quite general as it applies not only to the ideal gas but also to systems of interacting particles.

Particles with spin:

If the particles have a non-zero spin S, we use, as a basis, the kets $|\mathbf{r}, \nu\rangle$ where ν takes the $(2S+1)$ values $-S$, $-S+1$, .., $+S$, and we add a ν index to the field operators. We then define $(2S+1)^2$ two-point field correlation functions as the average values:

$$\left\langle \Psi^\dagger_\nu(\mathbf{r}_0)\Psi_{\nu'}(\mathbf{r}'_0) \right\rangle = G_1(\mathbf{r}_0, \nu; \mathbf{r}'_0, \nu') \tag{B-26}$$

[2]Note the inversion of the variable order between the function G_1 (or the variables of $\Psi^\dagger$ and Ψ) and the matrix element of ρ_I.

The same computation that led to (B-23) for spinless particles, can be repeated with no changes other than the simple replacement of the kets (or bras) $|\mathbf{r}\rangle$ by $|\mathbf{r},\nu\rangle$; it shows that these average values yield the matrix elements of the one-particle density operator:

$$\left\langle \mathbf{r}_0',\nu' \right| \rho_I \left| \mathbf{r}_0,\nu \right\rangle = \left\langle \Psi_\nu^\dagger(\mathbf{r}_0)\Psi_{\nu'}(\mathbf{r}_0') \right\rangle \tag{B-27}$$

(here again we have chosen to normalize to N the trace of the one-particle density operator).

B-3-b. Higher order correlation functions

One can also start with the two-particle operator, which now depends on four positions:

$$\widehat{g}(1:\mathbf{r}_0,\mathbf{r}_0';2:\mathbf{r}_0'',\mathbf{r}_0''') = |1:\mathbf{r}_0\rangle\langle 1:\mathbf{r}_0'| \otimes |2:\mathbf{r}_0''\rangle\langle 2:\mathbf{r}_0'''| \tag{B-28}$$

In this case, the expression of $\widehat{g}$ is not symmetric with respect to the exchange of particles 1 and 2, as opposed to what happens for an interaction energy. The operator $\widehat{G}$ is then defined without the 1/2 factor of relation (C-1) of Chapter XV:

$$\widehat{G}^{(N)}(\mathbf{r}_0,\mathbf{r}_0',\mathbf{r}_0'',\mathbf{r}_0''') = \sum_{q,q'=1;\ q\neq q'}^{N} \widehat{g}(q:\mathbf{r}_0,\mathbf{r}_0';q':\mathbf{r}_0'',\mathbf{r}_0''') \tag{B-29}$$

and yields in the Fock space the operator $\widehat{G}(\mathbf{r}_0,\mathbf{r}_0',\mathbf{r}_0'',\mathbf{r}_0''')$. Relation (B-2), without this factor 1/2, then leads to:

$$\widehat{G}(\mathbf{r}_0,\mathbf{r}_0',\mathbf{r}_0'',\mathbf{r}_0''') = \Psi^\dagger(\mathbf{r}_0)\Psi^\dagger(\mathbf{r}_0'')\Psi(\mathbf{r}_0''')\Psi(\mathbf{r}_0') \tag{B-30}$$

In this case, the operator annihilates two particles at two points and recreates them at two others.

A computation very similar to the one leading to (B-22) and (B-23) enables us to show, using (B-2), that the matrix elements of the two-particle density operator ρ_{II} can be written[3]:

$$\langle 1:\mathbf{r}_0';2:\mathbf{r}_0'''|\ \rho_{II}\ |1:\mathbf{r}_0;2:\mathbf{r}_0''\rangle = \left\langle \Psi^\dagger(\mathbf{r}_0)\Psi^\dagger(\mathbf{r}_0'')\Psi(\mathbf{r}_0''')\Psi(\mathbf{r}_0') \right\rangle \tag{B-31}$$

This density operator, whose trace is equal to $N(N-1)$, plays an essential role in the study of correlations between particles.

A particularly important example of a higher order correlation function corresponds to the case where $\mathbf{r}_0 = \mathbf{r}_0'$ and $\mathbf{r}_0'' = \mathbf{r}_0'''$. We then get:

$$\begin{aligned}\widehat{G}^{(N)}(\mathbf{r}_0,\mathbf{r}_0,\mathbf{r}_0'',\mathbf{r}_0'') &= \sum_{q,q'=1;\ q\neq q'}^{N} |q:\mathbf{r}_0\rangle\langle q:\mathbf{r}_0| \otimes |q':\mathbf{r}_0''\rangle\langle q':\mathbf{r}_0''| \\ &= \sum_{q,q'=1;\ q\neq q'}^{N} |q:\mathbf{r}_0;\ q':\mathbf{r}_0''\rangle\langle q:\mathbf{r}_0;\ q':\mathbf{r}_0''|\end{aligned} \tag{B-32}$$

[3] One can also use relation (C-19) of Chapter XV to get the same result.

which yields in the Fock space the operator:

$$\widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') = \Psi^\dagger(\mathbf{r}_0)\Psi^\dagger(\mathbf{r}_0'')\Psi(\mathbf{r}_0'')\Psi(\mathbf{r}_0) \tag{B-33}$$

The expression on the right-hand side of (B-32) characterizes the probability of finding any particle at $\mathbf{r}_0$ and any other one at $\mathbf{r}_0''$. In the same way as the average value (B-9) gives the one-particle density, the average value:

$$\left\langle \widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') \right\rangle = \langle \Phi | \, \widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') \, | \Phi \rangle = G_2(\mathbf{r}_0, \mathbf{r}_0'') \tag{B-34}$$

gives the two-particle "double density", which contains information on all the binary correlations between the particle positions.

We are now in a position to again obtain expression (C-28) of Chapter XV, and more precisely justify the interpretation we gave of the average value of the interaction energy written in (C-27) of that chapter. We replace in (B-33) the field operators (or their adjoints) by their expansion (A-3) on the operators a_i (or the $a_i^\dagger$); we then get (C-28) of Chapter XV, $\mathbf{r}_0$ being replaced by $\mathbf{r}_1$ and $\mathbf{r}_0'$ by $\mathbf{r}_2$.

If $\mathbf{r}_0 \neq \mathbf{r}_0''$, we can check[4] that this operator is equal to the product of the simple densities defined in (B-8):

$$\widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') = \widehat{D}(\mathbf{r}_0) \; \widehat{D}(\mathbf{r}_0'') \tag{B-35}$$

Obviously this relation between operators does not mean that the double density $G_2(\mathbf{r}_0, \mathbf{r}_0'')$ is merely the product $\left\langle \widehat{D}(\mathbf{r}_0) \right\rangle \left\langle \widehat{D}(\mathbf{r}_0'') \right\rangle$ of the simple densities: the average value of a product of operators is not, in general, equal to the product of the average values. When we studied the function G_2 (§ C-5-b in Chapter XV), we did find the presence of an exchange term that introduces "statistical correlations" between particles, even in the absence of interactions.

Particles with spin:

For particles with non-zero spin, we just have to add an index ν to each of the kets or bras, as well as to the field operators; this brings up to $(2S+1)^4$ the number of 4-point correlation functions. The matrix elements of the two-body density operator are then given by the average values:

$$\left\langle 1 : \mathbf{r}', \nu'; 2 : \mathbf{r}''', \nu''' \right| \rho_{II} \left| 1 : \mathbf{r}, \nu; 2 : \mathbf{r}'', \nu'' \right\rangle = \left\langle \Psi_\nu^\dagger(\mathbf{r}) \Psi_{\nu''}^\dagger(\mathbf{r}'') \Psi_{\nu'''}(\mathbf{r}''') \Psi_{\nu'}(\mathbf{r}') \right\rangle \tag{B-36}$$

B-4. Hamiltonian operator

We now establish the expression, in terms of the field operator, of the Hamiltonian operator for a system of identical (spinless) particles. Two formulas will be useful for this computation. The first one transposes to three dimensions the formula (34) of Appendix II:

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k \; e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \tag{B-37}$$

[4] If the particles are bosons, we just permute the commuting operators to bring $\Psi(\mathbf{r}_0)$ to the second position and obtain the result. If we are dealing with fermions, two successive anticommutations are necessary to get that result, and the corresponding two minus signs cancel each other.

The second one is obtained from (B-37) by a double derivation with respect to $\mathbf{r}$:

$$\Delta\delta(\mathbf{r}-\mathbf{r}') = -\frac{1}{(2\pi)^3}\int \mathrm{d}^3k\ k^2\ e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \tag{B-38}$$

The matrix elements of a single particle's kinetic energy are written as:

$$\langle\mathbf{r}|\,\frac{P^2}{2m}\,|\mathbf{r}'\rangle = \int \mathrm{d}^3k\ \langle\mathbf{r}|\,\mathbf{k}\rangle\,\frac{\hbar^2k^2}{2m}\,\langle\mathbf{k}\,|\mathbf{r}'\rangle = -\frac{\hbar^2}{2m}\Delta\delta(\mathbf{r}-\mathbf{r}') \tag{B-39}$$

In the Fock space, it corresponds to the following operators (an integration by parts[5] was used to go from the first to the second relation):

$$\widehat{H}_0 = -\frac{\hbar^2}{2m}\int \mathrm{d}^3r\ \Psi^\dagger(\mathbf{r})\Delta\Psi(\mathbf{r}) = \frac{\hbar^2}{2m}\int \mathrm{d}^3r\ \boldsymbol{\nabla}\Psi^\dagger(\mathbf{r})\cdot\boldsymbol{\nabla}\Psi(\mathbf{r}) \tag{B-40}$$

As in § B-1, we obtain here an expression similar to the average value of an operator (here the kinetic energy) for one particle; but the gradient of the wave function must be replaced by that of a field operator, and the order of the operators can matter.

The system Hamiltonian includes in general an interaction term, which makes it a two-particle operator and requires using formula (B-2). For a two-particle system, we know that the interaction yields an operator that is diagonal in the $\{|\mathbf{r},\mathbf{r}'\rangle\}$ representation; furthermore, it only depends on the relative position $\mathbf{r}-\mathbf{r}'$ (and not on $\mathbf{r}$ and $\mathbf{r}'$ separately). Consequently, the matrix element in (B-2) takes the form:

$$\langle\mathbf{r},\mathbf{r}'|\,g\,|\mathbf{r}'',\mathbf{r}'''\rangle = \delta(\mathbf{r}-\mathbf{r}'')\ \delta(\mathbf{r}'-\mathbf{r}''')\times W_{2.}(\mathbf{r}-\mathbf{r}') \tag{B-41}$$

where $W_2(\mathbf{r}-\mathbf{r}')$ is the interaction potential energy between two particles located at a relative position $\mathbf{r}-\mathbf{r}'$ (this interaction is often isotropic, in which case W_2 only depends on the relative distance $|\mathbf{r}-\mathbf{r}'|$). Finally, starting from (B-2), we get the following expression for the Hamiltonian operator:

$$\begin{aligned}\widehat{H} = \int \mathrm{d}^3r\ &\left[\frac{\hbar^2}{2m}\boldsymbol{\nabla}\Psi^\dagger(\mathbf{r})\cdot\boldsymbol{\nabla}\Psi(\mathbf{r})\ + V_1(\mathbf{r})\Psi^\dagger(\mathbf{r})\Psi(\mathbf{r})\right]\\ &+\frac{1}{2}\int \mathrm{d}^3r'\int \mathrm{d}^3r''\ W_2(\mathbf{r}'-\mathbf{r}'')\ \Psi^\dagger(\mathbf{r}')\Psi^\dagger(\mathbf{r}'')\Psi(\mathbf{r}'')\Psi(\mathbf{r}')\end{aligned} \tag{B-42}$$

The first term corresponds to the particles' kinetic energy, the second to the external potential $V_1(\mathbf{r})$ acting separately on each particle, and the third one to the mutual interaction between particles; note that this last term involves four field operators, whereas the first two involve only two. The same comment as above still applies: this expression is reminiscent of the average energy of a system of two particles, both described by the same wave function; but now we are dealing with operators that do not commute.

This Hamiltonian can also be expressed directly via the one-particle simple density $\widehat{D}(\mathbf{r}_0,\mathbf{r}_0')$ and the two-particle double density $\widehat{G}(\mathbf{r}_0,\mathbf{r}_0,\mathbf{r}_0'',\mathbf{r}_0'')$ operators, as we now show. Inserting relations (B-21) and (B-30) in expression (B-42), we obtain:

$$\begin{aligned}\widehat{H} = \int \mathrm{d}^3r\ &\left[\frac{\hbar^2}{2m}\boldsymbol{\nabla}_{\mathbf{r}_0}\cdot\boldsymbol{\nabla}_{\mathbf{r}_0'}\ \widehat{D}(\mathbf{r}_0,\mathbf{r}_0')\Big|_{\mathbf{r}_0=\mathbf{r}_0'=\mathbf{r}} + V_1(\mathbf{r})\widehat{D}(\mathbf{r},\mathbf{r})\right]\\ &+\frac{1}{2}\int \mathrm{d}^3r'\int d^3r''\ W_2(\mathbf{r}'-\mathbf{r}'')\ G_2(\mathbf{r}',\mathbf{r}',\mathbf{r}'',\mathbf{r}'')\end{aligned} \tag{B-43}$$

[5] The value of the already integrated terms must be taken at infinity, and we assume that all the states of the physical system are limited to a finite volume; as a result, those terms do not play any role and can be ignored.

where the notation $\boldsymbol{\nabla}_{\mathbf{r}_0}$ and $\boldsymbol{\nabla}_{\mathbf{r}'_0}$ represents the gradient taken with respect to the variables $\mathbf{r}_0$ (for the first one), and $\mathbf{r}'_0$ (for the second); once these gradients have been computed in the kinetic energy term of (B-43), both variables take on the same value $\mathbf{r}$. The fact that the Hamiltonian operator can be directly expressed in terms of the simple and double density operators can be useful. For example, to determine the ground state energy of an N-particle system, we do not have to compute the state wave function, which involves all the correlations of order 1 to N between the particles; it is sufficient to know the average values of these two densities. There exist, in certain cases, approximation methods that yield directly good estimates of these simple and double densities, hence allowing an access to the N-body energy. Complements $\mathrm{E_{XV}}$ and $\mathrm{G_{XV}}$ discuss the Hartree-Fock method, which is based on an approximation where the two-particle density operator is simply expressed as a function of the one-particle density operator, i.e. the double density as a function of the simple density (Complement $\mathrm{G_{XV}}$, § 2-b-γ); this allows convenient mean field calculations.

C. Time evolution of the field operator (Heisenberg picture)

The operators we have considered until now correspond to the "Schrödinger picture", where the time evolution of the system is determined by the time evolution of its state vector. It may, however, be more convenient to adopt the Heisenberg picture (Complement $\mathrm{G_{III}}$), where this time evolution is transferred to the operators associated with the system's physical quantities. For spinless particles, let us call $\Psi_H(\mathbf{r};t)$ the operator corresponding, in the Heisenberg picture, to $\Psi(\mathbf{r})$:

$$\Psi_H(\mathbf{r};t) = e^{i\widehat{H}t/\hbar}\ \Psi(\mathbf{r})\ e^{-i\widehat{H}t/\hbar} \tag{C-1}$$

($\widehat{H}$ is the Hamiltonian operator), and whose time dependence follows the equation:

$$i\hbar\frac{\partial}{\partial t}\Psi_H(\mathbf{r};t) = \left[\Psi_H(\mathbf{r};t)\ ,\ \widehat{H}\right] \tag{C-2}$$

We are going to compute successively the commutator of $\Psi_H(\mathbf{r};t)$ with each of the three terms on the right-hand side of (B-42). The evolution equation for the field operator involves all the terms of the Hamiltonian (B-42): the kinetic, potential and interaction energies.

C-1. Contribution of the kinetic energy

In order to determine the commutator of the field operator with the kinetic energy, we first transpose the equations (A-17), (A-18) and (A-21) to the Heisenberg picture. Actually, they can be used without any changes: the unitary transform of a product by (C-1) is the product of the unitary transforms, that of the commutator (or of the anticommutator) is the commutator (or the anticommutator) of the transforms, and numbers like zero or the function $\delta(\mathbf{r}-\mathbf{r}')$ are invariant. Those three relations are therefore still valid in the Heisenberg picture, if we simply add an index H to the field operators. We now take their derivative with respect to the positions; only (A-21) yields a non-zero result:

$$\left[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}_{\mathbf{r}'}\Psi_H^\dagger(\mathbf{r}';t)\right]_{-\eta} = \boldsymbol{\nabla}_{\mathbf{r}'}\ \delta(\mathbf{r}-\mathbf{r}') = -\boldsymbol{\nabla}_{\mathbf{r}}\ \delta(\mathbf{r}-\mathbf{r}') \tag{C-3}$$

Taking (B-40) into account, we can write the commutator to be evaluated as:

$$\Big[\Psi_H(\mathbf{r};t), \widehat{H}_{0H}(t)\Big] = \frac{\hbar^2}{2m}\int \mathrm{d}^3r' \ \Big[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)\Big] \tag{C-4}$$

In the term to be integrated on the right-hand side, a sign η is introduced each time we permute two field operators, or two adjoints; when we permute a field operator and an adjoint, we must add to the result the right-hand side of (C-3). Adding and subtracting two equal terms, we then obtain for the function to be integrated:

$$\begin{array}{l}\Psi_H(\mathbf{r};t)\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t) - \eta\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot(\Psi_H(\mathbf{r};t)\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)) \\ +\eta\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot(\Psi_H(\mathbf{r};t)\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)) - \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)\Psi_H(\mathbf{r};t)\end{array} \tag{C-5}$$

that is:

$$\begin{array}{l}\Big[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\Big]_{-\eta}\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t) \\ \qquad +\eta\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)]_{-\eta} \\ \qquad\qquad = \boldsymbol{\nabla}_{\mathbf{r}'}\ \delta(\mathbf{r}-\mathbf{r}')\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}',t)\end{array} \tag{C-6}$$

The integration over d^3r' then yields the Laplacian at $\mathbf{r}$ of the field operator, and we finally get:

$$\Big[\Psi_H(\mathbf{r};t),\ \widehat{H}_{0H}(t)\Big] = -\frac{\hbar^2}{2m}\Delta\Psi_H(\mathbf{r};t) \tag{C-7}$$

C-2. Contribution of the potential energy

Instead of (C-4), it is now the commutator:

$$\Big[\Psi_H(\mathbf{r};t),\ \widehat{V}_{1H}(t)\Big] = \int \mathrm{d}^3r'\ V_1(\mathbf{r}')\ \Big[\Psi_H(\mathbf{r};t),\ \Psi_H^\dagger(\mathbf{r}';t)\Psi_H(\mathbf{r}';t)\Big] \tag{C-8}$$

which comes into play. The calculation is similar to the previous one, but without the gradients which were applied to the field operators depending on $\mathbf{r}'$. The right-hand side of (C-6) now becomes simply:

$$\delta(\mathbf{r}-\mathbf{r}')\Psi_H(\mathbf{r}') \tag{C-9}$$

and the integration over d^3r' is straightforward, so that:

$$\Big[\Psi_H(\mathbf{r};t),\ \widehat{V}_{1H}(t)\Big] = V_1(\mathbf{r})\ \Psi_H(\mathbf{r};t) \tag{C-10}$$

C-3. Contribution of the interaction energy

It is now the commutator of $\Psi_H(\mathbf{r};t)$ with a product of four field operators that will have to be integrated:

$$\Big[\Psi_H(\mathbf{r};t),\ \Psi_H^\dagger(\mathbf{r}';t)\Psi_H^\dagger(\mathbf{r}'';t)\Psi_H(\mathbf{r}'';t)\Psi_H(\mathbf{r}';t)\Big] \tag{C-11}$$

We will not go through the details of the calculation, a bit long but without any real difficulties; as in the previous two cases, it involves the repeated application of the commutation relations. The result is that the commutator of the field with the interaction energy can be written as:

$$\int \mathrm{d}^3 r' \; W_2(\mathbf{r}-\mathbf{r}')\Psi_H^\dagger(\mathbf{r}';t)\Psi_H(\mathbf{r}';t)\Psi_H(\mathbf{r};t) \tag{C-12}$$

C-4. Global evolution

Regrouping the three previous terms, we get the evolution equation for the field operator:

$$\left[i\hbar\frac{\partial}{\partial t}+\frac{\hbar^2}{2m}\Delta - V_1(\mathbf{r})\right]\Psi_H(\mathbf{r};t) = \int \mathrm{d}^3 r' \; W_2(\mathbf{r}-\mathbf{r}') \; \Psi_H^\dagger(\mathbf{r}';t)\Psi_H(\mathbf{r}';t)\Psi_H(\mathbf{r};t) \tag{C-13}$$

The left hand-side includes the differential operator of the usual Schrödinger equation for a single particle in a potential $V_1(\mathbf{r})$; however, as already pointed out, Ψ is not a simple function here, but an operator. The right-hand side includes the binary interaction effects; its presence implies that the evolution equation of the field operator is not "closed". Its evolution depends not only on the operator itself, but also on a term containing the product of three fields (or their conjugates).

Analyzing in a similar way the evolution of such a product of three factors, we see that it depends on that product, and also on the product of 5 fields (or their conjugates); in turn, the evolution of a product of 5 fields will involve 7 others, etc. We thus get a series of more and more complex equations, often called a "hierarchy"of equations. They are in general very hard to solve exactly. This is why it is frequent to use an approximation by truncating this hierarchy at a certain stage, or else by eliminating the coupling term at a certain order, or replacing it by a more convenient expression. Many different methods have been proposed to accomplish this, the most well-known being the mean field approximation (Complements $\mathrm{E_{XV}}$ and $\mathrm{G_{XV}}$).

D. Relation to field quantization

In conclusion, we make some remarks concerning the field quantization procedures, and their relation to the identical particle concept. Consider a single spinless particle in an external potential well, and call $\varphi_i(\mathbf{r})$ the wave functions associated with its stationary states in that potential (the subscript i runs from 1 to infinity; we assume, for the sake of simplicity, that the spectrum is entirely discrete). The $\varphi_i(\mathbf{r})$ form a basis on which we can expand any particle wave function $\psi(\mathbf{r})$:

$$\psi(\mathbf{r}) = \sum_i c_i \varphi_i(\mathbf{r}) \tag{D-1}$$

We already noted the similarity between this formula and equality (A-3), where the only difference is that the numbers c_i are replaced by the operators a_i. Along the same line, relations (B-8) and (B-15) are reminiscent of a particle's probability density of presence, and of its probability current. Finally, expression (B-42) is very similar to the average

value of the energy of a system of two particles, both placed in the same state described by the wave function $\psi(\mathbf{r})$, once we replace that usual wave function by an operator depending on the parameter $\mathbf{r}$. Consequently, the creation and annihilation operator method has an air of "second quantization": we start with the quantum wave function for one (or two) particle(s) ("first quantization"), and in a second stage, we replace the wave functions coefficients by operators ("second quantization"). However, we must keep in mind that we do not, in reality, quantize the same physical system twice; the main difference comes from the fact that we go from a very small number of particles, one or two, to a very large number of identical particles.

Field operators can also appear when quantizing a classical field, such as the electromagnetic field. This is the object of Chapters XIX and XX, where we will show how the concept of a photon emerges, as the elementary excitation of the electromagnetic field. We shall also see how the electric and magnetic fields, which were classical functions, become operators defined at each point in space, creating and annihilating photons.

Generally speaking, a system consisting of an ensemble of identical bosons and a system obtained by quantizing a classical field obey exactly the same equations. The particles of the first system play the role of field quanta for the second system, and the field operators then satisfy commutation relations. The two physical systems are therefore perfectly equivalent. In the case of the electromagnetic field, the particles in questions are the photons and they have a zero mass. However, this is not necessarily the case for all fields. Moreover, quantum fields associated with a system of identical fermions also exist. These fields do not have a direct classical correspondence and their operators obey anticommutation relations. In particle physics, one simultaneously takes into account fermonic and bosonic fields, associated in general with non-zero mass particles.

COMPLEMENTS OF CHAPTER XVI, READER'S GUIDE

A_{XVI} : SPATIAL CORRELATIONS IN A BOSON OR FERMION IDEAL GAS	In this complement we study the properties of the spatial correlation functions in systems of fermions or bosons. For fermions, we establish the existence of an "exchange hole" which corresponds to the impossibility for fermions with parallel spins to be found at the same point in space. For bosons, we discuss their tendency to *bunch* (group together). *Recommended in a first reading*
B_{XVI} : SPATIO-TEMPORAL COORELATION FUNCTIONS, GREEN'S FUNCTIONS	Green's functions are a very general tool for the theoretical study of N-body systems. In this complement, they are first introduced in ordinary space, then in reciprocal space (the Fourier momentum space). Knowledge of these functions allows calculation of numerous physical properties of the system. *Slightly more difficult than the previous complement*
C_{XVI} : WICK'S THEOREM	Wick's theorem permits calculating average values of any product of creation and annihilation operators, for an ideal gas system in thermal equilibrium. The calculation involves a very useful concept, the operator "contraction".

Complement A_{XVI}

Spatial correlations in an ideal gas of bosons or fermions

1 System in a Fock state 1769
1-a Two-point correlations 1770
1-b Four-point correlations 1771
2 Fermions in the ground state 1771
2-a Two-point correlations 1772
2-b Correlations between two particles 1774
3 Bosons in a Fock state 1775
3-a Ground state 1775
3-b Fragmented state 1776
3-c Other states 1777

In this complement, we establish a certain number of properties of the correlation functions, arising solely from the particle statistics (i.e. from the fact they are either bosons or fermions), and independent of their possible interactions. To keep the calculations simple, we assume that the N-particle system is described by a Fock state, characterized by the occupation numbers n_i of each individual state $|u_i\rangle$. We shall see that fermions and bosons behave very differently: whereas the latter tend to bunch, the former tend to avoid each other, as indicated by the existence of an "exchange hole".

We give in § 1 the general expression of these correlation functions, without making any hypothesis concerning the nature of the individual states; the physical system is not necessarily homogeneous in space. In §§ 2 and 3, we study successively bosons and fermions, assuming the physical system to be contained in a box of volume $\mathcal{V}$ in which the particles are free. The periodic boundary conditions (Complement C_{XIV}) allow taking into account the confinement while maintaining translation invariance (the system is perfectly homogeneous in space), which makes the calculations easier. For spinless particles, the individual states $|u_i\rangle$ correspond to plane waves normalized in the volume $\mathcal{V}$:

$$u_i(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}_i \cdot \mathbf{r}} \tag{1}$$

where the $\mathbf{k}_i$ are chosen to satisfy the periodic boundary conditions.

1. System in a Fock state

We assume the state $|\Phi\rangle$ of the N-particle system to be a Fock state built from the basis of individual states $\{|u_i\rangle\}$, with occupation numbers n_i (not greater than 1 for fermions):

$$|\Phi\rangle = |n_1 : u_1;\ n_2 : u_2;\ ..;\ n_i : u_i;\ ..\rangle \tag{2}$$

This will be the case, for example, if the particles do not interact (ideal gas) and if the system is in a stationary state, such as its ground state.

1-a. Two-point correlations

For spinless particles, relations (A-3) and (B-21) of Chapter XVI lead to:

$$\widehat{D}(\mathbf{r},\mathbf{r}') = \sum_{i,j} u_i^*(\mathbf{r}) u_j(\mathbf{r}')\ a_i^\dagger a_j \tag{3}$$

Now the average value in a Fock state of the operator product $a_i^\dagger a_j$ is zero if $i \neq j$: the successive action of the two operators leads to another Fock state with the same particle number, but with two different occupation numbers – therefore to an orthogonal state. If, on the other hand, $i = j$, that product becomes the particle number operator $\widehat{n}_i$ that now acts on its eigenket, with eigenvalue n_i. We thus get:

$$\langle\Phi|\ a_i^\dagger a_j\ |\Phi\rangle = \delta_{ij}\ n_i \tag{4}$$

(where δ_{ij} is the Kronecker delta); this yields:

$$G_1(\mathbf{r},\mathbf{r}') = \left\langle \widehat{D}(\mathbf{r},\mathbf{r}') \right\rangle = \sum_i n_i\ u_i^*(\mathbf{r}) u_i(\mathbf{r}') \tag{5}$$

The physical interpretation of this result is the following: for a single particle in the individual state $|u_i\rangle$, the function G_1 would simply be $u_i^*(\mathbf{r})u_i(\mathbf{r}')$; for an N-particle system in a Fock state, each individual state gives a contribution, multiplied by a coefficient equal to its population.

Particles with non-zero spin:

For particles with spin, we define the $u_i^\nu(\mathbf{r})$ by:

$$u_i^\nu(\mathbf{r}) = \langle \mathbf{r}, \nu\ |u_i\rangle \tag{6}$$

When we add a discrete spin index ν to $\mathbf{r}$, formula (3) becomes:

$$\widehat{D}(\mathbf{r},\nu;\ \mathbf{r}',\nu') = \sum_{i,j} [u_i^\nu(\mathbf{r})]^*\ u_j^{\nu'}(\mathbf{r}')\ a_i^\dagger a_j \tag{7}$$

We obtain, with the same reasoning:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \left\langle \widehat{D}(\mathbf{r},\nu;\ \mathbf{r}',\nu') \right\rangle = \sum_i n_i\ \ [u_i^\nu(\mathbf{r})]^*\ u_i^{\nu'}(\mathbf{r}') \tag{8}$$

whose physical interpretation is similar to the previous one.

One often chooses a basis of individual states $|u_i\rangle$, such that each ket corresponds to a well defined value of the spin: each index i indicates both an individual orbital state and a value of ν (which then becomes a function ν_i of i). In that case, for a given i, the wave function $u_i^\nu(\mathbf{r})$ is only defined for a single value of the index ν; conversely, for a given ν, the wave functions are different from zero only if the index i (or j) belongs to a certain domain $D(\nu)$. In expression (8), ν and ν' are fixed, and the index i must necessarily belong to both $D(\nu)$ and $D(\nu')$, or else the result is zero. This leads to:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \delta_{\nu\nu'} \sum_{i\in D(\nu)} n_i\ \ [u_i^\nu(\mathbf{r})]^*\ u_i^{\nu'}(\mathbf{r}') \tag{9}$$

This correlation function is therefore zero if $\nu' \neq \nu$.

1-b. Four-point correlations

We limit ourselves to the calculation of $\left\langle \widehat{G}(\mathbf{r},\mathbf{r}',\mathbf{r}'',\mathbf{r}''') \right\rangle$ in the "diagonal case" $\mathbf{r}' = \mathbf{r}$ and $\mathbf{r}''' = \mathbf{r}''$; we call $\mathbf{r}_1$ the common value of $\mathbf{r}$ and $\mathbf{r}'$, and $\mathbf{r}_2$ the common value of $\mathbf{r}''$ and $\mathbf{r}'''$. Such a diagonal correlation function was already written in (B-20) of Chapter XVI and in § C-5-a of Chapter XV; this is the only one that plays a role in the particle interaction energy, as the associated operator is also diagonal in the position representation.

In the absence of spin, and for a Fock state, the calculation of the correlation function $G_2(\mathbf{r}_1,\mathbf{r}_2)$ was carried out in § C-5-b of Chapter XV, where relations (C-32) to (C-34) yield the value of this function:

$$\begin{aligned} G_2(\mathbf{r}_1,\mathbf{r}_2) = &\sum_i n_i\,(n_i-1)\ \ |u_i(\mathbf{r}_1)|^2\,|u_i(\mathbf{r}_2)|^2\ + \\ &+\sum_{i\neq j} n_i n_j\left[|u_i(\mathbf{r}_1)|^2\,|u_j(\mathbf{r}_2)|^2 + \eta\ u_i^*(\mathbf{r}_1)u_i(\mathbf{r}_2)u_j^*(\mathbf{r}_2)u_j(\mathbf{r}_1)\right] \end{aligned} \tag{10}$$

where $\eta = +1$ for bosons, $\eta = -1$ for fermions. The second line of this equation contains a "direct term" only involving the moduli squared of the wave functions; it also contains an "exchange term" where the phase of the wave functions come into play, and which changes sign depending on whether we are dealing with a system of bosons or fermions.

Particles with non-zero spin:

In the presence of spins, we must add, as previously, the corresponding spin indices ν, and the correlation function becomes:

$$\begin{aligned} &G_2(\mathbf{r}_1,\nu_1;\mathbf{r}_2,\nu_2) = \textstyle\sum_i n_i\,(n_i-1)\ \ |u_i^{\nu_1}(\mathbf{r}_1)|^2\,|u_i^{\nu_2}(\mathbf{r}_2)|^2\ + \\ &+\textstyle\sum_{i\neq j} n_i n_j\left[\left|u_i^{\nu_1}(\mathbf{r}_1)\right|^2\left|u_j^{\nu_2}(\mathbf{r}_2)\right|^2 + \eta\ \left[u_i^{\nu_1}(\mathbf{r}_1)\right]^* u_i^{\nu_2}(\mathbf{r}_2)\left[u_j^{\nu_2}(\mathbf{r}_2)\right]^* u_j^{\nu_1}(\mathbf{r}_1)\right] \end{aligned} \tag{11}$$

If we choose, as in (9), a basis of individual states $|u_i\rangle$ in which each ket has a well defined value of the spin, the summation is simpler and we get:

$$\begin{aligned} G_2(\mathbf{r}_1,\nu_1;\mathbf{r}_2,\nu_2) = &\,\delta_{\nu_1\nu_2}\textstyle\sum_{i\in D(\nu_1)} n_i\,(n_i-1)\ \ |u_i^{\nu_1}(\mathbf{r}_1)|^2\,|u_i^{\nu_2}(\mathbf{r}_2)|^2\ + \\ &+\textstyle\sum_{i\in D(\nu_1),\ j\in D(\nu_2);\ i\neq j} n_i n_j\left[\left|u_i^{\nu_1}(\mathbf{r}_1)\right|^2\left|u_j^{\nu_2}(\mathbf{r}_2\right|^2 + \right. \\ &\left. +\eta\delta_{\nu_1\nu_2}\left[u_i^{\nu_1}(\mathbf{r}_1)\right]^* u_i^{\nu_2}(\mathbf{r}_2)\left[u_j^{\nu_2}(\mathbf{r}_2)\right]^* u_j^{\nu_1}(\mathbf{r}_1)\right] \end{aligned} \tag{12}$$

As above, $D(\nu)$ corresponds to the domain of the index i for the wave function $u_i^\nu(\mathbf{r})$ to exist (otherwise, it is not defined). If the spin states are different ($\nu_1 \neq \nu_2$), the only contribution to the correlation function comes from the second line (the direct term). The exchange term which follows only concerns particles being in the same spin state; it changes sign depending on whether they are bosons or fermions. No exchange term exists for particles having orthogonal spins. This comes from the fact that to behave as strictly identical objects, two particles must occupy the same spin state; otherwise, their spin direction could, at least in principle, be used to distinguish them.

2. Fermions in the ground state

Consider an ideal gas of fermions, contained in a volume $\mathcal{V}$, and having a spin S equal to 1/2 (as for electrons); the index ν can only take on two values $\pm 1/2$. If we assume

the gas to be in its ground state, this corresponds, for an ideal gas, to a Fock state: for each of the two spin states, all the individual states with an energy lower than a certain value (called the Fermi energy) have an occupation number equal to 1, all the others being empty. We shall proceed as in Complement C_{XIV} (in particular for the study of the magnetic susceptibility) and attribute to each of the two spin states a different Fermi energy – this is useful to account for an average spin orientation (under the influence of a magnetic field for example). For all the values of the index i corresponding to $\nu = +1/2$, we thus assume that the occupation numbers are equal to 1 if they correspond to plane waves (1) having a wave vector smaller than the Fermi vector k_F^+, and zero otherwise. This Fermi vector is linked to the Fermi energy by the relation:

$$E_F^+ = \frac{\hbar^2 \left(k_F^+\right)^2}{2m} \tag{13}$$

where m is the mass of each particle. In a similar way, for all the spin values $\nu = -1/2$, we assume that only the states with wave vectors smaller than the Fermi vector k_F^- are occupied, with a relation similar to (13) in which the index + is replaced by −. The total particle number $N_\pm$ in each of the two spin states are then:

$$N_\pm = \sum_{|\mathbf{k}_i| \leq k_F^\pm} 1 \tag{14}$$

(the summation runs over all the states having a population equal to 1). In the limit of large volumes $\mathcal{V}$, this expression becomes an integral:

$$N_\pm = \frac{\mathcal{V}}{(2\pi)^3} \int_{|\mathbf{k}_i| \leq k_F^\pm} \mathrm{d}^3 k = \frac{\mathcal{V}}{2\pi^2} \int_0^{k_F^\pm} k^2 \mathrm{d}k = \frac{\left(k_F^\pm\right)^3}{6\pi^2} \mathcal{V} \tag{15}$$

Depending on whether k_F^+ is larger or smaller than k_F^-, the spins + or − will make up the majority, the populations being equal if $k_F^+ = k_F^-$.

2-a. Two-point correlations

Let us compute the average value of the operator $\widehat{D}_{(\mathbf{r},\nu;\ \mathbf{r}',\nu')}$ defined in (3), distinguishing the two cases where ν and ν' are equal or different.

(i) Same spin states

Taking (1) into account, relation (9) then yields:

$$G_1(\mathbf{r},\pm;\mathbf{r}',\pm) = \left\langle \widehat{D}(\mathbf{r},\pm;\ \mathbf{r}',\pm) \right\rangle = \frac{1}{\mathcal{V}} \sum_{|\mathbf{k}_i| \leq k_F^\pm} e^{i\mathbf{k}_i \cdot (\mathbf{r}' - \mathbf{r})} \tag{16}$$

where the notation for the spins is simplified from $\pm 1/2$ to $\pm$. In the limit of large volumes, the summation over $\mathbf{k}_i$ becomes an integral, and we get:

$$G_1(\mathbf{r},\pm;\mathbf{r}',\pm) = \frac{1}{(2\pi)^3} \int_{|\mathbf{k}_i| \leq k_F^\pm} \mathrm{d}^3 k\ e^{i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r})} \tag{17}$$

For $\mathbf{r} = \mathbf{r}'$, this function simply yields the particle density $N_\pm/\mathcal{V}$, already computed. For $\mathbf{r} \neq \mathbf{r}'$, the function to be computed is the Fourier transform of a function of $\mathbf{k}$ that only

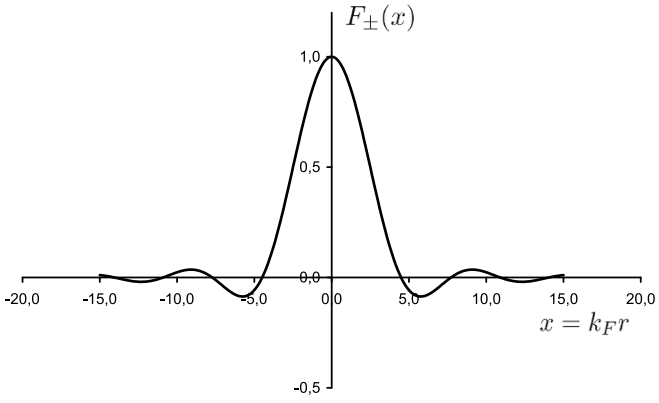


Figure 1: Plot of the function $F_{\pm}(x)$ as a function of the dimensionless variable $x = k_F r$.

depends on its modulus. Using relation (59) of Appendix I in volume II, we can write:

$$G_1(\mathbf{r}, \pm; \mathbf{r}', \pm) = \frac{N_{\pm}}{\mathcal{V}} \frac{F_{\pm}(|\mathbf{r}' - \mathbf{r}|)}{F_{\pm}(\mathbf{0})} \tag{18}$$

with[1]:

$$\begin{aligned} F_{\pm}(r) &= \frac{3}{\left(k_F^{\pm}\right)^3 r} \int_0^{k_F^{\pm}} k \mathrm{d}k \ \sin kr = \frac{3}{\left(k_F^{\pm}\right)^3 r} \left\{ -\frac{k}{r} \cos kr + \frac{1}{r^2} \sin kr \right\}_0^{k_F^{\pm}} \\ &= \frac{3}{\left(k_F^{\pm} r\right)^3} \left[\sin k_F^{\pm} r - \left(k_F^{\pm} r\right) \cos k_F^{\pm} r \right] \end{aligned} \tag{19}$$

Finally, we get:

$$G_1(\mathbf{r}, \pm; \mathbf{r}', \pm) = \frac{N_{\pm}}{\mathcal{V}} \ F_{\pm}(|\mathbf{r}' - \mathbf{r}|) \tag{20}$$

Figure 1 is a plot of $F_{\pm}(r)$ as a function of the mutual distance r between the two points. It shows that, for each spin state, the "non-diagonal" one-particle correlation function presents a maximum at $\mathbf{r} = \mathbf{r}'$, then rapidly decreases to zero over a distance of the order of a few Fermi wavelengths, $\lambda_F^{\pm} = 2\pi / k_F^{\pm}$. A system of free fermions, in its ground state, does not show any long-range "non-diagonal order".

(ii) Opposed spin states.

Relation (9) shows that the two-point correlation function is zero between two states of different spins; there is no non-diagonal order.

[1] An arbitrary coefficient $3/\left(k_F^{\pm}\right)^3$ has been introduced in the function F to make it tend towards 1 when its variable tends towards zero, which allows dropping the factor $F_{\pm}(\mathbf{0})$.

2-b. Correlations between two particles

We start from relation (12). In the second line, the condition $i \neq j$ may be ignored as, for fermions, the $i = j$ terms exactly cancel those on the third line; we can therefore consider the indices i and j as independent. Two cases must be distinguished:

(i) If $\nu_1 = \nu_2$, the three terms in (12) remain, but their behavior as a function of the volume $\mathcal{V}$ are different. This is because, in the limit of large volumes, each of the summations over i or j is proportional to the volume, whereas the moduli squared of the wave functions are each proportional to $1/\mathcal{V}$. For a large system, as the first of the three terms only contains a single sum over i, it varies as $1/\mathcal{V}$ and is thus negligible. We are left with the two other terms:

$$\begin{aligned} G_2\left(\mathbf{r},\pm;\mathbf{r}'',\pm\right) &= \left[\sum_{i\in D(\nu=\pm)} n_i \ \left|u_i^{\pm}(\mathbf{r})\right|^2\right]^2 - \left|\sum_{i\in D(\nu=\pm)} n_i \ \left[u_i^{\pm}(\mathbf{r})\right]^* u_i^{\pm}(\mathbf{r}'')\right|^2 \\ &= \left[\frac{N_\pm}{\mathcal{V}}\right]^2 - \left|\frac{1}{\mathcal{V}}\sum_{|\mathbf{k}_i|\leq k_F^{\pm}} e^{i\mathbf{k}_i\cdot\left(\mathbf{r}''-\mathbf{r}\right)}\right|^2 \end{aligned} \tag{21}$$

The same sum as (17) appears again in this relation. This leads to, in the limit of large volumes:

$$G_2\left(\mathbf{r},\pm;\mathbf{r}'',\pm\right) = \left[\frac{N_\pm}{\mathcal{V}}\right]^2 \left\{1 - \left[F_\pm(|\mathbf{r}''-\mathbf{r}|)\right]^2\right\} \tag{22}$$

The Pauli principle forbids particles in the same spin states to be at the same point in space; as expected, expression (22) goes to zero when $\mathbf{r} = \mathbf{r}''$. As the distance between particles increases, the function $F_\pm(|\mathbf{r}''-\mathbf{r}|)$ goes to zero, and the two-body correlation function tends towards the square of the one-body density $N_\pm/\mathcal{V}$, indicating that the long-range correlations disappear. This change of behavior occurs over a characteristic distance of the order of $\lambda_F^{\pm}$, comparable to the distance over which the non-diagonal order disappears. A plot of the spatial variations of the correlation function is given in Figure 2; it shows clearly the existence of an "exchange hole" corresponding to the mutual particle exclusion over this characteristic distance.

(ii) If $\nu_1 \neq \nu_2$, of the three terms of (12) only the second one (the direct term) is non-zero and yields a constant:

$$G_2\left(\mathbf{r},\pm;\mathbf{r}'',\mp\right) = \frac{N_+ N_-}{\mathcal{V}^2} \tag{23}$$

It is simply the product of the densities of the two kinds of spins; in the absence of interactions, particles with different spins do not show any correlation. This is because physically two particles at positions $\mathbf{r}$ and $\mathbf{r}''$ and in different spin states, can in principle be identified by the direction of their spin; consequently, they no longer behave as really indistinguishable quantum particles, and no Fermi statistical effects may be observed. As we assumed the particles did not interact with each other, no spatial correlations can develop.

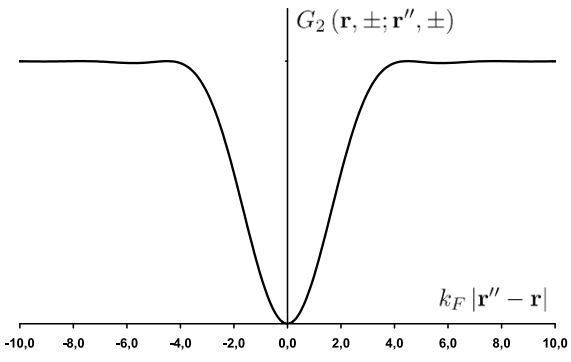


Figure 2: Plot, as a function of the dimensionless variable $k_F\,|\mathbf{r}'' - \mathbf{r}|$, of the correlation function $G_2(\mathbf{r}, \pm; \mathbf{r}'', \pm)$ between the positions $\mathbf{r}$ and $\mathbf{r}''$ of two particles in the same spin state, in a free fermion gas. As the Pauli principle forbids two particles to be at the same point in space, this function goes to zero at the origin, which creates an "exchange hole". As the distance increases, the function approaches 1 over a distance of the order of the inverse of the Fermi vector associated with this spin state.

Comment:

To keep the computation simple, we considered a system of non-interacting fermions, contained in a cubic box and in its ground state. The properties we discussed are, however, more general. In particular, it can be shown that a fermion system always exhibits an exchange hole for particles with identical spins, whether they interact or not; for a system at thermal equilibrium, the hole width gets smaller as the temperature increases, and goes from the Fermi wavelength at low temperature (degenerate system) to the thermal wavelength at high temperature (non-degenerate system).

3. Bosons in a Fock state

The situation is radically different for bosons, as there is no upper boundary for the occupation numbers n_i.

3-a. Ground state

For non-interacting spinless bosons in their ground state, the occupation number n_0 of the individual state $|u_0\rangle$ having the lowest energy is equal to the total particle number N, all the other occupation numbers being equal to zero. Relation (5) then yields:

$$G_1(\mathbf{r};\mathbf{r}') = \left\langle \widehat{D}_{(\mathbf{r},\mathbf{r}')} \right\rangle = N\; u_0^*(\mathbf{r}) u_0(\mathbf{r}') \tag{24}$$

As the wave function $u_0^*(\mathbf{r})$ extends over the entire volume $\mathcal{V}$, the modulus of this wave function does not decrease as the distance between $\mathbf{r}$ and $\mathbf{r}'$ increases and becomes comparable to the size of the system, as opposed to what occurs for fermions. This asymptotic behavior of $G_1(\mathbf{r};\mathbf{r}')$ has been used by Penrose and Onsager to define a general criterion for Bose-Einstein condensation, valid also for interacting bosons .

As for the two-particle correlation function, formula (10) yields:

$$G_2(\mathbf{r},\mathbf{r}'') = \left\langle \widehat{G}_{(\mathbf{r},\mathbf{r},\mathbf{r}'',\mathbf{r}'')} \right\rangle = N\,(N-1)\ \ |u_0(\mathbf{r})|^2\,|u_0(\mathbf{r}'')|^2 \tag{25}$$

If the ground state wave function is of the form $e^{-i\mathbf{k}_0\cdot\mathbf{r}}/\sqrt{\mathcal{V}}$, this function is simply equal to the constant $N\,(N-1)\,/\mathcal{V}^2$, independent of $\mathbf{r}$ and $\mathbf{r}'$. A system of bosons that are all in the same quantum state does not show any spatial correlations.

3-b. Fragmented state

We now assume the N-boson system to be in a "fragmented" state: instead of all the particles being in the same individual state, n_1 particles are in the state $|u_1\rangle$ and n_2 in the state $|u_2\rangle$, with $N = n_1 + n_2$. Relation (5) then yields:

$$G_1(\mathbf{r};\mathbf{r}') = n_1\ u_1^*(\mathbf{r})u_1(\mathbf{r}') + n_2\ u_2^*(\mathbf{r})u_2(\mathbf{r}') \tag{26}$$

When $\mathbf{r} = \mathbf{r}'$, expressions (24) and (26) contain the moduli squared of the wave functions $u_i(\mathbf{r}')$, which are all equal to $1/\mathcal{V}$ (for a system contained in a box of volume $\mathcal{V}$ with periodic boundary conditions); both expressions (24) and (26) are therefore equal. On the other hand, when $\mathbf{r}$ and $\mathbf{r}'$ are different, the phases of the two terms in (26) do not coincide any longer, and (destructive) interference effects can lower the modulus of $G_1(\mathbf{r};\mathbf{r}')$. Consequently, the fragmentation of a physical system into two states decreases the modulus of the non-diagonal terms of $G_1(\mathbf{r};\mathbf{r}')$. Obviously, the more fragmented states there are, the more noticeable the decrease.

Relation (10) now becomes:

$$\begin{aligned} G_2(\mathbf{r},\mathbf{r}'') \;=\; & n_1\,(n_1-1)\ \ |u_1(\mathbf{r})|^2\,|u_1(\mathbf{r}'')|^2 + n_2\,(n_2-1)\ \ |u_2(\mathbf{r})|^2\,|u_2(\mathbf{r}'')|^2 \\ & + n_1 n_2 \left[2\,|u_1(\mathbf{r})|^2\,|u_2(\mathbf{r}'')|^2 + u_1^*(\mathbf{r})u_1(\mathbf{r}'')u_2^*(\mathbf{r}'')u_2(\mathbf{r}) + \text{c.c.} \right] \end{aligned} \tag{27}$$

where the factor 2 in the second line comes from the fact that either $i = 1$ and $j = 2$, or the opposite; the last two terms of this expression correspond to the exchange term, and the notation c.c. indicates the complex conjugate of the previous term. Replacing the wave functions by $e^{-i\mathbf{k}_{1,2}\cdot\mathbf{r}}/\sqrt{\mathcal{V}}$, and assuming n_1 and n_2 to be very large compared to 1, we obtain the square of the sum $(n_1+n_2)^2 = N^2$, and we can write:

$$G_2(\mathbf{r},\mathbf{r}'') \simeq \frac{N^2}{\mathcal{V}^2} + 2\frac{n_1 n_2}{\mathcal{V}^2}\cos\left[(\mathbf{k}_2-\mathbf{k}_1)\cdot(\mathbf{r}-\mathbf{r}'')\right] \tag{28}$$

The first term is simply the square of the one-particle density $N/\mathcal{V}$; it does not have any spatial dependence and is what we expect in the absence of any particle correlation. On the other hand, the exchange term is position dependent; it presents a maximum when $\mathbf{r} = \mathbf{r}''$, and oscillates at the spatial frequency $|\mathbf{k}_2 - \mathbf{k}_1|$. This exchange term enhances the probability of finding two bosons close to one another (bunching effect

coming from the Bose-Einstein statistics); the probability of finding them at a greater distance is then lower, and then increases again, etc. For a short interaction range, only the first maximum plays a role, and increases the average value of the interaction. The consequences of that effect, in terms of the internal interaction energy, has been discussed in § 4-c of Complement C_{XV}.

3-c. Other states

We now consider situations described by Fock states where n_0 is still very large, but where other states $|u_i\rangle$ are also occupied, with populations n_i much smaller than n_0.

(i) One could for instance place a finite fraction n_0/N of particles in the ground state, and distribute the remaining fraction $1 - n_0/N$ among a large number of states, whose individual populations remain small and vary regularly with the index i. This leads to:

$$G_1(\mathbf{r};\mathbf{r}') = n_0 \; u_0^*(\mathbf{r})u_0(\mathbf{r}') + Q(\mathbf{r}-\mathbf{r}') \tag{29}$$

where $Q(\mathbf{r}-\mathbf{r}')$ is given by:

$$Q(\mathbf{r}-\mathbf{r}') = \frac{1}{\mathcal{V}}\sum_{i\neq 0} n_i \; e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})} = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k_i \; n(\mathbf{k}_i) \; e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})} \tag{30}$$

This function is the Fourier transform of the distribution $n_i = n(\mathbf{k}_i)$ for $\mathbf{k}_i \neq 0$. As all the n_i are positive or zero, the function $Q(\mathbf{r}-\mathbf{r}')$ presents a maximum when $\mathbf{r} = \mathbf{r}'$, since this is where all the exponentials $e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})}$ are in phase; it then decreases when the difference $|\mathbf{r}-\mathbf{r}'|$ increases, as all the phases spread out. If the n_i distribution is a regular function of width Δk (a Gaussian for example), the function $Q(\mathbf{r}-\mathbf{r}')$ tends towards zero over a distance $\Delta l \simeq 1/\Delta k$, in general much smaller than the size of the system.

Figure 3 is a plot of the function $G_1(\mathbf{r};\mathbf{r}')$ when the particles are contained in a box, so that we can use (1). As we assumed $|u_0\rangle$ to be the ground state, the corresponding wave function is the inverse of the square root of the volume, and (29) becomes:

$$G_1(\mathbf{r};\mathbf{r}') = x_0 \; + Q(\mathbf{r}-\mathbf{r}') \tag{31}$$

where x_0 is the density of atoms in the ground state:

$$x_0 = \frac{n_0}{\mathcal{V}} \tag{32}$$

After the decrease linked to that of $Q(\mathbf{r}-\mathbf{r}')$, and occurring over the interval $\Delta l \simeq 1/\Delta k$, the function does not tend towards zero (as it would for fermions), but towards a constant proportional to the population n_0. As already mentioned, this particular behavior is the base for the Penrose and Onsager criterion that defines, in a general way, the appearance of Bose-Einstein condensation.

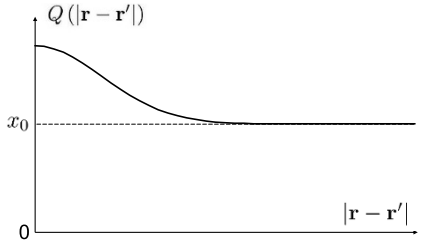


Figure 3: Plot of the function $G_1(\mathbf{r};\mathbf{r}') = x_0 + Q(\mathbf{r}-\mathbf{r}')$ for bosons, as a function of the distance $|\mathbf{r}-\mathbf{r}'|$. The function starts by decreasing over an interval of the order of the inverse of Δk; at a larger distance, it tends towards a constant x_0 proportional to the ground state population. The fact that it does not go to zero indicates the presence of a long-range non-diagonal order, and the existence of a highly populated individual level.

As for the two-body correlation function, relation (10) shows the existence of three kinds of terms for a system having only one single highly populated state $|u_0\rangle$:

– the terms corresponding to two particles in the highly populated state (condensate), which come from the first line[2] of (10), and yield again (25), replacing N by n_0.

– the crossed terms in $n_0 n_i$, which yield:

$$n_0 \sum_{i\neq 0} n_i \left[2\left|u_i(\mathbf{r})\right|^2 \left|u_0(\mathbf{r}'')\right|^2 + \; u_0^*(\mathbf{r})u_0(\mathbf{r}'')u_i^*(\mathbf{r}'')u_i(\mathbf{r}) + u_i^*(\mathbf{r})u_i(\mathbf{r}'')u_0^*(\mathbf{r}'')u_0(\mathbf{r})\right] \tag{33}$$

Inserting the value (1) for the wave functions (assuming $\mathbf{k}_0 = 0$), we get a contribution to $G_2(\mathbf{r},\mathbf{r}')$ equal to:

$$\frac{n_0}{\mathcal{V}^2} \sum_{i\neq 0} n_i \left[2 + e^{i\mathbf{k}_i\cdot\left(\mathbf{r}-\mathbf{r}''\right)} + e^{i\mathbf{k}_i\cdot\left(\mathbf{r}''-\mathbf{r}\right)}\right] \tag{34}$$

Using relation $\sum_{i\neq 0} n_i = N - n_0$, we can write this result in the form:

$$\frac{2n_0}{\mathcal{V}}\left[\frac{N-n_0}{\mathcal{V}} + \mathrm{Re}\left[Q(\mathbf{r}-\mathbf{r}')\right]\right] \tag{35}$$

[2]In the limit of large volumes, we have assumed that n_0 is the only population proportional to the volume. In the first line of (10), the term $i = 0$ then contains the product of n_0^2 and the two wave functions squared, each proportional to the inverse of the volume; this term is therefore independent of the volume. On the other hand, the terms $i \neq 0$ of the second line contains one summation over i, introducing a factor proportional to the volume (in the limit of large volumes), but also two wave functions squared, each inversely proportional to the volume. The net result is a contribution inversely proportional to the volume, hence negligible compared to the previous one.

where Re means "real part" and Q is the function defined in (30).

– the terms corresponding to two particles in states other than the ground state, and which yield:

$$\sum_{i \neq j} \frac{n_i n_j}{\mathcal{V}^2} \left[1 + e^{i(\mathbf{k}_i - \mathbf{k}_j)\cdot(\mathbf{r}'' - \mathbf{r})} \right] \tag{36}$$

In (34), as well as in (36), we notice that the contributions from all the states $i \neq 0$ have various phases in general; they are, however, all in phase when $\mathbf{r} = \mathbf{r}'$ and the correlation function then presents a maximum. The bosons have thus a tendency for bunching, and this effect is felt over a distance $\Delta l \simeq 1/\Delta k$, as if they were attracted to one another. It is, however, a purely statistical effect linked to the bosonic character of the particles, since we assumed there were no interactions between the particles.

(ii) One could also imagine the population distribution to be regular, without favoring any individual state, in which case the n_0 contribution vanishes; we are then left with the contribution from the $n_i n_j$ terms of (36), which is maximum when $\mathbf{r} = \mathbf{r}'$ for the same reasons as above. The general behavior of the two-body correlation function is shown in Figure 3: it presents a maximum at the origin, and then tends towards zero at large distance (in this case, $x_0 = 0$). Once again, identical bosons exhibit a bunching tendency.

Comment:

Suppose that, instead of assuming the bosonic system to be in a Fock state (a pure state), it is at thermal equilibrium, described by the thermodynamic equilibrium. This would lead to results similar to those we just derived, but with $\Delta l \simeq \lambda_T$, the thermal wavelength of the particles [24]. The boson bunching tendency is a quite general property.

Complement B_{XVI}

Spatio-temporal correlation functions, Green's functions

1 Green's functions in ordinary space 1781
1-a Spatio-temporal correlation functions 1782
1-b Two- and four-point Green's functions 1786
1-c An example, the ideal gas 1787
2 Fourier transforms 1790
2-a General definition 1790
2-b Ideal gas example 1791
2-c General expression in the presence of interactions 1792
2-d Discussion 1794
3 Spectral function, sum rule 1795
3-a Expression of the one-particle correlation functions 1795
3-b Sum rule 1796
3-c Expression of various physical quantities 1797

This complement discusses the properties of the spatio-temporal correlation functions of an ensemble of identical particles, generalizing the spatial correlation functions defined in § C-5-b of Chapter XV and § B-3 of Chapter XVI; the corresponding Green's functions shall also be introduced. We first study (§ 1) the normal and anti-normal spatio-temporal correlation functions, then the Green's function, and discuss some of their properties, illustrated with the example of an ideal gas. We then study in § 2 the Fourier transforms of these functions for physical systems that are translation invariant both in space and time; we shall write their general expression in the presence of interactions. In § 3, we finally introduce the "spectral function", which leads, for interacting particles, to very simple expressions for various physical quantities, in a form similar to the one used for an ideal gas.

1. Green's functions in ordinary space

In the previous complement, we studied the spatial dependence of the correlation functions, taken at a given time t. We now take into account the temporal dependence, using the Heisenberg picture (Complement G_{III}) where the operators are time-dependent. To keep the notation simple, we assume, from now on, that either the spin is zero for bosons ($S = 0$), or else, in the general case (fermions and bosons), that all the particles are in the same spin state. As mentioned before, the generalization to the case where S is non-zero would only require adding an index ν to all the field operators. In the Heisenberg picture, the field operator $\Psi(\mathbf{r})$ becomes a time-dependent operator $\Psi_H(\mathbf{r}, \mathbf{t})$:

$$\Psi_H(\mathbf{r}, t) = e^{iHt/\hbar}\ \Psi(\mathbf{r})\ e^{-iHt/\hbar} \tag{1}$$

where $e^{-iHt/\hbar}$ is the evolution operator, expressed as a function of the system Hamiltonian H (including the particle interactions when present), that we assume to be time-independent.

Consider a system of N identical particles, fermions or bosons, described by a density operator ρ. The spatio-temporal correlation functions and the Green's functions are defined as the average values, computed with ρ, of the products of a number of field operators $\Psi_H(\mathbf{r},t)$ and their Hermitian conjugates $\Psi_H^\dagger(\mathbf{r},t)$ taken at different space-time points $(\mathbf{r},t)$, $(\mathbf{r}',t')$,... etc.

1-a. Spatio-temporal correlation functions

The density operator ρ may contain very complex correlations between particles, and its time evolution can be very complicated in the presence of interactions. We are going to define a certain number of functions that characterize its most simple and useful properties, as they only pertain to a small number of particles.

α. *Two-point normal and anti-normal functions*

The one-particle spatio-temporal "normal" G_1^N and "anti-normal" G_1^{AN} correlation functions are defined by[1]:

$$\boxed{\begin{aligned} G_1^N(\mathbf{r},t;\mathbf{r}',t') &= \mathrm{Tr}\left\{\rho\ \Psi_H^\dagger(\mathbf{r},t)\Psi_H(\mathbf{r}',t')\right\} = \left\langle \Psi_H^\dagger(\mathbf{r},t)\Psi_H(\mathbf{r}',t')\right\rangle \\ G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') &= \mathrm{Tr}\left\{\rho\ \Psi_H(\mathbf{r}',t')\Psi_H^\dagger(\mathbf{r},t)\right\} = \left\langle \Psi_H(\mathbf{r}',t')\Psi_H^\dagger(\mathbf{r},t)\right\rangle \end{aligned}} \tag{2}$$

The normal ordering is obtained when the creation operator is on the left and the annihilation on the right; it is the opposite for the anti-normal order. Note that it is only the order of the two operators that changes between G_1^N and G_1^{AN}; the position and time variables attributed to each of the two operators Ψ_H and $\Psi_H^\dagger$ remain the same.

In the particular case where $t' = t$, the normal correlation function simply yields the matrix elements of the one-particle density operator – see formulas (B-22) and (B-23) of Chapter XVI. The normal correlation function is therefore a generalization, at different times, of this matrix element, which will prove to be useful.

To understand the physical meaning of these two definitions in an intuitive way, we start with the anti-normal function and consider the simple case where an N-particle system is in a pure state $|\Phi_0\rangle$ (its ground state, for example). We then get:

$$\begin{aligned} G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') &= \langle\Phi_0|\,\Psi_H(\mathbf{r}',t')\Psi_H^\dagger(\mathbf{r},t)\,|\Phi_0\rangle \\ &= \langle\Phi_0|\,e^{iHt'/\hbar}\Psi(\mathbf{r}')e^{iH\left(t-t'\right)/\hbar}\Psi^\dagger(\mathbf{r})e^{-iHt/\hbar}\,|\Phi_0\rangle \end{aligned} \tag{3}$$

The right-hand side reads as follows (from right to left): starting from the system initial state $|\Phi_0\rangle$, we let it evolve according to its own Hamiltonian until the time t, when we create a particle at point $\mathbf{r}$; we then let the $(N+1)$-particle system freely evolve until time t', and we finally annihilate a particle at point $\mathbf{r}'$. Consequently, the function G_1^{AN} is the scalar product of the ket thus obtained with the state $e^{-iHt'/\hbar}\,|\Phi_0\rangle$, result of the

[1] To be consistent with the notation of Chapter XVI (§ B), we choose a notation where, for the trace, the first group of variables $(\mathbf{r},t)$ of function G_1 is associated with the operator $\Psi^\dagger$, the second group $(\mathbf{r}',t')$ to the operator Ψ. Note however that the opposite convention can also be found in the literature.

free evolution of the initial ket over the same time interval (without the creation or annihilation of a particle). In other words, the perturbation created by the creation of a particle, followed by its later destruction, changes the state of the physical system; the value of G_1^{AN} is given by the probability amplitude of finding the system in the same state as the one it would have reached in the absence of this perturbation.

The previous interpretation is natural when $t' \geq t$; if this is not the case, the mathematical definition of G_1^{AN} is the same, but the intermediate evolution stage goes backward in time. In this process, and as expected, the dynamics of the system remains unchanged, including the particle interactions. Furthermore, the additional particle is not simply a particle juxtaposed to the pre-existing system, it is indistinguishable from the others and hence undergoes indistinguishability effects (more details on this point will be given in § 1-a-β).

For the normal function, $\Psi_H(\mathbf{r}', t')$ acts before $\Psi_H^\dagger(\mathbf{r}, t)$, which means this function has a natural interpretation if $t \geq t'$: the system evolves freely until t', at which time a particle is annihilated (which amounts to the creation of a "hole", see Complement $\mathrm{A_{XV}}$ $\mathrm{A_{XV}}$); the system, with $N-1$ particles, then evolves freely until time t, when a particle is created (which annihilates the hole). The normal function is therefore the analog of the anti-normal function, provided we replace the additional particle by a hole, and we invert the times.

β. *Physical discussion*

The definition of the correlation functions contains particle creation and annihilation operators, but this does not imply that such physical processes really occur in our system. Going from an N-particle system to another one with $N \pm 1$ particles can be mathematically useful but only plays an intermediate role since we finally return, via a second operator, to the same number of particles. Furthermore, the action of an operator $\Psi(\mathbf{r})$ is not merely a local destruction of a particle, neither is the action of $\Psi^\dagger(\mathbf{r})$ the simple creation of a particle at point $\mathbf{r}$, juxtaposed with the already present particles: the quantum indistinguishability that concerns all the particles (including the new one) plays an essential role.

A few simple examples

(i) For the normal correlation function, the perturbation starts with the annihilation of a particle (creation of a hole), followed by a later creation of a particle (destruction of a hole). At time $t = 0$, the annihilation operator $\Psi_H(\mathbf{r}', t)$ can be expanded according to formula (A-3) of Chapter XVI:

$$\Psi(\mathbf{r}') = \sum_i u_i(\mathbf{r}')\ a_i \tag{4}$$

where the effect of each operator a_i depends on the population n_i of the state $|u_i\rangle$ as it introduces the factor $\sqrt{n_i}$. Consider a very simple case: a gas of bosons, all in the same quantum state $|u_0\rangle$ (ideal gas of totally condensed bosons). Acting on a state $|\Phi_0\rangle$ where only the individual state $|u_0\rangle$ is occupied, all the terms of the sum (4) yield zero, except for the $i = 0$ term. The effect of the operator $\Psi(\mathbf{r}')$ on $|\Phi_0\rangle$ is to actually destroy a particle in the state $|u_0\rangle$; as $\mathbf{r}'$ varies, the result is still the same, simply multiplied by the coefficient $u_0(\mathbf{r}')$. If, for example, the ideal gas is placed in a trap where the individual ground state is $|u_0\rangle$, the operator $\Psi(\mathbf{r}')$ yields a ket with an appreciable norm only if

$\mathbf{r}'$ falls in a domain where $|u_0(\mathbf{r}')|$ is not negligible; if it falls outside this domain, the resulting ket is practically zero. In a general way, for fermions as for bosons, $\Psi(\mathbf{r}')$ can obviously destroy particles only while acting on an already occupied individual state. For bosons, in addition, the factor $\sqrt{n_i}$ means that the operator $\Psi(\mathbf{r}')$ gives more weight to the highly populated individual states rather than those with low occupation numbers. Consequently, the creation of a hole is not a local process at point $\mathbf{r}'$.

(ii) For the anti-normal correlation function, we start with the creation of a particle by the operator $\Psi^\dagger(\mathbf{r})$. In the case of bosons, and because of the factor $\sqrt{n_i+1}$ introduced by $a_i^\dagger$, $\Psi^\dagger(\mathbf{r})$ tends to preferentially create particles in states $|u_i\rangle$ having a high population n_i. Let us go back to the previous example of a large number of bosons in a trap, all in the same individual ground state $|u_0\rangle$. When $|u_0(\mathbf{r})|$ is not negligible, the supplementary boson is created in the same ground state. If, on the other hand, $\mathbf{r}$ is far away from the trap center and falls in a domain where $|u_0(\mathbf{r})|$ is practically zero, the boson is actually created at point $\mathbf{r}$ but without perturbing very much the bosons already present.

For fermions, on the contrary, it is impossible to create a particle in an already occupied state; in a Fock state, an additional fermion can only be created in a state orthogonal to all the initially occupied states. Let us assume the ideal gas of fermions is in its ground state, and contained in a harmonic trap; the energy levels, up to the Fermi level, are all occupied. The effect of the creation operator $\Psi^\dagger(\mathbf{r})$ at a point close to the trap center is to create an additional particle in a state that can be expanded on all the individual stationary states of the trap; as this state must be orthogonal to all the already occupied states, it only has components on the non-occupied states, which have an energy higher than the Fermi level. Now the corresponding wave functions take on small values at the center of the trap, and are maximum in the classical turning point regions[2], which means, in this case, at the edge of the existing fermion cloud (or even further). If position $\mathbf{r}$ is close to the center of the trap, the additional fermion will be added on the periphery or outside the cloud of fermions. On the other hand, if $\mathbf{r}$ falls outside, in a region of space where all the wave functions of the occupied states are practically zero, the additional particle will be created practically right at point $\mathbf{r}$. Examples of these various situations will be given in § 1-c.

γ. *Properties*

The complex conjugate of G_1^N is obtained by changing the field operators' order in (2) and replacing them by their Hermitian conjugates:

$$\left[G_1^N(\mathbf{r},t;\ \mathbf{r}',t')\right]^* = \left\langle \Psi_H^\dagger(\mathbf{r}',t')\Psi_H(\mathbf{r},t)\right\rangle = G_1^N(\mathbf{r}',t';\ \mathbf{r},t) \tag{5}$$

The complex conjugation of the function G_1^N is therefore equivalent to exchanging the two points $(\mathbf{r},t)$ and $(\mathbf{r}',t')$, which amounts to a parity operation on the variables $\mathbf{r}-\mathbf{r}'$ and $t-t'$. A similar property is easily demonstrated for G_1^{AN}. As a result, the Fourier transforms of these functions with respect to the variables $\mathbf{r}-\mathbf{r}'$ and $t-t'$ must be real.

[2]The modulus squared of a stationary state wave function yields, at each point, the probability density of presence for this state. This probability is maximal in the regions of space where, classically, the particle spends the most time, i.e. where its velocity is small, as is the case for the classical turning point regions. Figure 6 of Chapter V gives an example of such a situation.

In addition, when the system is in a state that is translation invariant in space and time[3], the correlation functions only depend on the differences $\mathbf{r}-\mathbf{r}'$ and $t-t'$.

We note that the linear combination $G_1^{AN}(\mathbf{r},t;\mathbf{r}',t')-\eta G_1^N(\mathbf{r},t;\mathbf{r}',t')$ involves the average value of the operator:

$$\left[\Psi_H(\mathbf{r}',t')\ ,\ \Psi_H^\dagger(\mathbf{r},t)\right]_{-\eta} \tag{6}$$

where $\eta=+1$ for bosons, $\eta=-1$ for fermions. If $t=t'$, and taking into account relation (A-19) of Chapter XVI, this equality becomes:

$$G_1^{AN}(\mathbf{r},t;\mathbf{r}',t)-\eta G_1^N(\mathbf{r},t;\mathbf{r}',t)=\mathrm{Tr}\left\{\rho\ \delta\left(\mathbf{r}-\mathbf{r}'\right)\right\}=\delta\left(\mathbf{r}-\mathbf{r}'\right) \tag{7}$$

δ. Temporal evolution, four-point functions

We now use, in definition (2) of G_1^N, the Hermitian conjugate of the evolution equation for $\Psi_H(\mathbf{r},t)$, written in (C-13) of Chapter XVI; this leads to:

$$\begin{aligned}\left[-i\hbar\frac{\partial}{\partial t}+\frac{\hbar^2}{2m}\Delta_{\mathbf{r}}-V_1(\mathbf{r})\right]&G_1^N(\mathbf{r},t;\mathbf{r}',t')\\ &=\int \mathrm{d}^3r''\ W_2(\mathbf{r}-\mathbf{r}'')\ \left\langle\Psi_H^\dagger(\mathbf{r};t)\Psi_H^\dagger(\mathbf{r}'';t)\Psi_H(\mathbf{r}'';t)\Psi_H(\mathbf{r}';t')\right\rangle\\ &=\int \mathrm{d}^3r''\ W_2(\mathbf{r}-\mathbf{r}'')\ G_2^N(\mathbf{r},t;\mathbf{r}',t;\mathbf{r}'',t;\mathbf{r}'',t)\end{aligned} \tag{8}$$

which involves the normal two-particle (or four-point) correlation function, whose general expression is:

$$G_2^N(\mathbf{r},t;\mathbf{r}',t';\mathbf{r}'',t'';\mathbf{r}''',t''')=\mathrm{Tr}\left\{\rho\ \Psi_H^\dagger(\mathbf{r},t)\Psi_H^\dagger(\mathbf{r}'',t'')\Psi_H(\mathbf{r}''',t''')\Psi_H(\mathbf{r}',t')\right\} \tag{9}$$

Taking into account a time dependence generalizes formula (B-30) of Chapter XVI (or more precisely its average value). We obtain, in a similar way, the equation giving the variation of G_1^N with respect to the variables $\mathbf{r}'$ and t', or that giving the evolution of the anti-normal function G_1^{AN}. The four groups of space-time variables the function G_2^N depends on, are, in general, independent. Most of the time, however, we only need the "diagonal part" $G_2^N(\mathbf{r},t;\mathbf{r}'',t'')$ of the correlation function, obtained for $\mathbf{r}'=\mathbf{r}$, $t'=t$ and $\mathbf{r}'''=\mathbf{r}''$, $t'''=t''$; this diagonal part is analogous to the two-particle correlation function (two positions, two times) in classical statistical mechanics. When the system is translation-invariant both in space and time, this function only depends on the differences $\mathbf{r}-\mathbf{r}''$ and $t-t''$.

Whereas, for the function G_1^N, a hole is created and then destroyed, for the function G_2^N there are now two holes first created and then destroyed; the natural order of the increasing times is given by the time variables of the operators in (9), taken from right to left. One could define, in a similar way, a function G_2^{AN} where two particles would be created at the beginning and then destroyed at later times (compared to the normal correlation function, the role of particles and holes are interchanged). Consequently, when the particles interact, the evolution equation of G_1^N involves another higher order

[3]This is the case if the system Hamiltonian is translation invariant, and if the system is in an eigenstate of H, or described by a density operator ρ that is a function of H.

correlation function, G_2^N. In turn, the evolution of G_2^N involves correlation functions of an even higher order, G_3^N etc. This means that, because of the interactions, the set of equations is not "closed", but includes a complete "hierarchy" of a large number of equations, involving correlations of higher and higher order.

1-b. Two- and four-point Green's functions

Equation (8) is a linear partial differential equation, with a right-hand side sometimes called a "source term". In our case, this right-hand side does not contain any singular function. But when this right-hand side is modified to include a delta function, the new solutions of the equation are called the "Green's functions" $\mathcal{G}$. We now show how to introduce Green's functions in the problem we are concerned with.

α. *Two-point Green's function*

The two-point Green's function $\mathcal{G}_1$ is obtained by a combination of the two correlation functions of § 1-a. We saw in § 1-a that, when $t' > t$, the anti-normal correlation function is the most natural as it includes the propagation of a particle from t to t'. On the other hand, for $t' < t$, the most natural one is the normal function that involves the propagation of a hole from t' to t. We can combine those two possibilities into one by setting:

$$\boxed{\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = Y\,(t'-t)\,G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') + \eta Y\,(t-t')\,G_1^{N}(\mathbf{r},t;\mathbf{r}',t')} \tag{10}$$

where $Y\,(t-t')$ is the Heaviside function of the variable $t'-t$ (equal to 1 if $t'-t>0$, zero otherwise), and where η equals $+1$ for bosons, -1 for fermions; we shall see later on, for example in § 1-c-γ, that the introduction of the η factor simplifies the following computations.

When we take the derivative of the two-point Green's function $\mathcal{G}_1$ with respect to time, the discontinuities introduced by the Heaviside functions yield delta functions; the precise calculation will be done in § 1-c-γ for an ideal gas, allowing us to verify that $\mathcal{G}_1$ is indeed a Green's function. Using this type of function is quite useful in a number of problems, for example those involving Fourier transforms, or in perturbation calculations.

β. *Four-point Green's function*

By analogy with (9) we define a two-particle (or four-point) Green's function:

$$\boxed{\mathcal{G}_2(\mathbf{r},t;\mathbf{r}',t';\mathbf{r}'',t'';\mathbf{r}''',t''') = \mathrm{Tr}\left\{\rho\ T\ \Psi_H(\mathbf{r}',t')\Psi_H(\mathbf{r}''',t''')\Psi_H^\dagger(\mathbf{r}'',t'')\Psi_H^\dagger(\mathbf{r},t)\right\}} \tag{11}$$

where T is the time ordering operator, which orders the 4 times by decreasing values from left to right (by definition, this operator also includes a factor η^P, where P is the parity of the permutation necessary for this time ordering; this may result in a change of sign for fermions).

1-c. An example, the ideal gas

When the system considered is an ideal gas, it is possible to get explicit values of the previous functions. As before, we assume the gas to be contained in a box of volume $\mathcal{V}$, with periodic boundary conditions. Using relation (A-3) of Chapter XVI in the case where the $u_i(\mathbf{r})$ are plane waves, we can write the field operator $\Psi(\mathbf{r})$ as:

$$\Psi(\mathbf{r}) = \sum_i \frac{e^{i\mathbf{k}_i\cdot\mathbf{r}}}{\sqrt{\mathcal{V}}}\ a_{\mathbf{k}_i} \tag{12}$$

where the sum covers all the wave vectors $\mathbf{k}_i$ satisfying the periodic boundary conditions. This expansion is convenient as, for an ideal gas, the time dependence of the operators in the Heisenberg picture $a^H_{\mathbf{k}}$ is particularly simple, as we now show. The Hamiltonian can be written as:

$$H = \sum_j\ \hbar\omega_{\mathbf{k}_j}\left(a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_j}\right) \tag{13}$$

with:

$$\omega_{\mathbf{k}_j} = \frac{\hbar\,(\mathbf{k}_j)^2}{2m} \tag{14}$$

As the operator $\left(a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_j}\right)$ commutes with any annihilation operator $a_{\mathbf{k}_i}$ pertaining to a different momentum[4], and since $\left[a_{\mathbf{k}_i}, a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}\right] = a_{\mathbf{k}_i}$, we get:

$$[a_{\mathbf{k}_i}, H] = \hbar\omega_{\mathbf{k}_i}\ a_{\mathbf{k}_i} \tag{15}$$

This corresponds, in the Heisenberg picture, to an evolution:

$$a^H_{\mathbf{k}_i}(t) = e^{-i\omega_{\mathbf{k}_i}t}\ a_{\mathbf{k}_i} \tag{16}$$

The time evolution of $\Psi_H(\mathbf{r},t)$ is then written as:

$$\Psi_H(\mathbf{r},t) = \sum_i \frac{e^{i\left(\mathbf{k}_i\cdot\mathbf{r}-\omega_{\mathbf{k}_i}t\right)}}{\sqrt{\mathcal{V}}}a_{\mathbf{k}_i} \tag{17}$$

α. Normal correlation function

Inserting this result in (2), we get:

$$G^N_1(\mathbf{r},t;\mathbf{r}',t') = \sum_{i,j}\frac{e^{-i\left(\mathbf{k}_j\cdot\mathbf{r}-\omega_{\mathbf{k}_j}t\right)}e^{i\left(\mathbf{k}_i\cdot\mathbf{r}'-\omega_{\mathbf{k}_i}t'\right)}}{\mathcal{V}}\left\langle a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_i}\right\rangle \tag{18}$$

We now show that, because of the translation invariance, the average value $\left\langle a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_i}\right\rangle$ must be zero whenever $\mathbf{k}_j \neq \mathbf{k}_i$. Assume, for example, that the system density operator

[4]For fermions, two minus signs are introduced because of the anticommutations, but they cancel each other. If the momentum is the same, we have $a_{\mathbf{k}_i}a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i} = -a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}a_{\mathbf{k}_i} + a_{\mathbf{k}_i} = 0 + a_{\mathbf{k}_i}$ and $a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}a_{\mathbf{k}_i} = 0$, so that $\left[a_{\mathbf{k}_i}, a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}\right] = a_{\mathbf{k}_i}$.

is the canonical thermal equilibrium operator $\rho_{eq} = e^{-\beta H}/Z$. This operator is diagonal in the basis of the Fock states, and the trace of the product $\rho a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i}$ will be zero unless both annihilation and creation operators act on the same individual state: the non-zero condition is therefore $\mathbf{k}_j = \mathbf{k}_i$. In the same way, if we are now using the grand canonical equilibrium $\rho_{eq} = e^{-\beta(H-\mu N)}/Z$, the operator ρ is still diagonal in the same basis, and the same rule applies. In a general way, one can see[5] that the translation invariance of ρ requires:

$$\left\langle a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i} \right\rangle = \mathrm{Tr}\left\{ \rho a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i} \right\} = \delta_{ij} \times n_i \tag{19}$$

where n_i is the average value of the population operator $\widehat{n}_i$:

$$n_i = \langle \widehat{n}_i \rangle \tag{20}$$

We can then write:

$$G_1^N(\mathbf{r}, t; \mathbf{r}', t') = \frac{1}{\mathcal{V}} \sum_i n_i \; e^{i\left[\mathbf{k}_i \cdot (\mathbf{r}' - \mathbf{r}) - \omega_{\mathbf{k}_i}(t' - t)\right]} \tag{21}$$

The normal correlation function is simply the sum of all the contributions from the individual states occupied by free particles, labeled by the index $\mathbf{k}_i$. Each contributes proportionally to its average population n_i, and has a spatio-temporal dependence given by the progressive wave $e^{i\left[\mathbf{k}_i \cdot (\mathbf{r}' - \mathbf{r}) - \omega_{\mathbf{k}_i}(t' - t)\right]}$ it is associated with.

As expected for a translation invariant (both in space and time) system, this normal function only depends on the differences $\mathbf{r} - \mathbf{r}'$ and $t - t'$. Taking (14) into account, it obeys the partial differential equation:

$$\left[i \frac{\partial}{\partial t'} + \frac{\hbar}{2m} \Delta_{\mathbf{r}'} \right] G_1^N(\mathbf{r}, t; \mathbf{r}', t') = 0 \tag{22}$$

which corresponds to the free propagation of particles in an ideal gas (a similar equation exists for the variables t and $\mathbf{r}$, with a change of sign for the time derivative term).

Expression (21) allows shedding new light on the physical interpretation of the normal correlation function given in § 1-a, in terms of the creation, in the N-particle system, of a "hole", which then propagates until it is annihilated at a later time. In an ideal gas, each term of the sum in (21) is a free wave plane: in the absence of interactions, the particles can freely propagate along straight lines. The n_i appearing in the formula shows that the hole can only propagate along already populated individual states in the N-particle state: a hole can only be created in an already occupied quantum state, as pointed out already in § 1-a-β. As a result, the created hole is not a point-like object actually localized at point $\mathbf{r}'$: it is only built from superpositions of free occupied states, whereas for a truly point-like excitation, one would have to combine values of $\mathbf{k}_i$ extending to infinity.

[5] To prove this in a general way, we simply have to note that the operator $a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i}$ is translation invariant only if $\mathbf{k}_j = \mathbf{k}_i$. To show this, one can use the expression of the translation operator as an exponential of the operator associated with the total momentum (Complement E_{II}, § 3).

β. Anti-normal correlation function

For the anti-normal function $G_1^{AN}(\mathbf{r},t;\mathbf{r}',t')$, the calculation is practically the same, the only difference being the inversion of the order of the operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^{\dagger}$, which leads to:

$$G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') = \frac{1}{\mathcal{V}}\sum_i \left[1+\eta n_i\right]\ e^{i\left[\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})-\omega_{\mathbf{k}_i}(t'-t)\right]} \tag{23}$$

It obeys the same partial differential equation as the normal function. It follows from (21) and (23) that the linear combination:

$$G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') - \eta G_1^{N}(\mathbf{r},t;\mathbf{r}',t') = \frac{1}{\mathcal{V}}\sum_i e^{i\left[\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})-\omega_{\mathbf{k}_i}(t'-t)\right]} \tag{24}$$

is independent of the state of the system; for $t=t'$, it is simply equal to the function $\delta(\mathbf{r}-\mathbf{r}')$. We shall see later the relation between this expression and the spectral function.

Let us come back, here again, to the interpretation of the effect of the operator $\Psi^{\dagger}(\mathbf{r})$, discussed in § 1-a-β. For a fermion system ($\eta=-1$), relation (23) clearly shows that the particle creation does not involve any of the already occupied individual states $\mathbf{k}_i$, with $n_i=1$, since the corresponding term is zero. The object created by the excitation cannot have any component on the already occupied individual states, which simply means that its wave function must remain orthogonal to those of all the fermions already present. For a boson system, the effect is just the opposite: if, for instance, an individual state of bosons is highly populated compared to all the others, the term in (23) corresponding to this large n_i will be dominant: the additional particle will be mainly in the same individual state as all the other particles already present.

γ. Two-point Green's function

Using definition (10), we obtain the Green's function $\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t')$:

$$\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = \frac{1}{\mathcal{V}}\sum_i e^{i\left[\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})-\omega_{\mathbf{k}_i}(t'-t)\right]}\ \left\{Y\left(t'-t\right)\left[1+\eta n_i\right]+\eta Y\left(t-t'\right)n_i\right\} \tag{25}$$

If we take the derivative with respect to the time t', the two Heaviside step functions yield delta functions with opposite signs. We then get the partial differential equation:

$$\left[i\frac{\partial}{\partial t'}+\frac{\hbar}{2m}\Delta_{\mathbf{r}'}\right]\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = \frac{i\delta\left(t-t'\right)}{\mathcal{V}}\sum_i e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})}\left[1+\eta n_i-\eta n_i\right] \tag{26}$$

that is:

$$\left[i\frac{\partial}{\partial t'}+\frac{\hbar}{2m}\Delta_{\mathbf{r}'}\right]\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = i\ \delta\left(t-t'\right)\ \delta\left(\mathbf{r}'-\mathbf{r}\right) \tag{27}$$

As expected, the right-hand side is a product of delta functions of the set of variables characterizing a Green's function.

In the presence of interactions, this partial differential equation is no longer valid; we must add to its right-hand side the interaction contributions, which involve Green's functions of a higher order.

δ. Four-point functions

The computation of the four-point correlation functions is very similar to the one we just explained; we simply use again relation (17) to get their explicit expressions for an ideal gas contained in a box. The results are the same as those obtained in Complement A_{XVI}, as for example in relation (21), except for the fact that we must now multiply each spatial plane wave $e^{i\mathbf{k}_i\cdot\mathbf{r}}$ by the associated time evolution factor $e^{-i\omega_{\mathbf{k}_i}t}$.

2. Fourier transforms

From now on, we shall only study systems that are translation-invariant in space and time. Consequently, the correlation functions only depend on the difference of the variables $\mathbf{r}'-\mathbf{r}$ and $t'-t$, so that we can choose to cancel both $\mathbf{r}$ and t.

2-a. General definition

Let us introduce the two (double) Fourier transforms with respect to time and space:

$$\boxed{\overline{G}_1^{N,AN}(\mathbf{k},\omega)=\int \mathrm{d}^3r'\int \mathrm{d}t'\ e^{i(\omega t'-\mathbf{k}\cdot\mathbf{r}')}\ G_1^{N,AN}(\mathbf{0},0;\mathbf{r}',t')} \tag{28}$$

(we have set $\mathbf{r}=0$ and $t=0$ in the function $G_1^{N,AN}$). These functions are real, as shown by the parity relation (5).

The Fourier transform $\overline{\mathcal{G}}_1$ of the Green's function $\mathcal{G}_1$ introduced in (10) is called the "one-particle propagator"; it is defined by:

$$\boxed{\overline{\mathcal{G}}_1(\mathbf{k},\omega)=\int \mathrm{d}^3r'\int \mathrm{d}t'\ e^{i(\omega t'-\mathbf{k}\cdot\mathbf{r}')}\mathcal{G}_1(\mathbf{0},0;\mathbf{r}',t')} \tag{29}$$

For a system contained in a volume $\mathcal{V}$ with periodic boundary conditions (Complement F_{XI}), the integrals over d^3r' in the above formulas must be taken over the volume $\mathcal{V}$. They yield the coefficients of a Fourier series where the vectors $\mathbf{k}$ take the discrete values $\mathbf{k}_i$ corresponding to the boundary conditions. This series characterizes the spatial dependence. As for the time dependence, the Fourier transform is a continuous function[6]. The inverse transformation relations are:

$$G_1^{N,AN}(\mathbf{0},0;\mathbf{r}',t')=\frac{1}{\mathcal{V}}\sum_{\mathbf{k}_i}\int\frac{\mathrm{d}\omega}{2\pi}e^{i(\mathbf{k}_i\cdot\mathbf{r}'-\omega t')}\overline{G}_1^{N,AN}(\mathbf{k}_i,\omega) \tag{30}$$

where, in the limit of very large volumes, the discrete summation becomes an integral with the coefficient $\mathcal{V}/(2\pi)^3$:

$$G_1^{N,AN}(\mathbf{0},0;\mathbf{r}',t')=\int\frac{\mathrm{d}^3k}{(2\pi)^3}\int\frac{\mathrm{d}\omega}{2\pi}e^{i(\mathbf{k}\cdot\mathbf{r}'-\omega t')}\overline{G}_1^{N,AN}(\mathbf{k},\omega) \tag{31}$$

[6] As opposed to the space variables, the time variable is not confined to a finite variation domain, and time Fourier transforms may be singular; such an example, concerning an ideal gas, will be given in § 2-b, where we shall introduce, as a convergence factor, a decreasing exponential $e^{\pm\varepsilon t}$ (where ε tends toward zero through positive values).

Comment:

One can also express the functions $\overline{G}_1^{N,AN}(k,\omega)$, not as Fourier transforms as in (28), but directly from the average values of products of creation $a^\dagger_{\mathbf{k}_j}$ and annihilation $a_{\mathbf{k}_j}$ operators in the individual states $|\mathbf{k}\rangle$. Formulas (A-3) and (A-6) of Chapter XVI tell us that:

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_j e^{i\mathbf{k}_j\cdot\mathbf{r}} a_{\mathbf{k}_j} \qquad \text{and} \qquad \Psi^\dagger(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_{j'} e^{-i\mathbf{k}_{j'}\cdot\mathbf{r}} a^\dagger_{\mathbf{k}_{j'}} \tag{32}$$

Inserting these relations in the definitions (2), then in (28), we get:

$$\overline{G}_1^N(\mathbf{k}_i,\omega) = \frac{1}{\mathcal{V}} \int \mathrm{d}^3r' \int \mathrm{d}t'\ e^{i(\omega t' - \mathbf{k}_i\cdot\mathbf{r}')} \sum_{j,j'} e^{i\mathbf{k}_j\cdot\mathbf{r}'} \left\langle a^\dagger_{\mathbf{k}_{j'}} a^H_{\mathbf{k}_j}(t') \right\rangle \tag{33}$$

where $a^H_{\mathbf{k}_j}(t')$ is the operator $a_{\mathbf{k}_j}$ in the Heisenberg picture. The integral over the volume selects a single term, $j = i$, and cancels the volume $\mathcal{V}$. Furthermore, the translation invariance means that neither the density operator ρ, nor the system Hamiltonian, have matrix elements between state vectors with different total momentum. Now the operator $a^\dagger_{\mathbf{k}_{j'}}$ increases that total momentum by $\hbar k_{j'}$ and $a^H_{\mathbf{k}_j}(t')$ decreases it by $\hbar k_j$. The average value $\left\langle a^\dagger_{\mathbf{k}_{j'}} a^H_{\mathbf{k}_j}(t') \right\rangle$ is therefore zero unless j' is equal to j, and that eliminates the summation over j'. Finally:

$$\overline{G}_1^N(\mathbf{k}_i,\omega) = \int \mathrm{d}t'\ e^{i\omega t'} \left\langle a^\dagger_{\mathbf{k}_i} a^H_{\mathbf{k}_i}(t') \right\rangle \tag{34}$$

In a similar way, we can also show that:

$$\overline{G}_1^{AN}(\mathbf{k}_i,\omega) = \int \mathrm{d}t'\ e^{i\omega t'} \left\langle a^H_{\mathbf{k}_i}(t') a^\dagger_{\mathbf{k}_i} \right\rangle \tag{35}$$

2-b. Ideal gas example

For an ideal gas contained in a box (with periodic boundary conditions), the $\mathbf{k}$ are discrete. Replacing $\mathbf{k}$ by $\mathbf{k}_i$ in (28) and using expression (21), after replacing the dummy index $\mathbf{k}_i$ by $\mathbf{k}_j$, we get a product of exponentials to be integrated. The integral over d^3r', combined with the factor $1/\mathcal{V}$, yields a Kronecker delta $\delta_{\mathbf{k}_j,\mathbf{k}_i}$ that eliminates the summation over $\mathbf{k}_j$; the integral over $\mathrm{d}t'$ yields $2\pi\delta(\omega - \omega_{\mathbf{k}_i})$, and we get:

$$\overline{G}_1^N(\mathbf{k}_i,\omega) = 2\pi\ n_i\ \delta(\omega - \omega_{\mathbf{k}_i}) \tag{36}$$

In a similar way:

$$\overline{G}_1^{AN}(\mathbf{k}_i,\omega) = 2\pi\ [1 + \eta n_i]\ \delta(\omega - \omega_{\mathbf{k}_i}) \tag{37}$$

and we can finally write:

$$\overline{G}_1^{AN}(\mathbf{k}_i,\omega) - \eta\overline{G}_1^N(\mathbf{k}_i,\omega) = 2\pi\delta(\omega - \omega_{\mathbf{k}_i}) \tag{38}$$

These expressions are particularly simple, but no longer valid when the particles interact. They are, however, useful as a reference point to understand the interaction effects.

The one-particle propagator $\overline{\mathcal{G}}_1(\mathbf{k},\omega)$ defined in (29) can be obtained in a similar way; the integral over d^3r' is unchanged, but the integral over time now yields:

$$\int \mathrm{d}t'\ e^{i(\omega-\omega_{\mathbf{k}_i})t'}\ \{Y(t')\ [1+\eta n_i]+\eta Y(-t')\ n_i\} \tag{39}$$

which does not converge. For the term in $Y(t')$, a classical method is to introduce a convergence factor by changing ω into $\omega+i\varepsilon$, with $\varepsilon\to 0$ through positive values; for the $Y(-t')$ term, the change is to $\omega-i\varepsilon$. We get:

$$\begin{aligned}\overline{\mathcal{G}}_1(\mathbf{k}_i,\omega) &= -\frac{1+\eta n_i}{i\left[\omega-\omega_{\mathbf{k}_i}+i\varepsilon\right]}+\frac{\eta n_i}{i\left[\omega-\omega_{\mathbf{k}_i}-i\varepsilon\right]}\\ &= \frac{i}{\left[\omega-\omega_{\mathbf{k}_i}+i\varepsilon\right]}+\eta n_i\frac{2\varepsilon}{\left[\omega-\omega_{\mathbf{k}_i}\right]^2+\varepsilon^2}\end{aligned} \tag{40}$$

In the limit where $\varepsilon\to 0$, the two fractions on the right-hand side yield principal parts and delta functions – see relation (12) of Appendix II. The first term on the right-hand side yields both a principal part $\mathcal{P}\left[1/\left(\omega-\omega_{\mathbf{k}_i}\right)\right]$ and a delta function $\delta\left(\omega-\omega_{\mathbf{k}_i}\right)$, whereas the second one (in ηn_i) yields only a delta function $\delta\left(\omega-\omega_{\mathbf{k}_i}\right)$. If the system is diluted, it is in the classical regime (i.e. non-degenerate), where all the occupation numbers are small compared to 1 and where the exchange effects are weak; the second term, associated with the indistinguishability between the particles, is then negligible.

2-c. General expression in the presence of interactions

A translation invariant system has a Hamiltonian that commutes with its total momentum. We can then build a basis with state vectors that are, for each particle number N and for each value of the total momentum $\hbar\mathbf{K}$, eigenvectors of H with energies E; we shall note $E_{N,\mathbf{K}}$ these energies as it is often useful to explicitly keep track of the values of N and $\mathbf{K}$ that define the subspace corresponding to that eigenvalue E in the spectrum of H. We call $\left|\Phi^N_{E,K,\alpha}\right\rangle$ the corresponding eigenvectors, where the index α accounts for possible degeneracies of these eigenvalues.

We assume the system to be in a stationary state, and translation invariant. This means that its density operator ρ cannot have non-diagonal matrix elements between eigenvectors corresponding to different momenta or energies; in each of the subspaces common to both of these quantities, we can choose the basis $\left|\Phi^N_{E,K,\alpha}\right\rangle$ that diagonalizes the density operator and set:

$$\rho^N_{E,\mathbf{K},\alpha}=\left\langle\Phi^N_{E,\mathbf{K},\alpha}\right|\rho\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle \tag{41}$$

We now use this basis to compute the trace appearing in (2). We first insert expression (12) for the field operators (and their conjugates) as a function of the operators $a_{\mathbf{k}_i}$; taking into account the exponentials introduced by the operators in the Heisenberg picture, we get:

$$\begin{aligned}&G^N_1(0,0;\mathbf{r}',t')=\tfrac{1}{\mathcal{V}}\textstyle\sum_i\sum_j e^{i\mathbf{k}_j\cdot\mathbf{r}'}\sum_{N,\mathbf{K},E,\alpha}\sum_{N',\mathbf{K}',E',\alpha'}\rho^N_{E,\mathbf{K},\alpha}\\ &\times\left\langle\Phi^N_{E,\mathbf{K},\alpha}\right|a^\dagger_{\mathbf{k}_i}\left|\Phi^{N'}_{E',\mathbf{K}',\alpha'}\right\rangle\left\langle\Phi^{N'}_{E',\mathbf{K}',\alpha'}\right|a_{\mathbf{k}_j}\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle\ \mathrm{e}^{i\left[E'_{N',\mathbf{K}'}-E_{N,\mathbf{K}}\right]t'/\hbar}\end{aligned} \tag{42}$$

Several simplifications can be made on the right-hand side of this equation. First of all, as the operator $a_{\mathbf{k}_j}$ destroys a particle, we necessarily have $N'=N-1$, or else the

matrix element of $a_{\mathbf{k}_j}$ would be zero; the sum over N' disappears. Along the same line, as this operator decreases the total momentum by $\hbar\mathbf{k}_j$, we must also have $\mathbf{K}' = \mathbf{K} - \mathbf{k}_j$, and the sum over $\mathbf{K}'$ is also eliminated. Now, for the matrix element of $a^\dagger_{\mathbf{k}_i}$ to be non-zero, and since this operator adds $\hbar\mathbf{k}_i$ to the momentum $\hbar\,(\mathbf{K} - \mathbf{k}_j)$, to recover the initial momentum $\hbar\mathbf{K}$, we necessarily have $\mathbf{k}_i = \mathbf{k}_j$. Once these simplifications have been made, we insert the result in definition (28) of $\overline{G}_1^N(\mathbf{k},\omega)$, and get:

$$\begin{array}{l}\overline{G}_1^N(\mathbf{k},\omega) = \frac{1}{\mathcal{V}}\int \mathrm{d}^3r' \int \mathrm{d}t'\ e^{i\left(\omega t' - \mathbf{k}\cdot\mathbf{r}'\right)} \sum_j e^{i\mathbf{k}_j\cdot\mathbf{r}'} \\ \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{E,\mathbf{K},\alpha}\ \left|\left\langle \Phi^N_{E,\mathbf{K},\alpha}\right| a^\dagger_{\mathbf{k}_i}\left|\Phi^{N-1}_{E',\mathbf{K}-\mathbf{k},\alpha'}\right\rangle\right|^2\ \mathrm{e}^{i\left[E'_{N-1,\mathbf{K}-\mathbf{k}} - E_{N,\mathbf{K}}\right]t'/\hbar}\end{array} \tag{43}$$

On the first line, the integral over d^3r' yields a Kronecker delta that forces $\mathbf{k}_j = \mathbf{k}$; the integral then yields $\mathcal{V}$, which cancels the volume appearing in the denominator. The time integral in relation (43) yields a 2π coefficient multiplied by a delta function of the variable $\omega + \left[E'_{N-1,\mathbf{K}-\mathbf{k}} - E_{N,\mathbf{K}}\right]/\hbar$. We finally get:

$$\begin{array}{l}\overline{G}_1^N(\mathbf{k},\omega) = \\ 2\pi \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{E,\mathbf{K},\alpha}\ \left|\left\langle \Phi^N_{E,\mathbf{K},\alpha}\right| a^\dagger_{\mathbf{k}}\left|\Phi^{N-1}_{E',\mathbf{K}-\mathbf{k},\alpha'}\right\rangle\right|^2\ \delta\left[\omega - \frac{E_{N,\mathbf{K}} - E'_{N-1,\mathbf{K}-\mathbf{k}}}{\hbar}\right]\end{array} \tag{44}$$

The same type of calculation also yields:

$$\begin{array}{l}\overline{G}_1^{AN}(\mathbf{k},\omega) = \\ 2\pi \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{E,\mathbf{K},\alpha}\ \left|\left\langle \Phi^N_{E,\mathbf{K},\alpha}\right| a_{\mathbf{k}}\left|\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right\rangle\right|^2\ \delta\left[\omega - \frac{E'_{N+1,\mathbf{K}+\mathbf{k}} - E_{N,\mathbf{K}}}{\hbar}\right]\end{array} \tag{45}$$

Let us assume, in addition, that our system is at thermal equilibrium, and described by the grand canonical density operator:

$$\rho = \frac{1}{Z} e^{-\beta\left(H - \mu\widehat{N}\right)} \tag{46}$$

with the classical notation: $\widehat{N}$ is the total particle number, $\beta = 1/k_BT$ (k_B is the Boltzmann constant) and $Z = \mathrm{Tr}\left\{e^{-\beta\left(H-\mu\widehat{N}\right)}\right\}$ is the grand canonical partition function. The two functions $\overline{G}_1^N$ and $\overline{G}_1^{AN}$ then obey a simple relation, often called "boundary condition":

$$\boxed{\overline{G}_1^{AN}(\mathbf{k},\omega) = e^{\beta(\hbar\omega-\mu)}\overline{G}_1^N(\mathbf{k},\omega)} \tag{47}$$

This relation turns out to be crucial in many calculations involving Green's functions; we shall use it in § 3.

Demonstration:

To establish this relation, we rewrite equality (45) using the fact that $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^{\dagger}$ are Hermitian conjugates. We then get:

$$\overline{G}_1^{AN}(\mathbf{k},\omega) = 2\pi \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{\mathbf{K},E,\alpha} \times \\ \times \left| \left\langle \Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'} \right| a_{\mathbf{k}}^{\dagger} \left| \Phi^N_{E,\mathbf{K},\alpha} \right\rangle \right|^2 \delta \left[\omega - \frac{E'_{N+1,\mathbf{K}+\mathbf{k}} - E_{N,\mathbf{K}}}{\hbar} \right] \tag{48}$$

We now permute the indices E' and E, as well as the indices α and α', we change the dummy summation index N to $N' = N + 1$, and finally replace the dummy variable K by $K' = K + k$. This yields:

$$\overline{G}_1^{AN}(\mathbf{k},\omega) = 2\pi \sum_{N',\mathbf{K}'} \sum_{E,E',\alpha,\alpha'} \rho^{N'-1}_{\mathbf{K}'-\mathbf{k},E',\alpha'} \times \\ \times \left| \left\langle \Phi^{N'}_{E,\mathbf{K}',\alpha} \right| a_{\mathbf{k}}^{\dagger} \left| \Phi^{N'-1}_{E',\mathbf{K}'-\mathbf{k},\alpha'} \right\rangle \right|^2 \delta \left[\omega - \frac{E_{N',\mathbf{K}'} - E'_{N'-1,\mathbf{K}'-\mathbf{k}}}{\hbar} \right] \tag{49}$$

In this summation, just as in (44), the lowest value of the index N' that gives a contribution is $N' = 1$; we therefore get the same expression as (44), with (aside from the irrelevant change of the dummy index N into N') just one modification, the replacement of $\rho^N_{\mathbf{K},E,\alpha}$ by $\rho^{N-1}_{\mathbf{K}'-\mathbf{k},E',\alpha'}$. However, since:

$$\rho^N_{\mathbf{K},E,\alpha} = \frac{1}{Z} e^{-\beta\left[E_{N,\mathbf{K}} - \mu N\right]} \tag{50}$$

the ratio of these two diagonal elements $\rho^{N-1}_{\mathbf{K}'-\mathbf{k},E',\alpha'} / \rho^N_{\mathbf{K},E,\alpha}$ introduces in the integral a factor:

$$e^{\beta\left[-E'_{N-1,\mathbf{K}'-\mathbf{k}} + \mu(N-1) + E_{N,\mathbf{K}} - \mu N\right]} = e^{\beta\left[E_{N,\mathbf{K}} - E'_{N-1,\mathbf{K}'-\mathbf{k}} - \mu\right]} \tag{51}$$

Now, the delta function in (49) allows replacing, in the exponent of this factor, the energy difference by $\hbar\omega$:

$$e^{\beta\left[E_{N,\mathbf{K}} - E_{N-1,\mathbf{K}'-\mathbf{k}} - \mu\right]} \Longrightarrow e^{\beta[\hbar\omega - \mu]} \tag{52}$$

This factor comes out of the summation and relation (47) is established.

2-d. Discussion

Expressions (44) and (45) give an idea of the behavior of the functions $\overline{G}_1^N$ and $\overline{G}_1^{AN}$. For an ideal gas, relations (36) and (37) show that they are singular functions, actually delta functions forcing the energy $\hbar\omega$ to take on exactly the one-particle kinetic energy $\hbar^2 k^2/2m$. This is because, in an ideal gas, one can choose Fock states as stationary states $\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$, where each individual state, with a given momentum $\hbar\mathbf{k}$, has a well defined population. In such a case, the operator $a_{\mathbf{k}}$ in expression (45) of $\overline{G}_1^{AN}$ can only couple, via its matrix element $\left\langle \Phi^N_{E,\mathbf{K},\alpha} \right| a_{\mathbf{k}} \left| \Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'} \right\rangle$, one state $\left|\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right\rangle$ to the initial state $\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$: the Fock state, where the occupation number $n_{\mathbf{k}}$ is larger by one unit. Consequently, the energy difference $E'_{N+1,\mathbf{K}+\mathbf{k}} - E_{N,\mathbf{K}}$ always takes the same value, the energy $\hbar^2 k^2/2m$ of an individual particle with momentum $\hbar\mathbf{k}$. We find again the results of § 2-b.

Let us assume now that the system includes mutual interactions. The matrix element $\left\langle \Phi^N_{E,\mathbf{K},\alpha} \right| a_{\mathbf{k}} \left| \Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'} \right\rangle$ then yields the scalar product of an N-particle system

stationary state with another state where a particle with a well defined momentum $\hbar\mathbf{k}$ has been removed from a stationary state of a system with $N+1$ particles. This new state has no reason to also be stationary, because of the particle interactions. It is therefore expected to have a non-zero scalar product with a whole series of kets $\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$ with different energies, which can become more numerous as the system enlarges. The sum over E no longer reduces to a single value of the energy; in the limit of large systems, it becomes a continuous sum, which absorbs the delta function, so that the function $\overline{G}_1^{AN}(\mathbf{k},\omega)$ takes, for a given $\mathbf{k}$, non-negligible values in a whole ω domain. A priori, its variations can take any value in this domain; however, if the interactions are not too strong and by comparison with the ideal gas, we expect the scalar product to have non-negligible values mostly in an E' energy domain close to $E_{N,\mathbf{K}}+\hbar^2k^2/2m$. For each $\mathbf{k}$, the interaction effect is to enlarge the energy peak, infinitely narrow for an ideal gas, by giving it a width that increases as the interaction becomes stronger. This width is interpreted in terms of a finite lifetime of the excitation created when a free particle is removed from a stationary state of the $(N+1)$-particle system; as the moduli of the two matrix elements $\left\langle\Phi^N_{E,\mathbf{K},\alpha}\right| a_{\mathbf{k}} \left|\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right\rangle$ and $\left\langle\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right| a^\dagger_{\mathbf{k}} \left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$ are equal, this lifetime can also be interpreted as that of the excitation created when a free particle is added to a stationary state of the N-particle system.

3. Spectral function, sum rule

In view of formula (7), it is natural to introduce the real function $A(\mathbf{k},\omega)$ as:

$$\boxed{\begin{aligned} A(\mathbf{k},\omega) &= \overline{G}_1^{AN}(\mathbf{k},\omega) - \eta\overline{G}_1^{N}(\mathbf{k},\omega) && \text{(53a)}\\ &= \int \mathrm{d}t'\ e^{i\omega t'} \left\langle \left[a^H_{\mathbf{k}_i}(t'), a^\dagger_{\mathbf{k}_i}\right]_{-\eta} \right\rangle && \text{(53b)} \end{aligned}}$$

where we have used (34) and (35). We call $A(\mathbf{k},\omega)$ the "spectral function"; it is real since both functions $\overline{G}_1^N$ and $\overline{G}_1^{AN}$ are real.

For an ideal gas, formula (38) shows that we simply have:

$$A(\mathbf{k},\omega) = 2\pi\ \delta(\omega-\omega_k) \tag{54}$$

but this equality is no longer valid as soon as the particles interact. We are going to show, however, that the spectral function allows expressing the correlation functions with a general formula, reminiscent of that established for an ideal gas, even in the presence of interactions.

3-a. Expression of the one-particle correlation functions

Inserting (47) in the definition of $A(\mathbf{k},\omega)$, we get:

$$A(\mathbf{k},\omega) = \overline{G}_1^N \left[e^{\beta(\hbar\omega-\mu)} - \eta\right] \tag{55}$$

that is:

$$\overline{G}_1^N(\mathbf{k},\omega) = A(\mathbf{k},\omega)\ f^\eta_\beta(\hbar\omega-\mu) \tag{56}$$

where $f_\beta^\eta(\hbar\omega - \mu)$ is the usual Bose-Einstein distribution for bosons, and the Fermi-Dirac distribution for fermions:

$$f_\beta^\eta(\hbar\omega - \mu) = \frac{1}{e^{\beta(\hbar\omega-\mu)} - \eta} \tag{57}$$

Using again (47), we get:

$$\overline{G}_1^{AN}(\mathbf{k},\omega) = e^{\beta(\hbar\omega-\mu)} A(\mathbf{k},\omega)\ f_\beta^\eta(\hbar\omega-\mu) = A(\mathbf{k},\omega)\ \left[1 + \eta f_\beta^\eta(\hbar\omega-\mu)\right] \tag{58}$$

Knowing the spectral function $A(\mathbf{k},\omega)$, we can determine the one-particle correlation functions with formulas similar to those of the ideal gas, containing the same quantum distribution functions f_β^η. Note, however, that in the presence of interactions, the energy $\hbar\omega$ and momentum $\hbar\mathbf{k}$ variables are independent, whereas for an ideal gas, they are constrained by relation (54).

3-b. Sum rule

We now insert relations (28) and (2) in (53a); since the operators Ψ and Ψ_H coincide at $t = 0$, we get:

$$A(\mathbf{k},\omega) = \int \mathrm{d}^3 r' \int \mathrm{d}t'\ e^{i(\omega t' - \mathbf{k}\cdot\mathbf{r}')}\ \mathrm{Tr}\left\{\rho\ \left[\Psi_H(\mathbf{r}',t')\ ,\ \Psi^\dagger(\mathbf{r}=\mathbf{0})\right]_{-\eta}\right\} \tag{59}$$

Taking a summation over ω, a time delta function is introduced:

$$\int \mathrm{d}\omega\ A(\mathbf{k},\omega) = 2\pi \int \mathrm{d}^3 r'\ e^{-i\mathbf{k}\cdot\mathbf{r}'}\ \mathrm{Tr}\left\{\rho\ \left[\Psi(\mathbf{r}')\ ,\ \Psi^\dagger(\mathbf{r}=\mathbf{0})\right]_{-\eta}\right\} \tag{60}$$

that is:

$$\int \frac{\mathrm{d}\omega}{2\pi}\ A(\mathbf{k},\omega) = \int \mathrm{d}^3 r'\ e^{-i\mathbf{k}\cdot\mathbf{r}'}\ \delta(\mathbf{r}') = 1 \tag{61}$$

In the presence of interactions, we do not know, a priori, how the spectral function depends on $\mathbf{k}$ and ω. However, for each value of $\mathbf{k}$, the interactions can only modify the frequency distribution, but not its integral over ω.

We have seen that for a gas of interacting particles, the effect of the interactions is to "spread" the spectral function over a certain domain of frequencies, all the while obeying the sum rule (61). There is no reasons for the spectral function to present any particular shape, or still contain delta functions of frequency. It often happens, however, that it presents marked peaks, whose narrowness signals the existence of excitations in the system, behaving almost like free particles (with long lifetimes). These excitations are called "quasi-particles" as they are, in a way, an extension – once the interactions have been introduced – of the independent particles of the ideal gas. The peak associated with a particle of momentum $\hbar\mathbf{k}$ is not, in general, centered on the energy $e_k = \hbar^2 k^2/2m$ of a free particle: in addition to the spreading, the energies of the quasi-particles are shifted by the interactions. This spreading and shifting is reminiscent of the results obtained in Complement D_{XIII}, where we studied, to lowest order, the coupling of a discrete state with a continuum. In other words, one can say that the interactions couple the state of a free particle with momentum $\hbar\mathbf{k}$, to a continuum of states having different momenta,

which explains the analogy. Note, however, that the present results concern not a single but an ensemble of identical particles, and its properties at thermal equilibrium; the physical situation is therefore different.

To sum up, to go from the ideal gas (where ω is necessarily equal to ω_k) to an interacting gas, we simply introduce in the Green's function, and for each value of $\mathbf{k}$, a weighting by a spectral function $A(\mathbf{k},\omega)$; this function distributes the ω dependence over a certain frequency domain. We are going to show that from the spectral function we can infer many properties of an interacting physical system. But this obviously does not mean that the spectral function is easy to compute! On the contrary, in most interacting physical systems, we do not know its exact value. Its very existence, however, independently of its precise mathematical form, is a very useful conceptual tool.

3-c. Expression of various physical quantities

The spectral function contains information on a great number of physical properties of interacting systems, in a form much more concise than the density operator of the N-body system: that operator contains everything but is mathematically far more complicated than a simple function.

Consider first the particle density, given by:

$$n(\mathbf{r}) = \left\langle \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) \right\rangle = G_1^N(\mathbf{r},0;\mathbf{r},0) \tag{62}$$

according to the definition (2) of the normal function G_1^N. For a translation invariant system, this density is independent of $\mathbf{r}$, and we can set $\mathbf{r}=0$. The density is then the value, at the origin, of the normal function (31), that is, taking (56) into account:

$$n = \int \frac{\mathrm{d}^3k}{(2\pi)^3} \int \frac{\mathrm{d}\omega}{2\pi}\, \overline{G}_1^N(\mathbf{k},\omega) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} \int \frac{\mathrm{d}\omega}{2\pi}\, A(\mathbf{k},\omega)\; f_\beta^\eta(\hbar\omega-\mu) \tag{63}$$

Let us now study a quantity furnishing more precise information, the particle momentum distribution, and compute the average number $\langle n_\mathbf{k}\rangle$ of particles having a momentum $\hbar\mathbf{k}$. We assume the system is contained in a cubic box of volume $\mathcal{V}$. Relations (A-9) and (A-10) of Chapter XVI, applied to the case where the basis wave functions are given by (12), yield:

$$\langle n_\mathbf{k}\rangle = \left\langle a_\mathbf{k}^\dagger a_\mathbf{k} \right\rangle = \frac{1}{\mathcal{V}} \int \mathrm{d}^3r' \int \mathrm{d}^3r\; e^{i\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})} \left\langle \Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}) \right\rangle \tag{64}$$

Replacing the integral variable $\mathbf{r}$ by $\mathbf{s}=\mathbf{r}-\mathbf{r}'$, the average value $\left\langle \Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}'+\mathbf{s}) \right\rangle$ appears, which is independent of $\mathbf{r}'$ because of the translation invariance; we can then replace, in this average value, $\mathbf{r}'$ by zero, and the integral over d^3r' simply yields the volume $\mathcal{V}$. We are left with:

$$\langle n_\mathbf{k}\rangle = \int \mathrm{d}^3s\; e^{-i\mathbf{k}\cdot\mathbf{s}} \left\langle \Psi^\dagger(\mathbf{0})\Psi(\mathbf{s}) \right\rangle \tag{65}$$

Now, performing the integral over ω of the definition (28) of $\overline{G}_1^N(\mathbf{k},\omega)$, and taking (2)

into account, we get:

$$\int \mathrm{d}\omega\ \overline{G}_1^N(\mathbf{k},\omega) = 2\pi \int \mathrm{d}^3 r' e^{-i\mathbf{k}\cdot\mathbf{r}'} G_1^N(0,0;\mathbf{r}',0)$$
$$= 2\pi \int \mathrm{d}^3 r' e^{-i\mathbf{k}\cdot\mathbf{r}'} \left\langle \Psi^\dagger(\mathbf{0})\Psi(\mathbf{r}')\right\rangle \tag{66}$$

which is identical to (65) within a factor of 2π. It then follows that[7]:

$$\langle n_{\mathbf{k}}\rangle = \frac{1}{2\pi}\int \mathrm{d}\omega\ \overline{G}_1^N(\mathbf{k},\omega) = \int \frac{\mathrm{d}\omega}{2\pi} A(\mathbf{k},\omega)\ f_\beta^\eta(\hbar\omega-\mu) \tag{67}$$

In the same way, one can show that the system average energy[8] per unit volume $\langle H\rangle/\mathcal{V}$ is given by:

$$\frac{\langle H\rangle}{\mathcal{V}} = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\int \frac{\mathrm{d}\omega}{2\pi}\left[\hbar\frac{\omega+\omega_k}{2}\right]\ A(\mathbf{k},\omega) f_\beta^\eta(\hbar\omega-\mu) \tag{68}$$

where ω_k is defined in (14). Once this function is known, one can get, by integration over β, the logarithm of the partition function, which in turn yields, by derivation, all the thermodynamic quantities (particle density, pressure, etc.). It is remarkable that the spectral function, whose definition comes from the one-particle Green's functions and could therefore be expected to only contain information on the one-particle density operator, actually allows computing all these physical quantities that depend on the correlations between the particles, and hence on their interactions. With this method, the study of N-body properties is reduced to the computation of functions mathematically defined for a single particle. It generalizes, in a way, the ideal gas equations, while taking rigorously into account the presence of an ensemble of indistinguishable particles at thermal equilibrium. It is therefore quite powerful.

Nevertheless, this obviously does not solve the problem of calculating the equilibrium properties of an interacting system; in practice, getting precise values for the spectral function can pose a very difficult mathematical problem. Numerous approximation methods have been developed to try and resolve this, using in particular the concept of "self energy" and of perturbation diagrams, but this is beyond the scope of this complement.

[7] For an ideal gas of bosons, the chemical potential is always below the lowest individual energy, so that the distribution function f_β^η never diverges. As in relation (67) ω is integrated from $-\infty$ to $+\infty$, this divergence now seems unavoidable. Relation (55), however, shows that the spectral function $A(\mathbf{k},\omega)$ for bosons goes to zero for $\hbar\omega=\mu$, as long as the function $\overline{G}_1^N$ remains regular at this point. Consequently, integrals (63), (67) and (68) do not present any divergences associated with that ω value

[8] Relation (68) can be demonstrated by first computing the time evolution of $\Psi_H(\mathbf{r},t)$ and $\Psi_H^\dagger(\mathbf{r},t')$ to get the expression for $i\hbar\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial t'}\right)\Psi_H^\dagger(\mathbf{r},t')\Psi_H(\mathbf{r},t)$; one then takes its average value and performs a Fourier transformation to get the average value of the energy – see for example § 2.2 of reference [7].

Complement C_{XVI}

Wick's theorem

1 Demonstration of the theorem 1799
1-a Statement of the problem 1799
1-b Recurrence relation 1800
1-c Contractions 1802
1-d Statement of the theorem 1804
2 Applications: correlation functions for an ideal gas 1804
2-a First order correlation function 1805
2-b Second order correlation functions 1806
2-c Higher order correlation functions 1808

For an ideal gas at thermal equilibrium, we computed in Complement B_{XV} the average values of one- and two-particle operators, and showed they could all be expressed in terms of the one-body quantum distributions f_β (the Fermi-Dirac distribution for fermions, and Bose-Einstein for bosons). In this complement we establish a theorem that allows generalizing those results to operators involving any number of particles. The demonstration of Wick's theorem is explained in § 1, and will be applied in § 2 to the calculation of correlation functions in an ideal gas.

1. Demonstration of the theorem

Let us consider an ideal gas at thermal equilibrium, described by the grand canonical ensemble (Appendix VI, § 1-c), with the density operator:

$$\rho_{eq} = \frac{1}{Z} e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \tag{1}$$

where $\beta = 1/k_B T$ is the inverse of the temperature multiplied by the Boltzmann constant k_B, μ the chemical potential, and $\widehat{H}$ the Hamiltonian:

$$\widehat{H} = \sum_{\mathbf{k}} e_k \; a_k^\dagger a_k \qquad \text{with:} \qquad e_k = \frac{\hbar^2 k^2}{2m} \tag{2}$$

The grand canonical partition function Z is defined by:

$$Z = \mathrm{Tr}\left\{ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \right\} \tag{3}$$

1-a. Statement of the problem

We wish to calculate the average value $\langle A_P \rangle$ of a product of operators:

$$A_P = b_1 b_2 .... b_P \tag{4}$$

where each of the operators b_k is, either an annihilation operator a_k, or a creation operator $a_k^\dagger$:

$$b_k = \begin{cases} a_k \\ a_k^\dagger \end{cases} \tag{5}$$

Taking into account relations (A-48) and (A-49) of Chapter XV, we have:

$$\begin{aligned} [b_i, b_j]_{-\eta} &= b_i b_j - \eta b_j b_i \\ &= \begin{cases} 0 \text{ if } b_i \text{ and } b_j \text{ are both creation or annihilation operators} \\ \delta_{ij} \text{ if } b_i \text{ is an annihilation, and } b_j \text{ a creation operator} \\ -\eta\delta_{ij} \text{ if } b_i \text{ is a creation, and } b_j \text{ an annihilation operator} \end{cases} \end{aligned} \tag{6}$$

with $\eta = +1$ for bosons and $\eta = -1$ for fermions.

Assuming the quantum state is described by the density operator ρ_{eq} given in (1), the average value of A_P is:

$$\langle A_P \rangle = \mathrm{Tr}\,\{\rho_{eq}\ b_1 b_2 .... b_P\} \tag{7}$$

As ρ_{eq} is diagonal in the basis of the Fock states associated with the a_k, this average value will be different from zero only if the series of operators b contains, for each creation operator $a_k^\dagger$, an annihilation operator a_k in that same individual state; they must exactly balance one another and must therefore appear the same number of times. In particular, the average value $\langle A_P \rangle$ will always be zero if P is odd; from now on, we shall assume that $P = 2n$, n being an integer.

1-b. Recurrence relation

We have to compute:

$$\langle A_{2n} \rangle = \mathrm{Tr}\,\{\rho_{eq}\ b_1 b_2 .... b_{2n}\} \tag{8}$$

We first start by changing the order of b_1 and b_2, using one of the relations (6); we will then continue to progressively shift b_1 towards the right, by permuting it first with b_3, then with b_4, etc. until the permutation with b_{2n} brings it to the very last position. As a trace is invariant under a circular permutation, the operator b_1 can then be moved back all the way to the first position, ahead of ρ_{eq}; a last commutation with ρ_{eq}, that we compute just below, returns it to its initial position, and allows computing the value of $\langle A_{2n} \rangle$ as a function of the average values of a product of $2(n-1)$ operators. The computation goes as follows:

$$\begin{aligned} \langle A_{2n} \rangle &= \mathrm{Tr}\left\{\rho_{eq}\ \ [b_1, b_2]_{-\eta}\, b_3 b_4 ... b_{2n}\right\} + \eta\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_1 b_3 b_4 ... b_{2n}\} \\ &= [b_1, b_2]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_3 b_4 ... b_{2n}\} + \eta\,[b_1, b_3]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_4 ... b_{2n}\} \\ &\qquad + \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_3 b_1 b_4 ... b_{2n}\} \\ &= [b_1, b_2]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_3 b_4 ... b_{2n}\} + \eta\,[b_1, b_3]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_4 ... b_{2n}\} + ... \\ &\qquad ... + \eta^{2n-2}\,[b_1, b_{2n}]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_3 b_4 ... b_{2n-1}\} \\ &\qquad\qquad + \eta^{2n-1}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_3 b_4 ... b_{2n} b_1\} \end{aligned} \tag{9}$$

Most of the terms on the right-hand side are in general zero: the first is non-zero only if b_1 and b_2 are two conjugated operators (an annihilation and a creation operator associated with the same individual quantum state); the second is non-zero if this is also the case for operators b_1 and b_3, etc.

After a circular permutation under the trace, the last term of the sum can be written as:

$$\eta^{2n-1}\mathrm{Tr}\left\{b_1\rho_{eq}\ b_2b_3....b_{2n}\right\} \tag{10}$$

We shall now relate both operators $b_1\rho_{eq}$ and $\rho_{eq}b_1$, showing that they are proportional to each other. Assume, for example, that b_1 is a creation operator $a_q^\dagger$; in the operator ρ_{eq} defined in (1), all the terms $k \neq q$ commute with $a_q^\dagger$ and the change of order for the operators leads to two expressions, to be compared:

$$a_q^\dagger\ e^{-\beta(e_q-\mu)\ a_q^\dagger a_q} \qquad \text{and} \quad e^{-\beta(e_q-\mu)\ a_q^\dagger a_q}\ a_q^\dagger \tag{11}$$

By action on the Fock vectors, it is easy to check (as we assumed the system was in thermal equilibrium) that:

$$a_q^\dagger\ e^{-\beta(e_q-\mu)\ a_q^\dagger a_q} = e^{\beta(e_q-\mu)} \times e^{-\beta(e_q-\mu)\ a_q^\dagger a_q}\ a_q^\dagger \tag{12}$$

which leads to:

$$a_q^\dagger\ \rho_{eq} = e^{\beta(e_q-\mu)}\rho_{eq}\ a_q^\dagger \tag{13}$$

If b_1 is an annihilation operator, the same reasoning shows that the change of order introduces the inverse factor: $e^{-\beta(e_n-\mu)}$. To sum up:

$$b_1\rho_{eq} = e^{\pm\beta(e_1-\mu)}\rho_{eq}b_1 \tag{14}$$

with a + sign in the exponential if b_1 is a creation operator, and a − sign if b_1 is an annihilation operator. Consequently, the last term in (9) is equal to $e^{\pm\beta(e_1-\mu)}\left\langle A_{2n}\right\rangle$, with a factor $\eta^{2n-1} = \eta$ since $\eta = \pm 1$.

Moving this last term to the left-hand side, we get:

$$\begin{aligned}\left\langle A_{2n}\right\rangle\left[1-\eta e^{\pm\beta(e_1-\mu)}\right] &= \left[b_1,b_2\right]_{-\eta}\mathrm{Tr}\left\{\rho_{eq}\ b_3...b_{2n}\right\}\\ &+\eta\left[b_1,b_3\right]_{-\eta}\mathrm{Tr}\left\{\rho_{eq}\ b_2b_4...b_{2n}\right\}+...\\ &+\eta^{2n-2}\left[b_1,b_{2n}\right]_{-\eta}\mathrm{Tr}\left\{\rho_{eq}\ b_2b_3...b_{2n-1}\right\}\end{aligned} \tag{15}$$

On the right-hand side of this equality, all the (anti)commutators $\left[b_1,b_2\right]_{-\eta}$ are actually numbers, and many of them are zero: as before, the only non-zero ones are those for which the two concerned operators b are conjugates of each other (a creation and an annihilation operator for the same individual quantum state). The average value of the product of $2n$ operators we are looking for can therefore be expressed as a linear combination of average values of products of $2n-2$ operators.

1-c. Contractions

We now define the "contraction" of two operators b_i and b_j as the number, written $\overbrace{b_i b_j}$, defined by:

$$\overbrace{b_i b_j} = \frac{1}{1-\eta e^{\pm\beta(e_i-\mu)}} \times [b_i, b_j]_{-\eta} = -\eta f_\beta\left(\pm(e_i-\mu)\right) \times [b_i, b_j]_{-\eta} \tag{16}$$

where, as above, in the denominator a + sign is chosen in the exponential if b_1 is a creation operator, and a − sign if b_1 is an annihilation operator. The function f_β is the Fermi-Dirac distribution f_β^{FD} for fermions, and that of Bose-Einstein f_β^{BE} for bosons:

$$f_\beta(e-\mu) = \frac{1}{e^{\beta(e-\mu)}-\eta} \tag{17}$$

The contraction is zero if it concerns two operators b_i and b_j acting on different individual quantum states; it is also zero if the operators are both creation or both annihilation operators in the same individual quantum state. If b_i is the creation operator, and b_j the annihilation operator in the same individual state, the contraction is simply equal to the distribution function $f_\beta(e_i-\mu)$ since:

$$\overbrace{b_i b_j} = \overbrace{a_i^\dagger a_i} = \frac{-\eta}{1-\eta e^{+\beta(e_i-\mu)}} = f_\beta(e_i-\mu) \tag{18}$$

In the opposite case (antinormal order), the contraction is given by:

$$\overbrace{b_i b_j} = \overbrace{a_i a_i^\dagger} = \frac{1}{1-\eta e^{-\beta(e_i-\mu)}} = 1+\eta f_\beta(e_i-\mu) \tag{19}$$

Relation (15) can thus be rewritten as:

$$\langle A_{2n}\rangle = \overbrace{b_1 b_2}\langle\, b_3 ... b_{2n}\rangle + \eta\, \overbrace{b_1 b_3}\langle b_2 b_4 ... b_{2n}\rangle + ... + \eta^{2n-2}\overbrace{b_1 b_{2n}}\langle b_2 b_3 ... b_{2n-1}\rangle \tag{20}$$

where the traces have been replaced with quantum averages.

We shall then reason by recurrence: each of the average values on the right-hand side of (20) is of the same type as $\langle A_{2n}\rangle$ written in (8), except for the fact that n has been lowered by one unit. Dealing with each of the average values $\langle A_{2n-2}\rangle$ as we did for $\langle A_{2n}\rangle$, that last average value now appears as a double sum of terms containing two contractions and average values $\langle A_{2n-4}\rangle$. Continuing as many times as necessary, we end up with an average value $\langle A_{2n}\rangle$ expressed as the sum of diverse products of n contractions.

As an illustration, let us consider a few simple examples. If $n=1$, we get directly:

$$\boxed{\langle b_1 b_2\rangle = \overbrace{b_1 b_2}} \tag{21}$$

This simple relation can actually be used as a definition of contractions, instead of (16). If b_1 is a creation operator and b_2 the corresponding annihilation operator, we get the result (18), equivalent to relations (19) and (23) of Complement B_{XV}; if the operator's order is reversed, we get (19) which comes directly from the previous result and from the

commutation or anticommutation relation (6). For all the other cases, we find zero on each side of the equality.

If $n = 2$, we use a first time relation (20), and obtain:

$$\langle b_1 b_2 b_3 b_4 \rangle = \overbrace{b_1 b_2} \langle b_3 b_4 \rangle + \eta\, \overbrace{b_1 b_3} \langle b_2 b_4 \rangle + \overbrace{b_1 b_4} \langle b_2 b_3 \rangle \tag{22}$$

Using again this same relation, we compute each of the average values of the product of two operators b, which yields:

$$\begin{aligned} \langle b_1 b_2 b_3 b_4 \rangle &= \overbrace{b_1 b_2}\, \overbrace{b_3 b_4} + \eta\, \overbrace{b_1 b_3}\, \overbrace{b_2 b_4} + \overbrace{b_1 b_4}\, \overbrace{b_2 b_3} \\ &= \overbrace{b_1 b_2 b_3 b_4} + \overbrace{b_1 \underbrace{b_2 b_3} b_4} + \overbrace{b_1 \underbrace{b_2 b_3 b_4}} \end{aligned} \tag{23}$$

In the second line, we have used a generalization of the notation of products of contractions. When two operators inside a contraction are separated, a permutation is needed to group them. For fermions, this introduces a sign given by the parity of the required permutation, but no sign change for bosons. When two contractions are embedded, we group together all pairs of operators belonging to the same contraction and, for fermions, we multiply by the parity of the corresponding permutation[1]; for bosons, no sign change is introduced. In the present case, we therefore have :

$$\overbrace{b_1 \underbrace{b_2 b_3} b_4} = \eta\, \overbrace{b_1 b_3}\, \overbrace{b_2 b_4} \qquad \text{and} \qquad \overbrace{b_1 \underbrace{b_2 b_3 b_4}} = \overbrace{b_1 b_4}\, \overbrace{b_2 b_3} \tag{24}$$

The final result (23) only contains products of contractions, i.e. of distribution functions. One can easily check that, among those three products, a maximum of two are non-zero.

Comments:

(i) Another notation is frequently used, where operators and contractions are embedded in the same average value, for instance:

$$\langle b_1 b_2 ... \overbrace{b_i ... b_j ... b_l} ... b_{2n} \rangle = (\varepsilon_\alpha)\, \overbrace{b_i b_l} \langle b_1 b_2 ... b_j ... b_{2n} \rangle \tag{25}$$

where, for fermions, ε_α is the parity of the permutation needed to bring operator b_l next to b_i; for bosons, $\varepsilon_\alpha = 1$. This can be generalized to cases where several contraction appear, embedded or not.

(ii) In the limit of zero temperature where $\beta \to \infty$, relation (16) simplifies into:

$$\begin{aligned} &\overbrace{b_i b_j} = 0 \text{ if the two operators are of the same nature (both creation,} \\ &\qquad\qquad \text{or both annihilation operators)} \end{aligned} \tag{26}$$

as well as:

$$\overbrace{a_i a_j^\dagger} = \delta_{ij} \times \begin{cases} 1 \text{ si } e_i > \mu \\ 0 \text{ si } e_i < \mu \end{cases} \tag{27}$$

[1]We multiply by -1 if, when writing the permutation, the number of crossings between brackets is odd. This is for instance the case in the permutation in the left of (24), but not that on the right.

and:

$$\overbrace{a_i^\dagger a_j} = \delta_{ij} \times \begin{cases} 0 \text{ si } e_i > \mu \\ 1 \text{ si } e_i < \mu \end{cases} \tag{28}$$

(the second lines ($e_i < \mu$) of these relations are useful only for fermions, since for bosons μ cannot be larger than e_i).

1-d. Statement of the theorem

The recurrence over n we have been using leads to Wick's theorem:

"The average value $\langle b_1 b_2 ... b_i .. b_l ... b_{2n} \rangle$ is the sum of all the complete systems of contractions that can be made on the string of operators $b_1 b_2 ... b_i .. b_l ... b_{2n}$. Each system is the product of binary contractions (16); for fermions, this product is multiplied by parity factors ε_α associated with each of them."

The word "complete" means in this case that in every considered system of contractions, each operator listed in the string of b_i operators is taken in one and only one contraction. The parity factor first includes the parity ε_{α_1} of the permutation that brings right after b_1 the operator it is contracted with; these two operators are then taken out of the list of the b_i. In the remaining list, we again compute the parity ε_{α_2} of the permutation needed to bring together the next two operators to be contracted, and it is multiplied by ε_{α_1}. We continue this until all the contractions have been taken into account, and obtain the product $\varepsilon_{\alpha_1} \varepsilon_{\alpha_2} ... \varepsilon_{\alpha_n}$ of all the parities involved. Among all the system of contractions, a very large number yield zero. The only non-zero ones are those for which every contraction contains a creation and an annihilation operator in the same individual quantum state. This rule significantly limits the number of contractions involved in the final result.

As seen above, the theorem yields again the results of Complement B_{XV}. For example, if (as is the case in the formula for the two-particle symmetric operators) the first two operators b are creation operators, and the last two annihilation operators, the first system of contraction in (23) yields zero, and we are left with the last two, corresponding to the two terms of equation (43) in Complement B_{XV}. The main interest of the theorem is, however, that it allows getting, almost without calculations, the average value of the product of any number of operators.

Comment:

Until now, we assumed that the operators b were creation or annihilation operators associated with the basis of individual states formed by the one-particle Hamiltonian eigenvectors. If this is not the case, and we wish to compute the average value of the product of creation $a_{v_s}^\dagger$ and annihilation a_{v_s} operators associated with any other basis, we first use formulas (A-51) and (A-52) of Chapter XV to express those operators in terms of the operators associated with the eigenbasis of the one-particle Hamiltonian, and then use Wick's theorem.

2. Applications: correlation functions for an ideal gas

As an illustration of the use of Wick's theorem, we now compute the N-order correlation functions in an ideal gas at thermal equilibrium. Thanks to Wick's theorem, they can

each be expressed as simple products of first order correlation functions. As a first step, we will derive, in a simpler way, a number of results already obtained in § 3 of Complement B_{XV}; these will then be generalized to correlation functions of a higher order.

Consider a gas of spinless particles, confined by a one-body potential inside a cubic box of edge length L; this potential is zero inside the box, and becomes infinite outside. We use periodic boundary conditions to account for this confinement (Complément C_{XIV}, § 1-c); the normalized eigenfunctions $u_{\mathbf{k}}(\mathbf{r})$ of the kinetic energy are then written:

$$u_{\mathbf{k}}(\mathbf{r}) = \frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{29}$$

where the possible wave vectors $\mathbf{k}$ are those whose three components are integer multiples of $2\pi/L$.

2-a. First order correlation function

Relation (B-21) of Chapter XVI defines the first order correlation function G_1, which depends on the two positions $\mathbf{r}_1$ and $\mathbf{r}'_1$:

$$G_1(\mathbf{r}_1, \mathbf{r}'_1) = \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi(\mathbf{r}'_1) \right\rangle \tag{30}$$

Using relations (A-3) and (A-6) of Chapter XVI, the field operator can be expressed as a function of the annihilation operators $a_{\mathbf{k}}$ in the state (29), and its adjoint, as a function of the creation operators $a^\dagger_{\mathbf{k}}$ in that same state. Taking into account (29), this leads to:

$$G_1(\mathbf{r}_1, \mathbf{r}'_1) = \frac{1}{L^3} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'_1 - \mathbf{k}\cdot\mathbf{r}_1\right)} \left\langle a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} \right\rangle \tag{31}$$

At thermal equilibrium, all the average values $\langle A \rangle$ of operators A are taken in the state described by the density operator ρ_{eq} written in (1):

$$\langle A \rangle = \mathrm{Tr}\left\{\rho_{eq} A\right\} \tag{32}$$

We can then use Wick's theorem in a particularly simple case, since in (31) the only contraction that comes into play is the one containing $\left[a^\dagger_{\mathbf{k}}, a_{\mathbf{k}'}\right]_{-\eta}$. Relation (18) thus applies and we get:

$$G_1(\mathbf{r}_1, \mathbf{r}'_1) = \frac{1}{L^3} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'_1 - \mathbf{r}_1\right)} \; f_\beta\left(e_{\mathbf{k}} - \mu\right) \tag{33}$$

The correlation function $G_1(\mathbf{r}, \mathbf{r}')$ is therefore directly (to within a constant factor) the Fourier transform of the distribution function $f_\beta(e_{\mathbf{k}} - \mu)$ itself.

The definition of G_1 can be generalized, using the expressions of the field operators in the Heisenberg picture; this leads to a correlation function depending on space and time:

$$G_1(\mathbf{r}_1, t; \mathbf{r}'_1, t') = \left\langle \Psi^\dagger_H(\mathbf{r}_1, t)\Psi_H(\mathbf{r}'_1, t') \right\rangle \tag{34}$$

For free particles (ideal gas), we have (§ 1-c of Complement B_{XV}):

$$\Psi_H(\mathbf{r},t) = \frac{1}{L^{3/2}} \sum_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} a_{\mathbf{k}} \tag{35}$$

where ω_k is the (angular) Bohr frequency associated with the energy of a particle of mass m, with wave vector $\mathbf{k}$:

$$\omega_k = \frac{\hbar k^2}{2m} \tag{36}$$

For an ideal gas, we simply multiply each exponential $e^{i\mathbf{k}\cdot\mathbf{r}}$ by $\mathrm{e}^{-i\omega_k t}$ to go from the Schrödinger to the Heisenberg representation. Expression (33) is then generalized as:

$$G_1(\mathbf{r}_1,t;\mathbf{r}'_1,t') = \frac{1}{L^3} \sum_{\mathbf{k}} e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}_1\right)-\omega_k\left(t'-t\right)\right]}\ f_\beta\left(e_{\mathbf{k}}-\mu\right) \tag{37}$$

Note that this correlation function only depends on the differences in positions (space homogeneity) and times (time translation invariance).

2-b. Second order correlation functions

α. Application of Wick's theorem

The second order correlation function is defined as:

$$G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) = \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi^\dagger(\mathbf{r}_2)\Psi(\mathbf{r}'_2)\Psi(\mathbf{r}'_1)\right\rangle \tag{38}$$

Here again, the average value is computed with the density operator at thermal equilibrium. The same calculation as in § 2-a leads to:

$$G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) = \frac{1}{L^6} \sum_{\mathbf{k}}\sum_{\mathbf{k}'}\sum_{\mathbf{k}''}\sum_{\mathbf{k}'''} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'_1+\mathbf{k}'''\cdot\mathbf{r}'_2-\mathbf{k}\cdot\mathbf{r}_1-\mathbf{k}''\cdot\mathbf{r}_2\right)} \left\langle a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{k}''} a_{\mathbf{k}'''} a_{\mathbf{k}'}\right\rangle \tag{39}$$

As we already saw in (23), using Wick's theorem yields two contraction systems, one where $a^\dagger_{\mathbf{k}}$ is contracted with $a_{\mathbf{k}'}$ (and hence $a^\dagger_{\mathbf{k}''}$ with $a_{\mathbf{k}'''}$), and another one where $a^\dagger_{\mathbf{k}}$ is contracted with $a_{\mathbf{k}'''}$ (and hence $a^\dagger_{\mathbf{k}''}$ with $a_{\mathbf{k}'}$):

$$\overbrace{a^\dagger_{\mathbf{k}}\underbrace{a^\dagger_{\mathbf{k}''}a_{\mathbf{k}'''}}a_{\mathbf{k}'}} \qquad \text{et} \qquad \overbrace{a^\dagger_{\mathbf{k}}\underbrace{a^\dagger_{\mathbf{k}''}a_{\mathbf{k}'''}a_{\mathbf{k}'}}}$$

The second contraction involves an odd permutation, and introduces a factor η. We therefore obtain (after changing the dummy index $\mathbf{k}''$ into $\mathbf{k}'$):

$$\begin{aligned} &G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) \\ &\quad= \frac{1}{L^6}\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\left[e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}_1\right)+\mathbf{k}'\cdot\left(\mathbf{r}'_2-\mathbf{r}_2\right)\right]} + \eta e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_2-\mathbf{r}_1\right)+\mathbf{k}'\cdot\left(\mathbf{r}'_1-\mathbf{r}_2\right)\right]}\right] \\ &\qquad\qquad \times f_\beta\left(e_{\mathbf{k}}-\mu\right) f_\beta\left(e_{\mathbf{k}'}-\mu\right) \end{aligned} \tag{40}$$

that is, taking (33) into account:

$$G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) = G_1(\mathbf{r}_1,\mathbf{r}'_1)\times G_1(\mathbf{r}_2,\mathbf{r}'_2) + \eta\ G_1(\mathbf{r}_1,\mathbf{r}'_2)\times G_1(\mathbf{r}_2,\mathbf{r}'_1) \tag{41}$$

This means that the second order correlation function is simply expressed as the sum of two products of first order correlation functions. The first is the direct term, and corresponds to totally uncorrelated particles. The second is the exchange term, a consequence of the quantum indistinguishability of the particles; it has an opposite sign for fermions and bosons. As in Complement A_{XVI}, we will show that this term introduces correlations between the particles.

β. *Double density*

Of particular interest is the "diagonal" case where $\mathbf{r}_1 = \mathbf{r}'_1$ and $\mathbf{r}_2 = \mathbf{r}'_2$, as the function $G_2(\mathbf{r}_1, \mathbf{r}'_1; \mathbf{r}_2, \mathbf{r}'_2)$ becomes very simple to interpret: it is the "double density" $G_2(\mathbf{r}_1, \mathbf{r}_2)$ characterizing the probability of finding a particle at point $\mathbf{r}_1$ and another one at point $\mathbf{r}_2$. The above relation takes on the simplified form:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = G_1(\mathbf{r}_1, \mathbf{r}_1) \times G_1(\mathbf{r}_2, \mathbf{r}_2) + \eta\ G_1(\mathbf{r}_1, \mathbf{r}_2) \times G_1(\mathbf{r}_2, \mathbf{r}_1) \tag{42}$$

If, in addition, $\mathbf{r}_1 = \mathbf{r}_2$, this function indicates the probability of finding two particles at the same point. We then get:

$$\begin{cases} \text{for fermions} & G_2(\mathbf{r}; \mathbf{r}) = 0 \\ \text{for bosons} & G_2(\mathbf{r}; \mathbf{r}) = 2\ \left[G_1(\mathbf{r}, \mathbf{r})\right]^2 \end{cases} \tag{43}$$

For fermions, we see as expected that one can never find two of them at the same point in space, a consequence of Pauli exclusion principle. For bosons, we find that the double density is twice the square of the one-body density. Now, if both particles were uncorrelated, this double density should simply be equal to the square, without the factor two. This factor two thus indicates an increase in the probability of finding two bosons at the same point in space; it expresses the bunching tendency of identical bosons, a tendency that comes from a pure quantum statistical effect since we assumed the particle's interactions to be zero. These results were already discussed in Complement A_{XVI} – see in particular Figure 3.

The Hartree-Fock method (mean field approximation), presented in Complements E_{XV} and F_{XV}, uses a variational ket (or a density operator) such that the binary correlation function $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is given by the sum of products of functions G_1, written in (42), even in the presence of interactions. Moreover, another way of introducing the Hartree-Fock approximation is to assume directly that the binary correlation function keeps this form even in the presence of interactions, which then allows a simple calculation of the interaction energy. Even though this method has numerous applications, and may be quite precise in certain cases, it does rely on an approximation: when the particles interact with each other, there is no general reason for G_2 to remain linked to G_1 by this relation, obtained with the assumption that the gas was ideal.

γ. *Time-dependent correlation function*

As we did for the first order correlation function, we can include time dependence in the field operators, and define:

$$G_2(\mathbf{r}_1, t_1; \mathbf{r}'_1, t'_1; \mathbf{r}_2, t_2; \mathbf{r}'_2, t'_2) = \left\langle \Psi_H^\dagger(\mathbf{r}_1, t_1) \Psi_H^\dagger(\mathbf{r}_2, t_2) \Psi_H(\mathbf{r}'_2, t'_2) \Psi_H(\mathbf{r}'_1, t'_1) \right\rangle \tag{44}$$

To include the time dependence, we simply add, as above, in each spatial exponential with wave vector $\mathbf{k}$, a time exponential with the corresponding angular frequency $-\omega_k$, which leads to:

$$\begin{aligned} G_2(\mathbf{r}_1,t_1;\mathbf{r}'_1,t'_1;\mathbf{r}_2,t_2;\mathbf{r}'_2,t'_2) = \\ \frac{1}{L^6}\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\Big\{ \mathrm{e}^{\mathrm{i}[\mathbf{k}\cdot(\mathbf{r}'_1-\mathbf{r}_1)-\omega_k(t'_1-t_1)+\mathbf{k}'\cdot(\mathbf{r}'_2-\mathbf{r}_2)-\omega_{k'}(t'_2-t_2)]} \\ + \eta\, \mathrm{e}^{\mathrm{i}[\mathbf{k}\cdot(\mathbf{r}'_2-\mathbf{r}_1)-\omega_k(t'_2-t_1)+\mathbf{k}'\cdot(\mathbf{r}'_1-\mathbf{r}_2)-\omega_{k'}(t'_1-t_2)]}\Big\} f_\beta(e_{\mathbf{k}}-\mu)\, f_\beta(e_{\mathbf{k}'}-\mu) \\ = G_1(\mathbf{r}_1,t_1;\mathbf{r}'_1,t'_1)\, G_1(\mathbf{r}_2,t_2;\mathbf{r}'_2,t'_2) + \eta\, G_1(\mathbf{r}_1,t_1;\mathbf{r}'_2,t'_2)\, G_1(\mathbf{r}_2,t_2;\mathbf{r}'_1,t'_1) \end{aligned} \tag{45}$$

Hence, when time dependence is included, we get a factored relation similar to (41). As before, because of the space homogeneity and the time translation invariance, only the differences in the space and time variables appear in the correlation function expression.

2-c. Higher order correlation functions

In a more general way[2], the N-order correlation function G_N is defined by:

$$G_N(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2;...;\mathbf{r}_N,\mathbf{r}'_N) = \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi^\dagger(\mathbf{r}_2)...\Psi^\dagger(\mathbf{r}_N)\Psi(\mathbf{r}'_N)...\Psi(\mathbf{r}'_2)\Psi(\mathbf{r}'_1)\right\rangle \tag{46}$$

These functions give information on the correlated behavior of groups of N particles in an ideal gas at equilibrium. Using Wick's theorem, each of them can be expressed in terms of the first order correlation function $G_1(\mathbf{r}_1,\mathbf{r}'_1)$. As an example, let us study the correlation function of order three:

$$\begin{aligned} G_3(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2;\mathbf{r}_3,\mathbf{r}'_3) &= \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi^\dagger(\mathbf{r}_2)\Psi^\dagger(\mathbf{r}_3)\Psi(\mathbf{r}'_3)\Psi(\mathbf{r}'_2)\Psi(\mathbf{r}'_1)\right\rangle \\ &= \frac{1}{L^9}\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\sum_{\mathbf{k}''}\sum_{\mathbf{k}'''}\sum_{\mathbf{k}''''}\sum_{\mathbf{k}'''''} \\ &\qquad e^{i\left(\mathbf{k}'\cdot\mathbf{r}'_1+\mathbf{k}'''\cdot\mathbf{r}'_2+\mathbf{k}'''''\cdot\mathbf{r}'_3-\mathbf{k}\cdot\mathbf{r}_1-\mathbf{k}''\cdot\mathbf{r}_2-\mathbf{k}''''\cdot\mathbf{r}_3\right)} \\ &\qquad \times \left\langle a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{k}''} a^\dagger_{\mathbf{k}''''} a_{\mathbf{k}'''''} a_{\mathbf{k}'''} a_{\mathbf{k}'}\right\rangle \end{aligned} \tag{47}$$

Six contraction systems must be considered to compute the average values. In the first system, $\mathbf{k}$ and $\mathbf{k}'$ are associated, $\mathbf{k}''$ with $\mathbf{k}'''$, and finally $\mathbf{k}''''$ with $\mathbf{k}'''''$. One can then permute the three vectors $\mathbf{k}'$, $\mathbf{k}'''$ and $\mathbf{k}'''''$ in 5 different ways, with odd or even permutations. In each of the terms thus obtained, the sixfold summation on the wave vectors is reduced to a triple sum, which yields a product of functions G_1. This leads to:

$$\begin{aligned} G_3(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2;\mathbf{r}_3,\mathbf{r}'_3) &= G_1(\mathbf{r}_1,\mathbf{r}'_1)\, G_1(\mathbf{r}_2,\mathbf{r}'_2)\, G_1(\mathbf{r}_3,\mathbf{r}'_3) \\ &+G_1(\mathbf{r}_1,\mathbf{r}'_2)\, G_1(\mathbf{r}_2,\mathbf{r}'_3)\, G_1(\mathbf{r}_3,\mathbf{r}'_1) + G_1(\mathbf{r}_1,\mathbf{r}'_3)\, G_1(\mathbf{r}_2,\mathbf{r}'_1)\, G_1(\mathbf{r}_3,\mathbf{r}'_2) \\ &+\eta\, G_1(\mathbf{r}_1,\mathbf{r}'_1)\, G_1(\mathbf{r}_2,\mathbf{r}'_3)\, G_1(\mathbf{r}_3,\mathbf{r}'_2) + \eta\, G_1(\mathbf{r}_1,\mathbf{r}'_3)\, G_1(\mathbf{r}_2,\mathbf{r}'_2)\, G_1(\mathbf{r}_3,\mathbf{r}'_1) \\ &+\eta\, G_1(\mathbf{r}_1,\mathbf{r}'_2)\, G_1(\mathbf{r}_2,\mathbf{r}'_1)\, G_1(\mathbf{r}_3,\mathbf{r}'_3) \end{aligned} \tag{48}$$

The computation can be generalized in the same way to correlation functions of any order; in an ideal gas, they are not independent since they are all simple products

[2]We only consider here the so-called "normal"correlation functions, those where the $\Psi^\dagger$ come before the Ψ. In Complement B_{XVI} we introduce more general correlation functions.

of first order correlation functions. In other words, the function G_1 contains all the information necessary for computing correlations of any order.

Finally, we can compute the triple density $G_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ by setting $\mathbf{r}_1 = \mathbf{r}'_1$, $\mathbf{r}_2 = \mathbf{r}'_2$ and $\mathbf{r}_3 = \mathbf{r}'_3$ in (48). The particular case $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}_3$ where all the positions are identical is interesting. For fermions, the triple density is zero, for the same reason as with the double density: Pauli principle does not allow several fermions to occupy the same point in space. For bosons, we find:

$$G_3(\mathbf{r}, \mathbf{r}; \mathbf{r}) = 6 \left[G_1 (\mathbf{r}, \mathbf{r})\right]^3 \tag{49}$$

For three identical bosons, the bunching tendency under the effect of their quantum statistics is even higher than for two bosons, introducing a factor 6 instead of 2.

Comment:

The results we have obtained are valid when the system's density operator is that of an ideal gas at thermal equilibrium as in relation (1), but they could be quite different if the system is in another state. If, for example, we assume (as in Complement $\mathrm{A_{XVI}}$) that the system is described by a Fock state, the relations between correlation functions can be totally different. The simplest case is that of an ideal gas of bosons in its ground state, where all the bosons occupy the same individual state; relations (24) and (25) of that complement indicate that:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{N-1}{N} G_1(\mathbf{r}_1, \mathbf{r}_1) \times G_1(\mathbf{r}_2, \mathbf{r}_2) \simeq G_1(\mathbf{r}_1, \mathbf{r}_1) \times G_1(\mathbf{r}_2, \mathbf{r}_2) \tag{50}$$

G_2 is thus simply the product of two functions G_1, without the exchange term of (42); consequently, the factor 2 in the second line of (43) no longer exists. In a similar way, one can show that the factor 6 of relation (49) is no longer present. In a general way, for an ensemble of bosons all in the same individual state, the bunching effects related to the indistinguishability of the particles are not present.

Following this line of thought, note that it is not possible to get the projector onto a Fock state (other than the vacuum) such as the one discussed above, by using the density operator (1) at thermal equilibrium, and taking its limit as the temperature T goes to zero, i.e. as $\beta \to \infty$. This is because this density operator associates with each individual state an occupation number distribution that is always a decreasing exponential, and never a narrow curve centered around a high value of the particle number. Consequently, there exist large fluctuations of the particle number in each mode, and hence the presence of the factors 2 in (43) and 6 in (49), whatever the value of β.

To conclude, let us mention that Wick's theorem can take on diverse forms, in particular at zero or non-zero temperadure (see for instance Chapter 4 of Reference [5]). As we saw, thanks to this theorem, and when dealing with independent particles, the computation of correlation functions of any order, time-dependent or not, can be reduced to computing the product of first order correlation functions. It is obviously a great simplification. This property is reminiscent of random Gaussian variables in classical statistics: for such variables, all moments of any order can be expressed in terms of products of the lowest order moment. These properties are characteristic of an ideal gas: in a system where particles interact, the correlation functions of successive orders remain independent in general. Nevertheless, the use of Wick's theorem is not limited to ideal gases; its range of application is much more general, and it is very useful in perturbation calculations where power series of the interaction potential are derived [5].

Chapter XVII

Paired states of identical particles

A Creation and annihilation operators of a pair of particles . 1813
A-1 Spinless particles, or particles in the same spin state 1813
A-2 Particles in different spin states 1816
B Building paired states 1818
B-1 Well determined particle number 1818
B-2 Undetermined particle number 1820
B-3 Pairs of particles and pairs of individual states 1822
C Properties of the kets characterizing the paired states . . . 1822
C-1 Normalization 1822
C-2 Average value and root mean square deviation of particle number 1825
C-3 "Anomalous" average values 1828
D Correlations between particles, pair wave function 1830
D-1 Particles in the same spin state 1831
D-2 Fermions in a singlet state 1834
E Paired states as a quasi-particle vacuum; Bogolubov-Valatin transformations 1836
E-1 Transformation of the creation and annihilation operators . . 1836
E-2 Effect on the kets $|\overline{\varphi}_{\mathbf{k}}\rangle$ 1838
E-3 Basis of excited states, quasi-particles 1840

Introduction

Fock states were introduced in Chapter XV by the action on the vacuum of a product of individual state creation operators. A certain number of their properties were studied in § C-5-b-β of Chapter XV and in Complement $\mathrm{A_{XVI}}$ (exchange hole for fermions, bunching

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

effect for bosons). We also used Fock states in Complements C_{XV}, D_{XV} and F_{XV} as variational kets to account, approximately, for the interactions between the particles. This led us, both for fermions and bosons, to a mean field theory where each particle can be seen as propagating in the mean field created by all the others.

We now introduce a larger class of variational states to improve the accuracy of these results, allowing us to study many more properties of physical systems of identical particles. It concerns the "paired states" obtained by the action on the vacuum of a product of creation operators, no longer of individual particles but rather of pairs of particles (if these particles form a molecule, we are dealing with molecule creation operators). As we shall see in the course of this chapter, these paired states are more general than the Fock states, since they can be reduced to Fock states for certain values of the parameters characteristic of the pair[1]. What is sought is an improvement of the variational method allowing us to ameliorate our treatment of the interactions, compared to that based on Fock states.

The additional flexibility introduced by the paired states plays an essential role for the following simple reason: changing the properties of the pair wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ used to build them, we modify the binary correlation function of the N-particle system. We therefore take advantage of the power of the mean field method, while retaining the possibility of taking into account any binary correlations. Whereas using variational Fock states allows taking into account only statistical correlations (due to particle indistinguishability), the paired states enable us to add dynamic correlations (due to interactions). These latter correlations are essential: when dealing with binary interactions (as is the case with a standard Hamiltonian), these correlations actually determine the average value of the potential energy (Chapter XV, § C-5-b). Three-body, four-body, etc.. correlations are indeed present in the system, playing their role; they are not, however, directly involved in the energy. This explains why the optimization with paired states of only the binary correlations can lead to fairly good results in the study of N-body systems. These possibilities have a wide range of applications for both fermions and bosons, which will be discussed in the complements.

This chapter is centered on the study of the general properties of paired states, and introduces the tools for handling such states. We study, in parallel, fermions and bosons to highlight the numerous analogies between results obtained for both cases. We first introduce (§ A) the creation and annihilation operators for pairs of particles. We then build (§ B) the paired states and discuss some of their properties; this permits introducing (§ C) the concept of "normal average values" (average values of operators conserving the particle number) or "anomalous average values" (average values of operators changing the particle number). We then show in § D how the paired states allow us to actually vary the spatial correlation functions of a system of identical particles. This will lead us to introduce a function playing an important role in what follows (in particular in the complements of this chapter), the pair wave function ϕ_{pair}, which is related to the anomalous average values. We then study in § E another interesting property of the paired states: they can be related to the concept of "quasi-particle" thanks to the introduction of new creation and annihilation operators resulting from a linear transformation of the initial operators (Bogolubov transformation). As the paired states are eigenkets of the new annihilation operators with a zero eigenvalue, they behave as

[1]For example, we shall clarify at the end of § C-1-a why the Hartree-Fock method can be viewed as a particular case of the pairing method.

a "quasi-particle vacuum". Furthermore, the creation operators can associate with each paired state an entire basis of other orthogonal states, which are interpreted as states occupied by quasi-particles.

This study of the necessary tools for handling the paired states will be continued in the first two complements, $\mathrm{A_{XVII}}$ and $\mathrm{B_{XVII}}$. Complement$\mathrm{A_{XVII}}$ discusses a complementary aspect of pairing, the introduction of the pair field operators. These operators have a non-zero average value in paired states, and highlight the cooperative effects existing in those states. This can lead to the spontaneous appearance in the system of an order parameter, described by the same pair wave function ϕ_{pair} as the one appearing in the computations of correlation functions in a paired state. In addition, Complement $\mathrm{A_{XVII}}$ will show that the commutation properties of these operators are reminiscent of those of a boson field: in a certain sense, a composite object built from two identical particles (whether they are bosons or fermions) behaves as a boson. It is, however, only an approximation, as can be inferred from the corrective terms appearing in the computation of the commutators, which can sometimes play an important role. Complement $\mathrm{B_{XVII}}$ discusses the computation of the energy average value in a paired state, whose expression is the basis of the following complements; it gives an example of how to deal with normal and anomalous average values in these computations.

The last three complements apply these results to the variational study of interacting boson or fermion systems. For fermions, the paired states play an essential role in the BCS (Bardeen-Cooper-Schrieffer, theory of supraconductivity) theory of supraconductivity (Complement $\mathrm{C_{XVII}}$), and explain the appearance of a pair field as a collective effect; paired states also come into play noticeably in nuclear physics, and in the study of ultra-cold fermionic atomic gases. For repulsive bosons (Complement $\mathrm{E_{XVII}}$), paired states can be quite useful for studying the ground state properties, and to obtain, for example, the Bogolubov linear spectrum. In that case, the paired state is associated with another state (a coherent state, for example), whose role is to describe the condensate as an accumulation of a notable fraction of particles in a single individual quantum state.

A. Creation and annihilation operators of a pair of particles

Let us introduce the creation or annihilation operators, no longer of a single particle, but of two identical particles in a bound state. We first assume the particles have no spin (or are both in the same spin state, so that no spin variable is needed).

A-1. Spinless particles, or particles in the same spin state

Consider two identical particles (bosons or fermions in the same spin state), with positions $\mathbf{r}_1$ and $\mathbf{r}_2$; the system is contained in a cubic box of edge length L and volume $\mathcal{V} = L^3$. These two particles occupy a bound state, characterized by a normalized wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$, forming a kind of binary "molecule". The state of the system is defined by this wave function (as far as its internal orbital variables are concerned), by spin variables identical for both particles (since those spin variables are of no importance here, they need not be written explicitly in what follows), and finally by its external orbital variables (center of mass). The normalized wave function of a "molecule" having

a total momentum $\hbar\mathbf{K}$ is then:

$$\begin{aligned}\phi_{\mathbf{K}}(\mathbf{r}_1,\mathbf{r}_2) &= (L)^{-3/2}\, e^{i\mathbf{K}\cdot(\mathbf{r}_1+\mathbf{r}_2)/2}\chi(\mathbf{r}_1-\mathbf{r}_2)\\ &= (L)^{-3}\sum_{\mathbf{k}} g_{\mathbf{k}}\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2}\end{aligned} \tag{A-1}$$

where $g_{\mathbf{k}}$ is the Fourier transform of χ:

$$\begin{aligned} g_{\mathbf{k}} &= \frac{1}{L^{3/2}}\int_{L^3} d^3r\; e^{-i\mathbf{k}\cdot\mathbf{r}}\chi(\mathbf{r})\\ \chi(\mathbf{r}) &= \frac{1}{L^{3/2}}\sum_{\mathbf{k}}\; g_{\mathbf{k}}\; e^{i\mathbf{k}\cdot\mathbf{r}}\end{aligned} \tag{A-2}$$

We assume that the individual wave functions of the particles obey the periodic boundary conditions (Complement C_{XIV}, § 1-c); in (A-1), each component of the wave vector of particle 1 or 2 can therefore take only the values $2\pi n_x/L$, $2\pi n_y/L$ and $2\pi n_z/L$, where n_x, n_y and n_z are any integer number (positive, negative, or zero). The normalization of the functions χ and g is written:

$$\int_{L^3} \mathrm{d}^3r\;\; |\chi(\mathbf{r})|^2 = \sum_{\mathbf{k}}\;\; |g_{\mathbf{k}}|^2 = 1 \tag{A-3}$$

Moreover, for identical particles, the symmetrization (or antisymmetrization) requires the function $\chi(\mathbf{r})$ and its Fourier transform $g(\mathbf{k})$ to have the parity η:

$$g_{-\mathbf{k}} = \eta\; g_{\mathbf{k}} \tag{A-4}$$

($\eta = +1$ for bosons, $\eta = -1$ for fermions).

In terms of kets, relation (A-1) becomes:

$$\begin{aligned}|\phi_{\mathbf{K}}(1,2)\rangle &= (L)^{-3}\sum_{\mathbf{k}}\int \mathrm{d}^3r_1\int \mathrm{d}^3r_2\; g_{\mathbf{k}}\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2}\,|1:\mathbf{r}_1;2:\mathbf{r}_2\rangle\\ &= \sum_{\mathbf{k}}\; g_{\mathbf{k}}\;\left|1:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right);2:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\right\rangle\end{aligned} \tag{A-5}$$

which, taking (A-4) into account, and changing the sign of the sum variable $\mathbf{k}$, can also be written as:

$$\begin{aligned}|\phi_{\mathbf{K}}(1,2)\rangle = \frac{1}{2}\sum_{\mathbf{k}}\; g_{\mathbf{k}}\;\Bigg[&\left|1:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right);2:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\right\rangle\\ &+\eta\left|1:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right);2:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\right\rangle\Bigg]\end{aligned} \tag{A-6}$$

The expression between brackets in the summation is simply the (anti)symmetrized ket of two particles, the first one of momentum $\hbar(\mathbf{k}+\mathbf{K}/2)$, and the other one of momentum $\hbar(-\mathbf{k}+\mathbf{K}/2)$. Two cases must be distinguished:

(i) If $\mathbf{k}\neq 0$, to normalize the ket between brackets, we divide it by $\sqrt{2}$; we then get a Fock state where two individual states with different momenta are occupied (see

the general definition of the Fock states in Chapter XV). The ket between brackets is thus equal to:

$$\sqrt{2}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}}\ |0\rangle \tag{A-7}$$

(ii) If $\mathbf{k} = 0$ and in the case of bosons, the ket between brackets is equal to twice the Fock state where a single individual level is occupied by two particles; this ket is equal to:

$$\sqrt{2}\left(a^\dagger_{\frac{\mathbf{K}}{2}}\right)^2 |0\rangle \tag{A-8}$$

For fermions, the ket between brackets must be zero, which is indeed the case of the ket in (A-8). To sum up, whether we are dealing with fermions or bosons, and whether $\mathbf{k}$ is zero or not, the ket between brackets can always be expressed as (A-7). This leads to:

$$|\phi_{\mathbf{K}}\rangle = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}}\ g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}}\ |0\rangle \tag{A-9}$$

If the particles are all in the same spin state, remember that in this expression the spin index is implicit: each creation operator is associated with an individual state whose momentum is specified by the operator index, and whose spin state is the common spin state of all the particles.

The creation operator $A^\dagger_{\mathbf{K}}$ of a "molecule" having a total momentum $\hbar\mathbf{K}$ can therefore be written as:

$$A^\dagger_{\mathbf{K}} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}}\ g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{A-10}$$

Its action is to create two particles of momenta $\hbar\left[(\mathbf{K}/2) \pm \mathbf{k}\right]$, with amplitudes given by the function $g_{\mathbf{k}}$. As this function has the parity η, we note that:

$$g_{-\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} = \eta g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} = g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{A-11}$$

Accordingly, the contributions of opposite values of $\mathbf{k}$ double each other in (A-10). Such a redundancy will cause a problem in § B-2, when we write a tensor product. It is thus preferable to eliminate it right now and this is why we restrict the summation over $\mathbf{k}$ to half the wave vector space. Calling D this half space, we shall write $A^\dagger_{\mathbf{K}}$ in the form:

$$A^\dagger_{\mathbf{K}} = \sqrt{2} \sum_{\mathbf{k}\in D}\ g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{A-12}$$

For a "molecule" having a zero total momentum, this relation becomes:

$$A^\dagger_{\mathbf{K}=\mathbf{0}} = \sqrt{2} \sum_{\mathbf{k}\in D}\ g_{\mathbf{k}}\ a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} \tag{A-13}$$

As for the annihilation operator of a molecule with total momentum $\hbar\mathbf{K}$, it is simply the adjoint of (A-12):

$$A_{\mathbf{K}} = \sqrt{2} \sum_{\mathbf{k}\in D}\ g^*_{\mathbf{k}}\ a_{\frac{\mathbf{K}}{2}-\mathbf{k}} a_{\frac{\mathbf{K}}{2}+\mathbf{k}} \tag{A-14}$$

We have reasoned in terms of "molecules" being created or annihilated, but the wave function $\chi(\mathbf{r})$ and its Fourier transform $g_{\mathbf{k}}$ are not related to any particular bound state, and do not imply the existence of any attraction potential between the two components of this "molecule". Actually, in what follows, the $g_{\mathbf{k}}$ will play the role of freely adjustable parameters, for example when using a variational method. To illustrate this generality, we shall, from now on, talk about "pairs".

Comment:

If we choose for $g_{\mathbf{k}}$ as in a Kronecker delta, with the symmetrization required by (A-4):

$$g_{\mathbf{k}} = \frac{1}{\sqrt{2}} \left[\delta_{\mathbf{k},\mathbf{k}_0} + \eta \delta_{\mathbf{k},-\mathbf{k}_0}\right] \tag{A-15}$$

we get, according to (A-12):

$$A^\dagger_{\mathbf{K}} = \frac{1}{2}\left[a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}_0} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}_0} + \eta\, a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}_0} a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}_0}\right] = a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}_0} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}_0} \tag{A-16}$$

In the right-hand side of this relation, the momenta appearing as indices of the creation operators can take any given values, obtained by varying $\mathbf{K}$ and $\mathbf{k}_0$. It is therefore possible, by a suitable choice of the pair's parameters, to create two particles in individual states having any given momenta, and thereby obtain a Fock state. Successive applications of P operators $A^\dagger_{\mathbf{K}}$ (having, in general, different values of $\mathbf{K}$ and $\mathbf{k}_0$) can thus yield a Fock state with $2P$ particles whose momenta can take on any values.

A-2. Particles in different spin states

We assume the internal state of the pair is a tensor product of an orbital state depending on $\mathbf{r}_1 - \mathbf{r}_2$ and a spin state $|\chi_S\rangle$. Equation (A-1) must then be replaced by:

$$\begin{aligned}
\phi^{\nu_1,\nu_2}_{\mathbf{K}}(\mathbf{r}_1,\mathbf{r}_2) &= \langle 1:\mathbf{r}_1,\nu_1;\ 2:\mathbf{r}_2,\nu_2\ |\Phi_{\mathbf{K}}\rangle \\
&= (L)^{-3/2}\, e^{i\mathbf{K}\cdot(\mathbf{r}_1+\mathbf{r}_2)/2} \chi(\mathbf{r}_1-\mathbf{r}_2)\, \langle \nu_1,\nu_2\ |\chi_S\rangle \\
&= (L)^{-3} \sum_{\mathbf{k}} g_{\mathbf{k}}\, e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2}\, \langle \nu_1,\nu_2\ |\chi_S\rangle
\end{aligned} \tag{A-17}$$

This means that relation (A-1) is to be multiplied by $\langle \nu_1,\nu_2\ |\chi_S\rangle$; relation (A-5) is now written:

$$\begin{aligned}
|\phi_{\mathbf{K}}(1,2)\rangle &= (L)^{-3} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \sum_{\mathbf{k}} g_{\mathbf{k}}\, e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2} \\
&\qquad \times \sum_{\nu_1,\nu_2} \langle \nu_1,\nu_2\ |\chi_S\rangle\ \ |1:\mathbf{r}_1,\nu_1;\ 2:\mathbf{r}_2,\nu_2\rangle \\
&= \sum_{\mathbf{k}} g_{\mathbf{k}} \sum_{\nu_1,\nu_2} \langle \nu_1,\nu_2\ |\chi_S\rangle \left|1:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right),\nu_1;\ 2:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right),\nu_2\right\rangle
\end{aligned} \tag{A-18}$$

The function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ is supposed to have an orbital parity equal to η_o, and the spin ket $|\chi_S\rangle$, a parity with respect to the exchange of spins equal to η_s, with, obviously:

$$\eta_o \ \eta_s \ = \eta \tag{A-19}$$

Hence:

$$|\phi_{\mathbf{K}}(1,2)\rangle = \frac{1}{2}\sum_{\mathbf{k}} \ g_{\mathbf{k}} \ \sum_{\nu_1,\nu_2} \langle \nu_1, \nu_2 \ |\chi_S\rangle \left\{ \left| 1 : \frac{\mathbf{K}}{2} + \mathbf{k}, \nu_1; 2 : \frac{\mathbf{K}}{2} - \mathbf{k}, \nu_2 \right\rangle \right.$$
$$\left. + \eta \left| 1 : \frac{\mathbf{K}}{2} - \mathbf{k}, \nu_2; 2 : \frac{\mathbf{K}}{2} + \mathbf{k}, \nu_1 \right\rangle \right\} \tag{A-20}$$

which shows that the creation operator of a pair is:

$$A^{\dagger}_{\mathbf{K}} = \frac{1}{\sqrt{2}}\sum_{\mathbf{k}} \ g_{\mathbf{k}} \ \sum_{\nu_1,\nu_2} \langle \nu_1, \nu_2 \ |\chi_S\rangle \ \ a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\nu_1} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\nu_2} \tag{A-21}$$

As an example, for two fermions of spin 1/2 in a singlet state:

$$A^{\dagger}_{\mathbf{K}} = \frac{1}{2}\sum_{\mathbf{k}} \ g_{\mathbf{k}} \ \left[a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\nu=+} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\nu=-} - a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\nu=-} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\nu=+} \right] \tag{A-22}$$

Since $\eta_s = -1$, the functions $\chi(\mathbf{r})$ and $g_{\mathbf{k}}$ are even. Using this parity, we can exchange the dummy indices $\mathbf{k}$ and $-\mathbf{k}$ in the second term on the right-hand side, and change the order of the two creation operators, with a sign change (anticommutation of fermionic operators). This second term then doubles the first one, and we get:

$$A^{\dagger}_{\mathbf{K}} = \sum_{\mathbf{k}} \ g_{\mathbf{k}} \ a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} \tag{A-23}$$

with the simplified notation we shall use from now on:

$$\begin{array}{lll} a_{\mathbf{k},\nu=+} & \text{noted:} & a_{\mathbf{k},\uparrow} \\ a_{\mathbf{k},\nu=-} & \text{noted:} & a_{\mathbf{k},\downarrow} \end{array} \tag{A-24}$$

(and, of course, a similar notation for the creation operators $a^{\dagger}$). Note in passing that, because of the presence of spins, no redundancy is present in the summation appearing in (A-23), and there is no need to restrict it to a half-space.

Comments:

(i) Taking for $g_{\mathbf{k}}$ a (symmetrized) delta function, as in (A-15), it is possible, as we pointed out before, to construct any Fock state with arbitrary momenta by successive application of operators $A^{\dagger}_{\mathbf{K}}$ on the vacuum; note, however, that the total occupation numbers of the two spin states must remain equal.

(ii) Choosing in (A-22) a function $g_{\mathbf{k}}$ that is even instead of odd for fermions, the operator written in (A-23) creates a fermion pair with a total spin state $S = 1$, and a $M = 0$ component. This is because replacing the minus sign by a plus sign in the middle of the bracket of relation (A-22) yields a triplet spin state; using the fact that $g_{\mathbf{k}}$ is now an odd function, the same reasoning leads to (A-23).

B. Building paired states

To avoid complex formulas, we will build the simplest possible paired states $|\Psi_P\rangle$. We shall be guided by the Gross-Pitaevskii variational method (Complement C_{XV}), where we assumed that the state of the N-particle system could be obtained from the vacuum by creating N particles in the same individual state. However, instead of applying many times the creation operator $a^\dagger_{\mathbf{k}}$ of a single particle to the particle vacuum, we shall now use the pair creation operator $A^\dagger_{\mathbf{K}}$. This difference is essential, in particular for fermions. As we know, it is impossible to create several fermions in the same individual quantum state, since the square, cube, etc. of any creation operator acting on a given individual state yields zero. We shall see, however, that the creation of P pairs of fermions, all in the same quantum state, does not lead to a zero state vector.

B-1. Well determined particle number

We define the paired state $|\Psi_P(\mathbf{K})\rangle$ as the (non-normalized) state vector where $N = 2P$ particles form P pairs, each having a total momentum $\hbar\mathbf{K}$:

$$|\Psi_P(\mathbf{K})\rangle = \left[A^\dagger_{\mathbf{K}}\right]^P |0\rangle \tag{B-1}$$

where $A^\dagger_{\mathbf{K}}$ has been defined in (A-12) or (A-21), depending on the case. To keep things simple, we assume in what follows that all the created pairs have zero total momentum; if this is not the case, we can change the reference frame and choose the one where the common value of the total momentum of all the pairs of particles is zero. The paired state $|\Psi_P\rangle$ is then written:

$$|\Psi_P\rangle = \left[A^\dagger_{\mathbf{K}=0}\right]^P |0\rangle \tag{B-2}$$

We shall first study (as in § A-1) the case of bosons or fermions in the same spin state. As for the case of particles in several spin states (as in § A-2), we shall, from now on, limit our study to fermions in a singlet state; this will allows exposing the general principle while avoiding more complex calculations. In both cases, the $2P$-particle state only depends on the values of the parameters $g_{\mathbf{k}}$. As soon as $P > 1$, we will see that the normalization of the ket $|\Psi_P(\mathbf{K})\rangle$ does not reduce to the simple condition (A-3), which required the sum of the $|g_{\mathbf{k}}|^2$ to be equal to unity. This is why we shall consider from now on that the $g_{\mathbf{k}}$ are totally free variational parameters. For example, multiplying them all by the same constant, one can choose to vary at will the norm of $|\Psi_P(\mathbf{K})\rangle$. This will offer a flexibility simplifying the computations.

B-1-a. Particles in the same spin state

For particles in the same spin state, we can use (A-13), which leads to:

$$|\Psi_P\rangle = \left\{\sum_{\mathbf{k}\in D} \sqrt{2} g_{\mathbf{k}}\, a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}}\right\}^P |0\rangle \tag{B-3}$$

where D is the summation domain defined previously (half of the $\mathbf{k}$ space); remember that the physical system is assumed to be contained in a cubic box of side length L and

volume $\mathcal{V} = L^3$; the periodic boundary conditions then fixes all the possible values for the summation over $\mathbf{k}$. Note also that the spin index is implicit: $a^\dagger_{\mathbf{k}}$ is the creation operator in the individual state defined by the momentum $\hbar\mathbf{k}$ and the unique spin state we are concerned with.

Initially, the parameters $g_{\mathbf{k}}$ were introduced as the Fourier components of the normalized pair wave function $\chi(\mathbf{r})$; the sum of their moduli squared was fixed to unity. This condition, however, does not ensure the normalization of $|\Psi_P\rangle$, as we now show. Because of the power P of the operator appearing in (B-3), factors containing square roots of occupation numbers will be introduced for each occurrence of the index $\mathbf{k}$; the ket $|\Psi_P\rangle$ is therefore not a simple tensor product, and its norm is not simply the sum of the squared moduli of the $g_{\mathbf{k}}$ raised to the power P. It will be simpler for the following computations to consider the $g_{\mathbf{k}}$ as entirely free parameters, and hence not impose a normalization of the state $|\Psi_P\rangle$.

On can choose to take a finite or infinite number of non-zero $g_{\mathbf{k}}$. The simplest case is the one already discussed above, where $g_{\mathbf{k}} \propto \delta_{\mathbf{k},\mathbf{k}_0}$; the ket $|\Psi_P\rangle$ then becomes proportional to a simple Fock state where only two states of opposite momenta are occupied. For other functions $g_{\mathbf{k}}$, the structure of the paired state will be more complex; adjusting those parameters allows a fine tuning of the particle correlation properties, which is not possible with a simple Fock state.

B-1-b. Fermions in a singlet state

Another frequently encountered case concerns fermionic particles in a singlet state; we must then use operator (A-23). The paired state is then:

$$|\Psi_P\rangle = \left[A^\dagger_{\mathbf{K}=0}\right]^P |0\rangle \quad = \left\{\sum_{\mathbf{k}} g_{\mathbf{k}}\, a^\dagger_{\mathbf{k},\uparrow} a^\dagger_{-\mathbf{k},\downarrow}\right\}^P |0\rangle \qquad \text{with } N = 2P \tag{B-4}$$

The summation over $\mathbf{k}$ runs over all the non-zero wave vectors, without the restriction (B-3) to the half-space D (because of the spins, the pairs of states $\mathbf{k},\uparrow$, $-\mathbf{k},\downarrow$ and $-\mathbf{k},\uparrow$, $\mathbf{k},\downarrow$ are different).

Here again we see that the normalization of $|\Psi_P\rangle$ does not simply reduce to condition (A-3). When $P > 1$, the same index $\mathbf{k}$ may appear twice (or more) in the expansion of the power P of the operator on the right-hand side of (B-4); the corresponding component cancels out since the square of any fermionic creation operator is zero. The norm of the ket $|\Psi_P\rangle$ is therefore a complex expression. Rather than imposing this norm to be equal to one, it is easier to let it vary, and consider the $g_{\mathbf{k}}$ to be totally free variational parameters.

B-1-c. Consequences of the symmetrization

The state vectors (B-3) for bosons, and (B-4) for fermions, are not simple juxtapositions of P pairs of particles, each being described by the relative wave function $\chi(\mathbf{r})$, with $g_{\mathbf{k}}$ as its Fourier transform, according to (A-2). As we already saw, the symmetrization or antisymmetrization of the $2P$-particle paired states strongly affect their norm; it also affects the very structure of these states, which are not merely the tensor product of P pair states. This is particularly obvious for fermions: expanding the sum of operators to the power P in the curly brackets of (B-4), we get the product of P sums

over indices $\mathbf{k}_1$, $\mathbf{k}_2$,.., $\mathbf{k}_P$, and many terms will cancel out: all those for which two (or more) summation indices $\mathbf{k}$ are equal (in which case we get squares of creation operators, which are zero for fermions).

There exists, however, a limiting case where the paired state practically describes the juxtaposition of P binary molecules. It occurs when the wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ varies over a very short range, and hence has a large number of significant Fourier components $g_{\mathbf{k}}$. When this number is much larger than the number of pairs P, most of the terms do not contain several occurrences of the same summation index $\mathbf{k}$, and consequently the paired state vector is very close to the tensor product of the pairs of particles. This state vector describes a gas formed by very strongly bound binary molecules, each moving in the mean field created by all the others (it is, in a certain sense, a "molecule Fock state"). This is actually a very special case; in general, when the pair wave function does not obey that criterion, we can study many other physical situations, hence the interest for introducing paired states.

Even though the values of $g_{\mathbf{k}}$ or the wave function $\chi(\mathbf{r})$ of a "molecule" are mathematically the starting ingredient that allows building $|\Psi_P\rangle$, the resulting state after symmetrization has a complex structure, hard to describe in terms of molecules. On the other hand, this state has a simple property: it contains exactly $N = 2P$ particles, since this is the case for each of its non-zero components; as all the particles are paired, it contains exactly P pairs.

B-2. Undetermined particle number

Computations with the ket $|\Psi_P\rangle$ written above (and in particular its normalization) are not easy: a great number of individual states $\mathbf{k}$ appear inside the curly brackets, which must be raised to a very large power P. This practical difficulty leads us to introduce another variational state $|\Psi_{\text{paired}}\rangle$ where the total number of particles is no longer fixed. This new state, which leads to simpler calculations[2], is defined, starting with (B-2), by:

$$|\Psi_{\text{paired}}\rangle = \sum_{P=0}^{\infty} \frac{1}{P!} |\Psi_P\rangle = \sum_{P=0}^{\infty} \frac{1}{P!} \left[A^{\dagger}_{\mathbf{K}=0}\right]^P |0\rangle \tag{B-5}$$

The $|\Psi_P\rangle$ are not normalized; multiplying all the $g_{\mathbf{k}}$, and hence $A_{\mathbf{K}=0}$, by the same constant α, changes their norm by the factor α^P. This results in varying the relative weights of the terms in the serie (B-5). The larger α, the more weight is placed on the high values of P, which is a way, for example, of modifying the average particle number.

In (B-5) we recognize the series expansion of an exponential, so that:

$$|\Psi_{\text{paired}}\rangle = \exp\left\{A^{\dagger}_{\mathbf{K}=0}\right\} |0\rangle \tag{B-6}$$

This property will greatly simplify the following calculations and is the major reason for letting the total particle number fluctuate.

Writing (B-5), we chose a state vector that is the superposition of states corresponding to different total particle numbers N; there are actually no physical processes taken into account in our approach that could create such a coherent superposition. This

[2]This does not mean that computations with a variational state having a fixed number of particles are always impossible, as shown for example in the treatment of the BCS theory in § 5.4 and Appendix 5C of the book [8].

operation reminds us of the passage from the canonical to the grand canonical ensemble where one introduces, for mathematical convenience, an (incoherent) statistical mixture of different N values. In our present case, however, we are dealing with a coherent superposition, introduced arbitrarily as we just did, and we may wonder whether it might radically change the physics of our problem. This is actually not the case for two reasons. The first is that, for very large values of N, we are going to show that the components of $|\Psi_{\text{paired}}\rangle$ are only important in a domain of N whose width is very small compared to the average value of the particle number; the distribution of the possible values for N is thus very narrow, in relative value, and the particle number remains quite well defined. The second reason is that we shall compute average values of operators that, such as $\widehat{H}$, conserve the total particle number, and for which the coherence of the state vector between kets of different N values is irrelevant. The average value in the coherent state $|\Psi_{\text{paired}}\rangle$ is therefore the weighted average of the average values obtained for each N which, when the average value of the particle number is very large, are approximately the same (since the N distribution is very narrow). In other words, the average values we are going to compute are good approximations of those we would obtain by projecting $|\Psi_{\text{paired}}\rangle$ onto one of its main components with fixed N; using the coherent superposition (B-5) is thus very convenient from a mathematical point of view, without greatly perturbing the results from a physical point of view. A more detailed discussion of this question will be presented in § 1 of Complement B_{XVII}.

B-2-a. Particles in the same spin state

When all the particles are in the same spin state, inserting (A-13) in (B-6) leads to:

$$|\Psi_{\text{paired}}\rangle = \exp\left\{\sum_{\mathbf{k}\in D} \sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right\}|0\rangle \tag{B-7}$$

The operators $a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}$ and $a^{\dagger}_{\mathbf{k}'} a^{\dagger}_{-\mathbf{k}'}$ commute with each other (for fermions in the same spin state, two minus signs cancel each other as we commute products of two operators). It then follows that the exponential of the sum is a product of exponentials, and we can write:

$$\begin{aligned} |\Psi_{\text{paired}}\rangle &= \left\{\prod_{\otimes \mathbf{k}\in D} \exp\left[\sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right]\right\}|0\rangle \\ &= \prod_{\otimes \mathbf{k}\in D} |\varphi_{\mathbf{k}}\rangle \end{aligned} \tag{B-8}$$

The state vector $|\Psi_{\text{paired}}\rangle$ is then simply a tensor product[3] of state vectors $|\varphi_{\mathbf{k}}\rangle$:

$$|\varphi_{\mathbf{k}}\rangle = \exp\left\{\sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right\}|0\rangle \tag{B-9}$$

[3]The Fock space is the tensor product of the states associated with all the individual quantum states $|\mathbf{k}\rangle$, each having any positive occupation number. One can regroup those spaces in pairs corresponding to opposite values of $\mathbf{k}$, and introduce spaces $\mathcal{E}(\mathbf{k})$ whose tensor product is also the Fock space. To build a basis in those spaces, one must vary two occupation numbers.

The restriction of the summation over $\mathbf{k}$ to a half-space D, introduced above, prevents each component of the tensor product from appearing twice in (B-8).

For fermions in a single spin state, the square of any creation operator is zero; the exponential reduces to the sum of the first two terms of its expansion:

$$|\varphi_{\mathbf{k}}\rangle = \left[1 + \sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right] |0\rangle \qquad \text{(fermions only)} \tag{B-10}$$

B-2-b. Fermions in a singlet state

For paired fermions in a singlet state, the state $|\Psi_{\text{paired}}\rangle$ will be called the "BCS state" (Complement C_{XVII}) and noted $|\Psi_{\text{BCS}}\rangle$; relation (A-23) must be used with $\mathbf{K} = 0$. As the exponential of a sum of commuting operators[4] is a product of operators, we get:

$$|\Psi_{\text{BCS}}\rangle = \exp\left\{\sum g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right\} |0\rangle = \prod_{\otimes \mathbf{k}} |\varphi_{\mathbf{k}}\rangle \tag{B-11}$$

with:

$$|\varphi_{\mathbf{k}}\rangle = \exp\left[g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{B-12}$$

As the square of any fermion creation operator is zero, the series expansion of the exponential is limited to its first two terms:

$$|\varphi_{\mathbf{k}}\rangle = \left[1 + g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \qquad \text{(fermions only)} \tag{B-13}$$

B-3. Pairs of particles and pairs of individual states

Pairs of states is an important concept not to be confused with pairs of particles. In (B-7) as well as in (B-11), the individual states intervene as "pairs of states" $(\mathbf{k}, -\mathbf{k})$. The number of those pairs (which can be infinite if D is infinite) is not related to the particle number. For fermions in a singlet state, it is convenient to label the pair of states by the momentum $\mathbf{k}$ associated with the spin state $\uparrow$, while remembering that the momentum associated with the spin state $\downarrow$ is the opposite, $-\mathbf{k}$. We shall systematically use this simplification in what follows.

C. Properties of the kets characterizing the paired states

Let us examine a few properties of the states $|\varphi_{\mathbf{k}}\rangle$ that will be useful in what follows. To keep things simple, we continue limiting in this § C the generality of the cases under study, and assume the particles in the same spin states are bosons; as for the particles in different spin states, we shall continue using the example of fermions in a singlet state. The generalization to other paired cases does not introduce any particular difficulties.

C-1. Normalization

The normalization of the states $|\varphi_{\mathbf{k}}\rangle$ is actually simpler for fermions than for bosons; this is because, as we shall see below, the series expansion of the exponential (B-12) contains only two terms for fermions, instead of an infinity for bosons. This is why we do not keep in this § C the same order as in § A and start with the study of spin 1/2 fermions.

[4] The operators $a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}$ and $a^{\dagger}_{\mathbf{k}',\uparrow} a^{\dagger}_{-\mathbf{k}',\downarrow}$ associated with different pairs commute, since they are products of two fermionic operators.

C-1-a. Fermions in a singlet state

We choose to normalize separately each of the $|\varphi_{\mathbf{k}}\rangle$ by multiplying them by a number $u_{\mathbf{k}}$. This operation amounts to replacing $|\varphi_{\mathbf{k}}\rangle$ by:

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{C-1}$$

with:

$$v_{\mathbf{k}} = u_{\mathbf{k}}\ g_{\mathbf{k}} \tag{C-2}$$

The normalization condition becomes:

$$|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1 \tag{C-3}$$

It then becomes natural to set:

$$\begin{aligned} u_{\mathbf{k}} &= \cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}} \end{aligned} \tag{C-4}$$

where $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are the two variables[5] the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ depends on. One can choose $\theta_{\mathbf{k}}$ between 0 and $\pi/2$:

$$0 \leq \theta_{\mathbf{k}} \leq \frac{\pi}{2} \tag{C-5}$$

so that $\cos\theta_{\mathbf{k}}$ and $\sin\theta_{\mathbf{k}}$ are positive and represent the moduli of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$. We saw in § A-2 that $g_{\mathbf{k}} = g_{-\mathbf{k}}$; the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are therefore even with respect to $\mathbf{k}$.

The variational ket $|\Psi_{\mathrm{BCS}}\rangle$ now becomes the normalized ket $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$:

$$\begin{aligned} \left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle &= \prod\nolimits_{\otimes \mathbf{k}} \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \\ &= \prod\nolimits_{\otimes \mathbf{k}} \left[\cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} + \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \end{aligned} \tag{C-6}$$

Comment:

A particular case occurs when all the $\theta_{\mathbf{k}}$ are either zero or equal to $\pi/2$. The ket $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$ then reduces to a simple Fock state, whose populations of individual states are either zero, or equal to one (for populations corresponding to states belonging to a pair for which $\theta_{\mathbf{k}} = \pi/2$). In that case, the phases $\zeta_{\mathbf{k}}$ no longer play any role: instead of fixing a relative phase, they only determine the global phase of the state vector.

If, furthermore, we choose $\theta_{\mathbf{k}} = \pi/2$ for all values of $\mathbf{k}$ whose modulus is less than a given value k_F, and zero otherwise, the paired state now describes an ensemble of fermions filling two Fermi spheres (one for each spin state), which is simply the ground state of an ideal gas of fermions. The ket $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$ then reduces to the trial ket of the Hartree-Fock method of Complement $\mathrm{B_{XV}}$; that method appears as a particular case of the more general pairing method used in this chapter.

[5] The variable $\zeta_{\mathbf{k}}$ determines the difference $2\zeta_{\mathbf{k}}$ between the phases of $v_{\mathbf{k}}$ and $u_{\mathbf{k}}$. We could also introduce a variable to determine their sum, but that would be pointless: such a variable would only change the total phase of the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$, without any physical consequences.

C-1-b. Bosons in the same spin state

For bosons, the results are slightly different. To maintain a certain analogy, we shall use the same parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ as for fermions, but it is now the hyperbolic sine and cosine of $\theta_{\mathbf{k}}$ that will come into play. Relation (B-9) leads to:

$$|\varphi_{\mathbf{k}}\rangle = \sum_{q=0}^{\infty} \frac{1}{q!} \left[\sqrt{2} g_{\mathbf{k}}\ a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right]^q |0\rangle = \sum_{q=0}^{\infty} \frac{1}{q!} \left[-x_{\mathbf{k}}\right]^q \left[a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right]^q |0\rangle \tag{C-7}$$

with[6]:

$$x_{\mathbf{k}} = -\sqrt{2}\ g_{\mathbf{k}} \tag{C-8}$$

As mentioned before, the spin index that comes in addition to the index $\mathbf{k}$ is not written explicitly as its value does not change.

Consequently:

$$\begin{aligned} \langle\varphi_{\mathbf{k}}\ |\varphi_{\mathbf{k}}\rangle &= \sum_{q=0}^{\infty} \left(\frac{1}{q!}\right)^2 |x_{\mathbf{k}}|^{2q} \left(\sqrt{q!}\right)^4 = \sum_{q=0}^{\infty} |x_{\mathbf{k}}|^{2q} \\ &= \frac{1}{1-|x_{\mathbf{k}}|^2} \end{aligned} \tag{C-9}$$

We assumed, to sum the series, that:

$$|x_{\mathbf{k}}|^2 < 1 \tag{C-10}$$

It is useful in what follows to characterize the complex variable $x_{\mathbf{k}}$ by two real variables: $\theta_{\mathbf{k}}$ to define its modulus, and an angle $\zeta_{\mathbf{k}}$ that characterizes its phase. We therefore set:

$$x_{\mathbf{k}} = \tanh\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \qquad \text{with:} \quad \theta_{\mathbf{k}} \geq 0 \tag{C-11}$$

Inequality (C-10) is automatically satisfied since the modulus of a hyperbolic tangent is always less than 1; as the function $g_{\mathbf{k}}$ is even – see relation (A-4) – so are the variable $x_{\mathbf{k}}$ and the functions[7] $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$. We then get:

$$\langle\varphi_{\mathbf{k}}\ |\varphi_{\mathbf{k}}\rangle = \frac{1}{1-\tanh^2\theta_{\mathbf{k}}} = \cosh^2\theta_{\mathbf{k}} \tag{C-12}$$

The normalized kets $|\overline{\varphi}_{\mathbf{k}}\rangle$ can be written as:

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{1}{\cosh\theta_{\mathbf{k}}} |\varphi_{\mathbf{k}}\rangle = \frac{1}{\cosh\theta_{\mathbf{k}}} \exp\left[-x_{\mathbf{k}}\ a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right] |0\rangle \tag{C-13}$$

Replacing the $|\varphi_{\mathbf{k}}\rangle$ by the $|\overline{\varphi}_{\mathbf{k}}\rangle$, the ket $|\Psi_{\text{paired}}\rangle$ becomes normalized to 1.

Initially, the kets $|\varphi_{\mathbf{k}}\rangle$, as well as their normalized version $|\overline{\varphi}_{\mathbf{k}}\rangle$, have been defined in the tensor product (B-8) only when $\mathbf{k}$ belongs to the half-space D. They can, however, be defined by relations (C-7) and (C-13) for any $\mathbf{k}$; we then have simply $|\overline{\varphi}_{\mathbf{k}}\rangle = |\overline{\varphi}_{-\mathbf{k}}\rangle$, which was to be expected since $|\overline{\varphi}_{\mathbf{k}}\rangle$ involves the two individual states $\mathbf{k}$ and $-\mathbf{k}$ in the same way.

[6] The minus sign in this definition is arbitrary – a change of sign of the wave function $\chi(\mathbf{r})$ or of its Fourier transform $g_{\mathbf{k}}$ has no physical consequences – but it is convenient to introduce this sign to ensure coherence with the calculations in § E.

[7] Furthermore, rotational invariance generally requires those functions to depend only on the modulus k of $\mathbf{k}$.

C-2. Average value and root mean square deviation of particle number

The particle number in the individual state $\mathbf{k}$ corresponds to the operator:

$$\widehat{n}_{\mathbf{k}} = a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}} \tag{C-14}$$

We are now going to compute the average value and the root mean square deviation of the particle number, first in a given pair of states, then for the system as a whole.

C-2-a. Fermions in a singlet state

Let us compute the average value of the particle number in the state $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$, which is the tensor product of the states $|\overline{\varphi}_{\mathbf{k}}\rangle$, each being associated with the pair of states $(\hbar\mathbf{k}, \nu = +; -\hbar\mathbf{k}, \nu = -)$; as defined above, each pair is labeled by the wave vector $\mathbf{k}$ of the spin $+$ particle. The particle number in each of these pairs of states corresponds to the operator:

$$\widehat{n}_{(\text{pair } \mathbf{k})} = \widehat{n}_{\mathbf{k},\uparrow} + \widehat{n}_{-\mathbf{k},\downarrow} = a^{\dagger}_{\mathbf{k},\uparrow} a_{\mathbf{k},\uparrow} + a^{\dagger}_{-\mathbf{k},\downarrow} a_{-\mathbf{k},\downarrow} \tag{C-15}$$

with eigenvalues 0, 1 and 2. Now $|\overline{\varphi}\rangle_{\mathbf{k}}$ is given by (C-1), the sum of two components, one with zero particles, and the other with two particles. This leads to:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, \widehat{n}_{(\text{pair } \mathbf{k})}\, |\overline{\varphi}_{\mathbf{k}}\rangle = 2\,|v_{\mathbf{k}}|^2 = 2\sin^2\theta_{\mathbf{k}} \tag{C-16}$$

and:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, \left[\widehat{n}_{(\text{pair } \mathbf{k})}\right]^2 |\overline{\varphi}_{\mathbf{k}}\rangle = 4\,|v_{\mathbf{k}}|^2 = 4\sin^2\theta_{\mathbf{k}} \tag{C-17}$$

The root mean square deviation $\Delta n_{(\text{pair } \mathbf{k})}$ of the particle number in a pair is thus:

$$\Delta n_{(\text{pair } \mathbf{k})} = \sqrt{4\,|v_{\mathbf{k}}|^2 \left[1 - |v_{\mathbf{k}}|^2\right]} = 2\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}} \tag{C-18}$$

Consequently, the fluctuations of the particle number in each pair of states can be large.

On the other hand, the fluctuations of the total particle number, obtained by summing over all the pairs, remain small. The average value of this total number is:

$$\left\langle \widehat{N} \right\rangle = 2\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = 2\sum_{\mathbf{k}} \sin^2\theta_{\mathbf{k}} \tag{C-19}$$

As we will show just below, the square of the fluctuation ΔN of $\langle N\rangle$ is given by:

$$[\Delta N]^2 = 4\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 \left[1 - |v_{\mathbf{k}}|^2\right] \tag{C-20}$$

Since $\left[1 - |v_{\mathbf{k}}|^2\right] \leq 1$, we get:

$$[\Delta N]^2 \leq 4\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = 2\left\langle \widehat{N} \right\rangle \tag{C-21}$$

so that:

$$\frac{\Delta N}{\langle \widehat{N} \rangle} \leq \sqrt{\frac{2}{\langle \widehat{N} \rangle}} \tag{C-22}$$

Hence, for large values of $\langle N \rangle$, the fluctuations of the particle number, in relative value, are very small, decreasing at least as fast as the inverse of the square root of the average value.

Demonstration:

The operator corresponding to the square of the particle number is:

$$\left[\widehat{N}\right]^2 = \sum_{\mathbf{k}} \left[\widehat{n}_{(\text{pair } \mathbf{k})}\right]^2 + \sum_{\mathbf{k} \neq \mathbf{k}'} \widehat{n}_{(\text{pair } \mathbf{k})}\ \widehat{n}_{(\text{pair } \mathbf{k}')} \tag{C-23}$$

As the state $\left|\overline{\Psi}_{\text{BCS}}\right\rangle$ is a product of states of pairs, the latter are not correlated and the average value of this operator is written:

$$\begin{aligned} &\left\langle \overline{\Psi}_{\text{BCS}} \right| \left[\widehat{N}\right]^2 \left| \overline{\Psi}_{\text{BCS}} \right\rangle \\ &= \sum_{\mathbf{k}} \langle \overline{\varphi_{\mathbf{k}}} | \left[\widehat{n}_{(\text{pair } \mathbf{k})}\right]^2 | \overline{\varphi}_{\mathbf{k}} \rangle + \sum_{\mathbf{k} \neq \mathbf{k}'} \langle \overline{\varphi}_{\mathbf{k}} | \widehat{n}_{(\text{pair } \mathbf{k})} | \overline{\varphi}_{\mathbf{k}} \rangle \langle \overline{\varphi}_{\mathbf{k}'} | \widehat{n}_{(\text{pair } \mathbf{k}')} | \overline{\varphi}_{\mathbf{k}'} \rangle \end{aligned} \tag{C-24}$$

Expression (C-1) for $|\overline{\varphi}_{\mathbf{k}}\rangle$ leads to:

$$\widehat{n}_{(\text{pair } \mathbf{k})} | \overline{\varphi}_{\mathbf{k}} \rangle = 2 v_{\mathbf{k}} | \mathbf{k}, \uparrow; -\mathbf{k}, \downarrow \rangle \tag{C-25}$$

so that:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = 4 \sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 + 4 \sum_{\mathbf{k} \neq \mathbf{k}'} |v_{\mathbf{k}}|^2 |v_{\mathbf{k}'}|^2 \tag{C-26}$$

Now the square of the average value $\langle N \rangle$ is equal to the last terms of this equality, but without the constraint $\mathbf{k} \neq \mathbf{k}'$ in the summation. It follows that the root mean square ΔN is written as:

$$[\Delta N]^2 = \left\langle \left[\widehat{N}\right]^2 \right\rangle - [\langle N \rangle]^2 = 4 \sum_{\mathbf{k}} \left[|v_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^4 \right] \tag{C-27}$$

which leads to (C-20).

C-2-b. Bosons in the same spin state

For bosons, each pair contains two individual states of opposite $\mathbf{k}$. We show below that for each of them, we have:

$$\langle \widehat{n}_{\mathbf{k}} \rangle = \sinh^2 \theta_{\mathbf{k}} \tag{C-28}$$

and that:

$$\left\langle [\widehat{n}_{\mathbf{k}}]^2 \right\rangle = 2 \langle \widehat{n}_{\mathbf{k}} \rangle^2 + \langle \widehat{n}_{\mathbf{k}} \rangle \tag{C-29}$$

The root mean square deviation of the distribution associated with the values of $n_{\mathbf{k}}$ is thus:

$$\begin{aligned}\Delta n_{\mathbf{k}} &= \sqrt{\left\langle \left[\widehat{n}_{\mathbf{k}}\right]^2 \right\rangle - \left\langle \widehat{n}_{\mathbf{k}} \right\rangle^2} = \sqrt{\left\langle \widehat{n}_{\mathbf{k}} \right\rangle^2 + \left\langle \widehat{n}_{\mathbf{k}} \right\rangle} \\ &= \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}} \end{aligned} \tag{C-30}$$

(the average value of the particle number in a pair of states is $2\left\langle \widehat{n}_{\mathbf{k}} \right\rangle$, and the root mean square deviation of that number is $2\Delta n_{\mathbf{k}}$).

Demonstration:

As $|\varphi_{\mathbf{k}}\rangle$ is symmetric with respect to the two individual states $\mathbf{k}$ and $-\mathbf{k}$, we have:

$$\left\langle \widehat{n}_{\mathbf{k}} \right\rangle = \left\langle \widehat{n}_{-\mathbf{k}} \right\rangle \tag{C-31}$$

with:

$$\begin{aligned}\left\langle \varphi_{\mathbf{k}} \right| \widehat{n}_{\mathbf{k}} \left| \varphi_{\mathbf{k}} \right\rangle &= \sum_{q=0}^{\infty} \left|x_{\mathbf{k}}\right|^{2q} \times q = \left|x_{\mathbf{k}}\right|^2 \frac{\partial}{\partial \left|x_{\mathbf{k}}\right|^2} \left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle \\ &= \frac{\left|x_{\mathbf{k}}\right|^2}{\left[1 - \left|x_{\mathbf{k}}\right|^2\right]^2} \end{aligned} \tag{C-32}$$

so that:

$$\left\langle \widehat{n}_{\mathbf{k}} \right\rangle = \frac{\left\langle \varphi_{\mathbf{k}} \right| \widehat{n}_{\mathbf{k}} \left| \varphi_{\mathbf{k}} \right\rangle}{\left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle} = \frac{\left|x_{\mathbf{k}}\right|^2}{1 - \left|x_{\mathbf{k}}\right|^2} \tag{C-33}$$

which leads to (C-28).

The average value of the particle number squared is computed in a similar way. Using the identity $q^2 = q(q-1) + q$ to bring up the second derivative with respect to $\left|x_{\mathbf{k}}\right|^2$, we can write:

$$\begin{aligned}\left\langle \varphi_{\mathbf{k}} \right| \left[\widehat{n}_{\mathbf{k}}\right]^2 \left| \varphi_{\mathbf{k}} \right\rangle &= \sum_{q=0}^{\infty} \left|x_{\mathbf{k}}\right|^{2q} \times q^2 \\ &= \left|x_{\mathbf{k}}\right|^4 \frac{\partial^2}{\left[\partial \left|x_{\mathbf{k}}\right|^2\right]^2} \left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle + \left|x_{\mathbf{k}}\right|^2 \frac{\partial}{\partial \left|x_{\mathbf{k}}\right|^2} \left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle \\ &= \frac{2\left|x_{\mathbf{k}}\right|^4}{\left[1 - \left|x_{\mathbf{k}}\right|^2\right]^3} + \frac{\left|x_{\mathbf{k}}\right|^2}{\left[1 - \left|x_{\mathbf{k}}\right|^2\right]^2} \end{aligned} \tag{C-34}$$

and hence:

$$\left\langle \left[\widehat{n}_{\mathbf{k}}\right]^2 \right\rangle = \frac{\left\langle \varphi_{\mathbf{k}} \right| \left[\widehat{n}_{\mathbf{k}}\right]^2 \left| \varphi_{\mathbf{k}} \right\rangle}{\left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle} = 2\left\langle \widehat{n}_{\mathbf{k}} \right\rangle^2 + \left\langle \widehat{n}_{\mathbf{k}} \right\rangle \tag{C-35}$$

The total number of particles is written:

$$\left\langle \widehat{N} \right\rangle = \sum_{\mathbf{k}} \left|v_{\mathbf{k}}\right|^2 = \sum_{\mathbf{k}} \sinh^2\theta_{\mathbf{k}} \tag{C-36}$$

(as expected, each pair of states appears twice in this summation; one can also restrict the summation to the half-space D provided we add a factor 2). We also have:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = \left\langle \sum_{\mathbf{k}} \widehat{n}_{\mathbf{k}} \sum_{\mathbf{k}'} \widehat{n}_{\mathbf{k}'} \right\rangle = \left\langle \sum_{\mathbf{k}} \widehat{n}_{\mathbf{k}} \left[\widehat{n}_{\mathbf{k}} + \widehat{n}_{-\mathbf{k}} + \sum_{\mathbf{k}' \neq \pm \mathbf{k}} \widehat{n}_{\mathbf{k}'} \right] \right\rangle$$
$$= \sum_{\mathbf{k}} \left[\langle [\widehat{n}_{\mathbf{k}}]^2 \rangle + \langle \widehat{n}_{\mathbf{k}} \widehat{n}_{-\mathbf{k}} \rangle \right] + \sum_{\mathbf{k}, \mathbf{k}' \neq \pm \mathbf{k}} \langle \widehat{n}_{\mathbf{k}} \rangle \langle \widehat{n}_{\mathbf{k}'} \rangle \tag{C-37a}$$

where, in the last term, we have used the fact that the paired state is the product of uncorrelated pairs. But this state is symmetric in $\mathbf{k}$ and $-\mathbf{k}$, and operators $\widehat{n}_{\mathbf{k}}$ and $\widehat{n}_{-\mathbf{k}}$ act on it in the same way. Therefore:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = \sum_{\mathbf{k}} 2 \left[\left\langle [\widehat{n}_{\mathbf{k}}]^2 \right\rangle - \langle \widehat{n}_{\mathbf{k}} \rangle^2 \right] + \sum_{\mathbf{k}, \mathbf{k}'} \langle \widehat{n}_{\mathbf{k}} \rangle \langle \widehat{n}_{\mathbf{k}'} \rangle \tag{C-37b}$$

(the constraint in the second summmation has been eliminated by subtracting a term in the first summation). We then get:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = 2 \sum_{\mathbf{k}} [\Delta n_{\mathbf{k}}]^2 + \langle N \rangle^2 \tag{C-37c}$$

The root mean square deviations $\Delta n_{\mathbf{k}}$ have been obtained in (C-30). Hence, the square of the root mean square deviation of the total number of particles is written as:

$$[\Delta N]^2 = 2 \sum_{\mathbf{k}} [\Delta n_{\mathbf{k}}]^2 = 2 \sum_{\mathbf{k}} \sinh^2 \theta_{\mathbf{k}} \cosh^2 \theta_{\mathbf{k}} \tag{C-38}$$

As for the fermion case, this square contains only a single summation on $\mathbf{k}$, whereas the square of the total particle number contains two. Now the number of non-zero terms in those summations is the number of Fourier components necessary to describe the pair of particles used, in § B, to build the paired state in a cube of edge length L (size of the momentum quantization box – see § A-1). This number is of the order of the cube of the ratio between L and the size of the pair, hence a very large number, as it is the ratio between a macroscopic and a microscopic volume. A double summation over $\mathbf{k}$ therefore contains many more terms than a simple summation, and since all the terms are positive and of comparable magnitude, we have:

$$\left\langle \widehat{N} \right\rangle^2 \gg [\Delta N]^2 \tag{C-39}$$

We again find, as for fermions, that $\Delta N \ll \left\langle \widehat{N} \right\rangle$.

C-3. "Anomalous" average values

For computing average values of the energy (in particular, in Complement $\mathrm{B_{XVII}}$), we will need the average values of products of two creation or annihilation operators. For example, for bosons we will need to calculate:

$$\langle \overline{\varphi}_{\mathbf{k}} |\, a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger \,| \overline{\varphi}_{\mathbf{k}} \rangle \qquad \text{and} \qquad \langle \overline{\varphi}_{\mathbf{k}} |\, a_{\mathbf{k}} a_{-\mathbf{k}} \,| \overline{\varphi}_{\mathbf{k}} \rangle \tag{C-40}$$

We note, right away, that they concern operators that do not conserve the particle number, and this is the reason they are often called "anomalous average values". One could be surprised that such average values come into play while studying physical processes that do not physically imply creation or destruction of particles. We will show that they actually occur in a very natural way in the calculation of the average value of a Hamiltonian that conserves the particle number. The reason is that $|\overline{\varphi}_{\mathbf{k}}\rangle$ is only a component of the total state vector (B-8), in which it is associated with many other $|\overline{\varphi}_{\mathbf{k}'}\rangle$; in the total state vector, the particle number in the state $|\overline{\varphi}_{\mathbf{k}}\rangle$ may, for example, decrease by 2 while the particle number in the state $|\overline{\varphi}_{\mathbf{k}'}\rangle$ simultaneously increases by the same quantity. We are, therefore, performing computations on the components of a state vector that has the same total particle number; the "anomalous" character is only apparent, and is due to the fact that we only consider part of the total state vector.

C-3-a. Fermions in a singlet state

Consider the action of the operator $a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}$ on the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ written in (C-1). Only one of its component, in $v(\mathbf{k})$, remains and, after two anticommutations, we can write:

$$\begin{aligned} a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}|\overline{\varphi}_{\mathbf{k}}\rangle &= v_{\mathbf{k}}a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}a^{\dagger}_{\mathbf{k},\uparrow}a^{\dagger}_{-\mathbf{k},\downarrow}|0\rangle = v_{\mathbf{k}}\left[a_{-\mathbf{k},\downarrow}a^{\dagger}_{-\mathbf{k},\downarrow}\right]\left[a_{\mathbf{k},\uparrow}a^{\dagger}_{\mathbf{k},\uparrow}\right]|0\rangle \\ &= v_{\mathbf{k}}|0\rangle \end{aligned} \tag{C-41}$$

Taking the scalar product of this ket with the bra $\langle\overline{\varphi}_{\mathbf{k}}|$, only its component $u^{*}(\mathbf{k})\langle 0|$ remains; the average value is thus:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}|\overline{\varphi}_{\mathbf{k}}\rangle = u^{*}_{\mathbf{k}}v_{\mathbf{k}} = \sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{2i\zeta_{\mathbf{k}}} \tag{C-42}$$

The anticommutation of these two operators then yields:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a_{\mathbf{k},\uparrow}a_{-\mathbf{k},\downarrow}|\overline{\varphi}_{\mathbf{k}}\rangle = -u^{*}_{\mathbf{k}}v_{\mathbf{k}} = -\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{2i\zeta_{\mathbf{k}}} \tag{C-43}$$

Taking the Hermitian conjugate of (C-42), we get:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a^{\dagger}_{\mathbf{k},\uparrow}a^{\dagger}_{-\mathbf{k},\downarrow}|\overline{\varphi}_{\mathbf{k}}\rangle = u_{\mathbf{k}}v^{*}_{\mathbf{k}} = \sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{-2i\zeta_{\mathbf{k}}} \tag{C-44}$$

whereas the average value of $a^{\dagger}_{-\mathbf{k},\downarrow}a^{\dagger}_{\mathbf{k},\uparrow}$ is the opposite (anticommutation):

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a^{\dagger}_{-\mathbf{k},\downarrow}a^{\dagger}_{\mathbf{k},\uparrow}|\overline{\varphi}_{\mathbf{k}}\rangle = -\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{-2i\zeta_{\mathbf{k}}} \tag{C-45}$$

We saw, in § C-1-a, that the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even; we can therefore change the sign of $\mathbf{k}$ on the left-hand side of the previous relations without changing the right-hand side.

C-3-b. Bosons in the same spin state

For bosons, it is easier to first compute the average value of a product of creation operators:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{1}{\cosh^{2}\theta_{\mathbf{k}}}\langle\varphi_{\mathbf{k}}|\, a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}|\varphi_{\mathbf{k}}\rangle \tag{C-46}$$

This expression contains the product of the ket:

$$a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} |\varphi_{\mathbf{k}}\rangle = a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} \sum_{q=0}^{\infty} [-x_{\mathbf{k}}]^q |n_{\mathbf{k}} = q; n_{-\mathbf{k}} = q\rangle$$
$$= \sum_{q=0}^{\infty} (q+1) [-x_{\mathbf{k}}]^q |n_{\mathbf{k}} = q+1; n_{-\mathbf{k}} = q+1\rangle \tag{C-47}$$

by the bra:

$$\sum_{q'=0}^{\infty} [-x^*_{\mathbf{k}}]^{q'} \langle n_{\mathbf{k}} = q'; n_{-\mathbf{k}} = q'| \tag{C-48}$$

To get a non-zero term, we must have $q' = q+1$, which leads to:

$$(q+1) [-x_{\mathbf{k}}]^q [-x^*_{\mathbf{k}}]^{q+1} \tag{C-49}$$

whose sum over q yields, taking (C-32) into account:

$$\sum_{q=0}^{\infty} (q+1) \quad [-x^*_{\mathbf{k}}] \quad |x_{\mathbf{k}}|^{2q} = [-x^*_{\mathbf{k}}] \quad [\langle n_{\mathbf{k}}\rangle + 1] \langle \varphi_{\mathbf{k}} | \varphi_{\mathbf{k}}\rangle \tag{C-50}$$

We finally divide by $\langle \varphi_{\mathbf{k}} | \varphi_{\mathbf{k}}\rangle$ as in (C-46), to obtain, inserting the value (C-11) of $x_{\mathbf{k}}$:

$$\langle \overline{\varphi}_{\mathbf{k}}| a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle = -\tanh\theta_{\mathbf{k}}\, e^{-2i\zeta_{\mathbf{k}}} \left[1 + \sinh^2\theta_{\mathbf{k}}\right]$$
$$= -e^{-2i\zeta_{\mathbf{k}}} \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}} \tag{C-51}$$

As for the other "anomalous" average value $\langle \overline{\varphi}_{\mathbf{k}}| a_{\mathbf{k}} a_{-\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle$, a simple Hermitian conjugation operation shows that it is the complex conjugate of the previous one:

$$\langle \overline{\varphi}_{\mathbf{k}}| a_{\mathbf{k}} a_{-\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle = -e^{2i\zeta_{\mathbf{k}}} \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}} \tag{C-52}$$

As was the case for fermions, the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even, which allows changing the sign of $\mathbf{k}$ on the left-hand side of the previous equations without changing the result.

D. Correlations between particles, pair wave function

As already mentioned in this chapter's introduction, one of the major interest of the paired states is to allow varying the spatial correlation functions of a system of identical particles. In addition to the purely statistical correlations, coming from the indistinguishability of the particles and already present in an ideal gas, we now have a way to include dynamic correlations due to the interactions. Using paired states instead of simple Fock states allows, for example, a better optimization of the energy. We shall limit our study to the two-particle diagonal correlation function, as it is the one that fixes the average value of the interaction Hamiltonian. This will lead us to introduce a new wave function, that we shall name the "pair wave function". In the complements following

this chapter we shall also study non-diagonal correlation functions; it will concern the one-particle correlation function, whose long range behavior may signal the existence of Bose-Einstein condensation, as well as the two-particle correlation function.

In a general way, one may wonder about the physical significance of correlation functions computed in states $|\Psi_{\text{paired}}\rangle$ or $|\Psi_{\text{BCS}}\rangle$, since these states are coherent superpositions of kets containing different particle numbers N. However, correlation functions are average values of operators keeping the particle number constant, and hence independent of the coherence between kets of different N values. Furthermore, we saw in § C-2 that for large values of the average particle number $\left\langle \widehat{N} \right\rangle$, the relative fluctuations of that number were negligible. In the limit of large $\left\langle \widehat{N} \right\rangle$, one can thus expect the results obtained with $|\Psi_{\text{paired}}\rangle$ or $|\Psi_{\text{BCS}}\rangle$ to be very close to those obtained with the $|\Psi_P\rangle$, for which these fluctuations are strictly zero. This question will be discussed in more detail in § 1 of Complement B_{XVII}.

When studying correlation functions in the case where the paired particles are in the same spin state, the only relevant indices concern the orbital variables. We shall start with this simpler case, and study later the case of paired particles in a singlet state.

D-1. Particles in the same spin state

Relation (B-34) of Chapter XVI indicates that the two-particle diagonal correlation function $G_2(\mathbf{r}_1, \mathbf{r}_2)$ can be written:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \left\langle \Psi^\dagger(\mathbf{r}_1)\, \Psi^\dagger(\mathbf{r}_2)\, \Psi(\mathbf{r}_2)\, \Psi(\mathbf{r}_1) \right\rangle \tag{D-1}$$

Replacing the field operators and their adjoints by expressions (A-3) and (A-6) of Chapter XVI, using as a basis the normalized plane waves, we get:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{L^6} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4} e^{i[(\mathbf{k}_4 - \mathbf{k}_1)\cdot \mathbf{r}_1 + (\mathbf{k}_3 - \mathbf{k}_2)\cdot \mathbf{r}_2]} \left\langle a^\dagger_{\mathbf{k}_1} a^\dagger_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4} \right\rangle \tag{D-2}$$

where the average value of the product of the 4 creation and annihilation operators must be taken in a paired state. Figure 1 symbolizes the different terms present in this correlation function.

D-1-a. Simplifications due to pairing

The computation is greatly simplified by noting that in a paired state, the populations of the two individual states having opposite wave numbers $\mathbf{k}$ and $-\mathbf{k}$ must always be equal. Consequently, only those combinations of the 4 operators that do not change this equality will lead to non-zero average values. Three cases are then possible:

– **Case I**: the two annihilation operators concern two individual states that do not belong to the same pair ($\mathbf{k}_3 \neq \pm \mathbf{k}_4$); the two creation operators must then restore to their initial values the populations of these two same states, or else their average value will be zero; these are the so called "forward scattering" terms. We then have either $\mathbf{k}_4 = \mathbf{k}_1$ and $\mathbf{k}_3 = \mathbf{k}_2$ (direct term), or $\mathbf{k}_4 = \mathbf{k}_2$ and $\mathbf{k}_3 = \mathbf{k}_1$ (exchange term).

– **Case II**: the two annihilation operators act on the two states of a first pair ($\mathbf{k}_4 = -\mathbf{k}_3$), and the creation operators on the two states of another pair ($\mathbf{k}_2 = -\mathbf{k}_1$). We then talk about a "pair annihilation-creation process" .

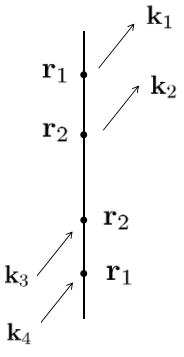


Figure 1: This diagram symbolizes the terms that come into play in the computation of the correlation function of two particles at points $\mathbf{r}_1$ *and* $\mathbf{r}_1$*. The two incoming arrows at the bottom left-hand side represent the two particles eliminated by the annihilation operators; they are associated with a positive imaginary exponent of the position. The two outgoing arrows on the top right-hand side represent the two particles resulting from the action of the creation operators, associated with a negative imaginary exponent. The correlation function is the sum of these terms over all the values of the* 4 $\mathbf{k}$ *vecteurs.*

– **Case III**: the two annihilation operators act on the two states of the same pair, and the creation operators replenish these same two states (this is a special case of the one we just discussed); another possibility is that the two annihilation operators act on the same individual state (all the wave numbers $\mathbf{k}$ must then be equal).

Using these conditions on the values of the wave numbers in (D-2), we note that the terms corresponding to cases I and II include two summations over the wave vectors, whereas there is only one summation in the terms corresponding to case III. Consequently, for a large (macroscopic) volume L^3, there are far fewer terms coming from case III than from cases I and II. We shall therefore only take into account terms arising from case I and II. For the same reason, we shall ignore in our computation of these terms the constraints $\mathbf{k}_3 \neq \pm\mathbf{k}_4$ or $\mathbf{k}_3 \neq \pm\mathbf{k}_1$, as this amounts to adding a negligible number of terms.

D-1-b. Expression of the correlation function

The direct term is obtained for $\mathbf{k}_4 = \mathbf{k}_1$ and $\mathbf{k}_3 = \mathbf{k}_2$; it no longer has any spatial dependence. Since $\mathbf{k}_1$ and $\mathbf{k}_2$ are different, the average value of the product of operators can also be written $\left\langle a^\dagger_{\mathbf{k}_1} a_{\mathbf{k}_1} a^\dagger_{\mathbf{k}_2} a_{\mathbf{k}_2} \right\rangle$ – for fermions, the two minus signs coming from the anticommutations cancel each other. Now relation (B-8) shows that the paired state is a tensor product of pairs of states. This means that the average value we wish to determine is simply the product of the average values of the first two operators and of the last two operators, i.e. the product of the average values of two occupation numbers. We thus

get a first contribution:

$$G_2^{\text{dir}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2}\langle\widehat{n}_{\mathbf{k}_1}\rangle\langle\widehat{n}_{\mathbf{k}_2}\rangle = \frac{\left\langle\widehat{N}\right\rangle^2}{L^6} \tag{D-3}$$

where the summation over $\mathbf{k}_1$ and $\mathbf{k}_2$ are considered as independent, since as we mentioned above, we can neglect the constraint linking these two indices.

The exchange term is obtained for $\mathbf{k}_4 = \mathbf{k}_2$ and $\mathbf{k}_3 = \mathbf{k}_1$; it exhibits a spatial dependence. As we did for the direct term, we regroup the creation and annihilation operators acting on the same individual states, but this operation now involves only one commutation between operators. We then introduce a factor η, equal to -1 for fermions, and we get:

$$G_2^{\text{ex}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\eta}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i(\mathbf{k}_2-\mathbf{k}_1)\cdot(\mathbf{r}_1-\mathbf{r}_2)}\ \langle\widehat{n}_{\mathbf{k}_1}\rangle\langle\widehat{n}_{\mathbf{k}_2}\rangle \tag{D-4}$$

The pair annihilation-creation term $\mathbf{k}_4 = -\mathbf{k}_3$ and $\mathbf{k}_2 = -\mathbf{k}_1$ also exhibits a spatial dependence, but no longer involves average values of occupation numbers. Its expression is:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_4} e^{i(\mathbf{k}_4-\mathbf{k}_1)\cdot(\mathbf{r}_1-\mathbf{r}_2)}\ \left\langle a^\dagger_{\mathbf{k}_1}a^\dagger_{-\mathbf{k}_1}\right\rangle\left\langle a_{-\mathbf{k}_4}a_{\mathbf{k}_4}\right\rangle \tag{D-5}$$

and its structure is schematized in Figure 2. Expression (D-5) contains average values of products of operators that do not conserve the particle number, but rather annihilate (or create) two of them. They are called "anomalous average values". As we explained in § C-3, these anomalous average values come into play quite naturally in the computation of the average value of an operator that does conserve the particle number. Defining the "pair wave function" ϕ_{pair} as:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}}\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle \tag{D-6}$$

this correlation function can be written as:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\mathbf{r}_2) = \left|\phi_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\right|^2 \tag{D-7}$$

The complete correlation function $G_2(\mathbf{r}_1,\mathbf{r}_2)$ is the sum of the three previous contributions:

$$G_2(\mathbf{r}_1,\mathbf{r}_2) = G_2^{\text{dir}}(\mathbf{r}_1,\mathbf{r}_2) + G_2^{\text{ex}}(\mathbf{r}_1,\mathbf{r}_2) + G_2^{\text{pair-pair}}(\mathbf{r}_1,\mathbf{r}_2) \tag{D-8}$$

For bosons in the same spin state, we can insert in this correlation function the average values given in (C-28), (C-51) and (C-52). We then get a binary correlation function that explicitly depends on the parameters $\theta_{\mathbf{k}}$, as well as on the phases $\zeta_{\mathbf{k}}$, which both define the paired state. This clearly verifies that these parameters introduce flexibility in the two-body correlation function. For example, we find:

$$\phi_{\text{pair}}(\mathbf{r}) = -\frac{1}{L^3}\sum_{\mathbf{k}}\sinh\theta_{\mathbf{k}}\ \cosh\theta_{\mathbf{k}}\ e^{-i(\mathbf{k}\cdot\mathbf{r}-2\zeta_{\mathbf{k}})} \tag{D-9}$$

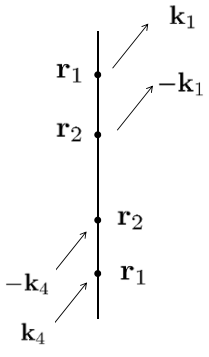


Figure 2: Diagram symbolizing the pair-pair term of the binary correlation function, with the same convention as for Figure 1.

The pair wave function thus directly depends on the phases $\zeta_{\mathbf{k}}$; as we shall see in the complements, these phases actually play a major role in the optimization of the energy. For bosons, this wave function is always even since, as we saw in § C-1-b, this is the case for the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ introduced in (C-11).

Comment:

To describe systems of interacting bosons undergoing Bose-Einstein condensation (see § 4 of Complement $\mathrm{B_{XVII}}$ and Complement $\mathrm{E_{XVII}}$, we shall add to the paired state $|\Psi_{\mathrm{paired}}\rangle$ another highly populated state with zero momentum ($\mathbf{k} = 0$). This will introduce new terms in the correlation functions, in addition to those computed in this chapter. When the population of that individual state with zero momentum is very high, these additional terms may become dominant.

D-2. Fermions in a singlet state

For fermions with spin 1/2, since each spin can point in two directions, there exists a larger number of correlation functions. Several among them will be studied in §2 of Complement $\mathrm{C_{XVII}}$. We shall only compute one of them here, involving opposite spins, as it plays the most significant role:

$$G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right) = \left\langle \Psi_\uparrow^\dagger\left(\mathbf{r}_1\right)\Psi_\downarrow^\dagger\left(\mathbf{r}_2\right)\Psi_\downarrow\left(\mathbf{r}_2\right)\Psi_\uparrow\left(\mathbf{r}_1\right)\right\rangle \tag{D-10}$$

Relation (D-2) now becomes:

$$G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} e^{i[(\mathbf{k}_4-\mathbf{k}_1)\cdot\mathbf{r}_1+(\mathbf{k}_3-\mathbf{k}_2)\cdot\mathbf{r}_2]}\left\langle a_{\mathbf{k}_1,\uparrow}^\dagger a_{\mathbf{k}_2,\downarrow}^\dagger a_{\mathbf{k}_3,\downarrow} a_{\mathbf{k}_4,\uparrow}\right\rangle \tag{D-11}$$

The diagram schematizing each term of this sum is obtained by adding spin indices to the positions in Figure 1 – as is done in Figure 4 of Complement $\mathrm{C_{XVII}}$.

The computation is then similar to that of § D-1. The direct term is written:

$$G_2^{\text{dir}}(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2}\langle\widehat{n}_{\mathbf{k}_1,\uparrow}\rangle\langle\widehat{n}_{\mathbf{k}_2,\downarrow}\rangle = \frac{\left\langle\widehat{N_\uparrow}\right\rangle\left\langle\widehat{N_\downarrow}\right\rangle}{L^6} \tag{D-12}$$

There is no exchange term where $\mathbf{k}_4 = \mathbf{k}_2$ and $\mathbf{k}_3 = \mathbf{k}_1$, as it would correspond to the average value of an operator changing the direction of one of the spins in two different pairs, hence destroying the equality between populations of opposite spins in each pair; this term does exist, however, in the special case where $\mathbf{k}_1 = -\mathbf{k}_2$, but its contribution is negligible. Finally, the pair annihilation-creation term corresponds to $\mathbf{k}_4 = -\mathbf{k}_3$ and $\mathbf{k}_2 = -\mathbf{k}_1$; it is written:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_4} e^{i(\mathbf{k}_4-\mathbf{k}_1)\cdot(\mathbf{r}_1-\mathbf{r}_2)}\ \left\langle a^\dagger_{\mathbf{k}_1,\uparrow}a^\dagger_{-\mathbf{k}_1,\downarrow}\right\rangle\left\langle a_{-\mathbf{k}_4,\downarrow}a_{\mathbf{k}_4,\uparrow}\right\rangle \tag{D-13}$$

Here again, the pair-pair term involves anomalous average values. As before, we can define a pair wave function ϕ_{pair} as:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}}\langle a_{\mathbf{k},\downarrow}a_{-\mathbf{k},\uparrow}\rangle = \frac{1}{L^3}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}\langle a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}\rangle \tag{D-14}$$

whose modulus squared appears in the correlation function:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \left|\phi_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\right|^2 \tag{D-15}$$

Inserting relations (C-42) into (D-14) yields:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} u^*_{\mathbf{k}}v_{\mathbf{k}}\ e^{i\mathbf{k}\cdot\mathbf{r}} = \frac{1}{L^3}\sum_{\mathbf{k}}\sin\theta_{\mathbf{k}}\ \cos\theta_{\mathbf{k}}\ e^{i(\mathbf{k}\cdot\mathbf{r}+2\zeta_{\mathbf{k}})} \tag{D-16}$$

The important role of this pair wave function in the BCS condensation phenomenon will be discussed in detail in Complement C_{XVII}. We will show in particular that this function not only plays a role in the diagonal binary correlation function; it also determines the long-range non-diagonal properties of the density operator, hence playing the role of an order parameter. We noted that the parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even functions of $\mathbf{k}$; consequently, the function $\phi_{\text{pair}}(\mathbf{r})$ is also an even function of $\mathbf{r}$.

The total correlation function is then:

$$G_2(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \frac{\left\langle\widehat{N_\uparrow}\right\rangle\left\langle\widehat{N_\downarrow}\right\rangle}{L^6} + \left|\phi_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\right|^2 \tag{D-17}$$

Inserting in this result expression (D-16) for $\phi_{\text{pair}}(\mathbf{r})$, we obtain the dependence of the correlation function on the parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$. This illustrates how these parameters, which define the paired state, allow changing the correlation function.

Comment:

In the particular case where all the $\theta_{\mathbf{k}}$ are either zero or equal to $\pi/2$, we already mentioned (see end of § C-1-a) that the paired state becomes a Fock state in which the phases $\zeta_{\mathbf{k}}$ no longer play any role. It is easy to check that the anomalous average values are then all equal to zero, as is, obviously, the function $\phi_{\text{pair}}(\mathbf{r})$. On the other hand, for a different choice of the parameters $\theta_{\mathbf{k}}$, the phases $\zeta_{\mathbf{k}}$ play an especially important role, as will be shown for example in Complement C_{XVII}.

E. Paired states as a quasi-particle vacuum; Bogolubov-Valatin transformations

The Hamiltonian of a noninteracting particle system can be written as:

$$\widehat{H}_0 = \sum_i (\hbar\omega_i)\, a_i^\dagger a_i \tag{E-1}$$

where $\hbar\omega_i$ is the energy of an individual state labeled by the index i. The ground state $|\Phi_0\rangle$ of $\widehat{H}_0$ is an eigenvector of all the annihilation operators a_i, with a zero eigenvalue:

$$a_i\,|\Phi_0\rangle = 0 \tag{E-2}$$

The paired ket $|\Psi_{\text{paired}}\rangle$ is not an eigenvector of the usual annihilation operators a_i. We shall, however, introduce in § E-1 a linear transformation of the a_i and $a_i^\dagger$ into new annihilation and creation operators, and show in § E-2 that $|\Psi_{\text{paired}}\rangle$ is an eigenvector, with a zero eigenvalue, of all the new annihilation operators. The paired state will then appear as a "particle vacuum". Furthermore, in § E-3, we shall associate with $|\Psi_{\text{paired}}\rangle$ a family of operators having the same form as the Hamiltonian (E-1), but where the a_i and $a_i^\dagger$ are replaced by the new annihilation and creation operators. The interest of that association is the possibility, in certain cases (illustrated in the complements), to identify – with certain approximations if needed – an operator in this family with the Hamiltonian of a given physical situation. The problem of finding the ground state and the excited states is then solved, as if dealing with a system of independent particles. The state $|\Psi_{\text{paired}}\rangle$ can then be considered as the ground state of the Hamiltonian of independent "quasi-particles", while the new creation operators permit building a complete orthogonal basis of excited states.

E-1. Transformation of the creation and annihilation operators

For bosons in the same spin state, the state $|\varphi_{\mathbf{k}}\rangle$ belongs to the space $\mathcal{E}_{\mathbf{k}}$ associated with the pair $(\mathbf{k}, -\mathbf{k})$; this space is generated by the action of two creation operators $a_{\mathbf{k}}^\dagger$ and $a_{-\mathbf{k}}^\dagger$ on the vacuum. This is also the case for fermions in opposite spin states, if we simplify the notation $\mathbf{k},\uparrow$ to $\mathbf{k}$, as well as $-\mathbf{k},\downarrow$ to $-\mathbf{k}$ (we have labeled each pair of individual states by the value of $\mathbf{k}$ associated with the spin $\uparrow$). For both cases, we now define two new couples of creation and annihilation operators that act in $\mathcal{E}_{\mathbf{k}}$.

We introduce the two annihilation operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$, defined for $\mathbf{k} \neq 0$, as well as the Hermitian conjugate operators $b_{\mathbf{k}}^\dagger$ and $b_{-\mathbf{k}}^\dagger$, as:

$$\begin{aligned} b_{\mathbf{k}} &= u_{\mathbf{k}} a_{\mathbf{k}} + \eta\, v_{\mathbf{k}} a_{-\mathbf{k}}^\dagger & b_{\mathbf{k}}^\dagger &= u_{\mathbf{k}}^* a_{\mathbf{k}}^\dagger + \eta\, v_{\mathbf{k}}^* a_{-\mathbf{k}} \\ b_{-\mathbf{k}} &= u_{\mathbf{k}} a_{-\mathbf{k}} + v_{\mathbf{k}} a_{\mathbf{k}}^\dagger & b_{-\mathbf{k}}^\dagger &= u_{\mathbf{k}}^* a_{-\mathbf{k}}^\dagger + v_{\mathbf{k}}^* a_{\mathbf{k}} \end{aligned} \tag{E-3}$$

or:

$$\begin{pmatrix} b_{\mathbf{k}} \\ b_{-\mathbf{k}} \\ b_{\mathbf{k}}^\dagger \\ b_{-\mathbf{k}}^\dagger \end{pmatrix} = \begin{pmatrix} u_{\mathbf{k}} & & & \eta\, v_{\mathbf{k}} \\ & u_{\mathbf{k}} & v_{\mathbf{k}} & \\ & \eta\, v_{\mathbf{k}}^* & u_{\mathbf{k}}^* & \\ v_{\mathbf{k}}^* & & & u_{\mathbf{k}}^* \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}} \\ a_{-\mathbf{k}} \\ a_{\mathbf{k}}^\dagger \\ a_{-\mathbf{k}}^\dagger \end{pmatrix} \tag{E-4}$$

As for now, $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are any two complex numbers.

As in Chapters XV and XVI, $[A, B]_{-\eta}$ denotes the commutator of A and B if $\eta = 1$ (bosons), and their anticommutator if $\eta = -1$ (fermions). We now compute $[b_{\mathbf{k}}, b_{-\mathbf{k}}]_{-\eta}$; as $a_{\mathbf{k}}$ (anti)commutes with $a_{-\mathbf{k}}$ and as $a^{\dagger}_{-\mathbf{k}}$ (anti)commutes with $a^{\dagger}_{\mathbf{k}}$, only the cross terms in $u_{\mathbf{k}} v_{\mathbf{k}}$ remain:

$$[b_{\mathbf{k}}, b_{-\mathbf{k}}]_{-\eta} = u_{\mathbf{k}} v_{\mathbf{k}} \left\{ \left[a_{\mathbf{k}}, a^{\dagger}_{\mathbf{k}} \right]_{-\eta} + \eta \left[a^{\dagger}_{-\mathbf{k}}, a_{-\mathbf{k}} \right]_{-\eta} \right\} \tag{E-5}$$

For bosons, the commutator of $a_{-\mathbf{k}}$ and $a^{\dagger}_{-\mathbf{k}}$ equals 1, and hence the commutator of $a^{\dagger}_{-\mathbf{k}}$ and $a_{-\mathbf{k}}$ equals -1; for fermions, the two anticommutators of those operators are equal to 1, so that we obtain, in both cases:

$$\begin{aligned} [b_{\mathbf{k}}, b_{-\mathbf{k}}]_{-\eta} &= u_{\mathbf{k}} v_{\mathbf{k}} \{1 - 1\} \\ &= 0 \end{aligned} \tag{E-6}$$

By Hermitian conjugation, we get:

$$\left[b^{\dagger}_{\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta} = 0 \tag{E-7}$$

We now compute $\left[b_{\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta}$. This time, we get the two squared terms in $|u_{\mathbf{k}}|^2$ and $|v_{\mathbf{k}}|^2$. The one in $|u_{\mathbf{k}}|^2$ contains the (anti)commutator $\left[a_{\mathbf{k}}, a^{\dagger}_{\mathbf{k}} \right]$ equal to 1; the one in $|v_{\mathbf{k}}|^2$ contains, for bosons, the commutator of $a^{\dagger}_{-\mathbf{k}}$ and $a_{-\mathbf{k}}$ which equals -1, and for fermions, the anticommutator of those two operators which is equal to $+1$. We therefore get:

$$\left[b_{\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta} = |u_{\mathbf{k}}|^2 - \eta \, |v_{\mathbf{k}}|^2 \tag{E-8}$$

In a similar way:

$$\left[b_{-\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta} = |u_{\mathbf{k}}|^2 - \eta \, |v_{\mathbf{k}}|^2 \tag{E-9}$$

Finally, we are left with the computation of $\left[b_{\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta}$ and $\left[b_{-\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta}$. The first is zero since $a_{\mathbf{k}}$ (anti)commutes both with $a^{\dagger}_{-\mathbf{k}}$ and itself[8], and that $a^{\dagger}_{-\mathbf{k}}$ (anti)commutes with itself and with $a_{-\mathbf{k}}$; the reasoning is the same for the second, so that:

$$\begin{aligned} \left[b_{\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta} &= 0 \\ \left[b_{-\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta} &= 0 \end{aligned} \tag{E-10}$$

To sum up, it suffices to impose, for all values of $\mathbf{k}$, the condition:

$$|u_{\mathbf{k}}|^2 - \eta \, |v_{\mathbf{k}}|^2 = 1 \tag{E-11}$$

[8] For fermions, its square is identically zero.

to get the relations:

$$\begin{aligned}\left[b_{\mathbf{k}}, b_{\mathbf{k}}^{\dagger}\right]_{-\eta} &= 1 \\ \left[b_{-\mathbf{k}}, b_{-\mathbf{k}}^{\dagger}\right]_{-\eta} &= 1\end{aligned} \tag{E-12}$$

so that the operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$, as well as their adjoints, obey the same relations of (anti)commutation as the usual annihilation and creation operators of identical particles.

For fermions, we find again condition (C-3), which allows us to simply set, as in (C-4):

$$\begin{aligned}u_{\mathbf{k}} &= \cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}\end{aligned} \tag{E-13}$$

We then see that the matrix in the right hand side of (E-4) is unitary. This unitary transformation of the creation and annihilation operators is called the "Bogolubov-Valatin transformation".

For bosons, we will set:

$$\begin{aligned}u_{\mathbf{k}} &= \cosh\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sinh\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}\end{aligned} \tag{E-14}$$

Comparison with relation (C-11) shows that:

$$x_{\mathbf{k}} = \frac{v_{\mathbf{k}}}{u_{\mathbf{k}}} \tag{E-15}$$

The transformation of the creation and annihilation operators for bosons is called the "Bogolubov transformation".

E-2. Effect on the kets $|\overline{\varphi}_{\mathbf{k}}\rangle$

We now show that the vectors $|\overline{\varphi}_{\mathbf{k}}\rangle$ are eigenkets of two annihilation operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ with eigenvalues zero. This property makes them similar to a usual vacuum state, which yields zero under the action of all the annihilation operators $a_{\mathbf{k}}$.

E-2-a. Fermions in a singlet state

Let us compute the effect of those operators on the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ defined by relation (C-1), that we write with the simplified notation already used above ($\mathbf{k}$ is associated with the spin index + and $-\mathbf{k}$ with the spin index −):

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right]|0\rangle \tag{E-16}$$

We start with the operator $b_{\mathbf{k}}$ defined in (E-3). Its $u_{\mathbf{k}} a_{\mathbf{k}}$ term yields zero when acting on the term in $u_{\mathbf{k}}$ of $|\overline{\varphi}_{\mathbf{k}}\rangle$; only the term in $v_{\mathbf{k}}$ remains, for which the operator lowers from one to zero the occupation number of the state $\mathbf{k}, \uparrow$, since:

$$a_{\mathbf{k}} a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}|0\rangle = a_{-\mathbf{k}}^{\dagger}|0\rangle \tag{E-17}$$

As for the operator $v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}$, it yields zero when acting on the $v_{\mathbf{k}}$ term of $|\overline{\varphi}_{\mathbf{k}}\rangle$ (for fermions, the square of a creation operator is zero), leaving only the term in $u_{\mathbf{k}}$. This leads to:

$$b_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}}v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}} - v_{\mathbf{k}}u_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right]|0\rangle = 0 \tag{E-18}$$

The computation is the same for the operator $b_{-\mathbf{k}}$, except for the fact that the operator $a_{-\mathbf{k}}$ must first anticommute with $a^{\dagger}_{\mathbf{k}}$ before it can be regrouped with $a^{\dagger}_{-\mathbf{k}}$ and lower from one to zero the occupation number of the state $-\mathbf{k}$. The anticommutation therefore introduces a sign change, but as the definition of $b_{-\mathbf{k}}$ does not contain any $-$ sign, we again find:

$$b_{-\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = u_{\mathbf{k}}v_{\mathbf{k}}\left[-a^{\dagger}_{\mathbf{k}} + a^{\dagger}_{\mathbf{k}}\right]|0\rangle = 0 \tag{E-19}$$

We have shown that the two operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ have the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ as an eigenvector, with eigenvalue zero.

E-2-b. Bosons in the same spin state

Taking into account (E-15), relation (C-7) is written:

$$|\varphi_{\mathbf{k}}\rangle = \sum_{q=0}^{\infty}\frac{1}{q!}\left[-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right]^{q}\left[a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle \tag{E-20}$$

Since:

$$a_{\mathbf{k}}\left[a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle = \left(a_{\mathbf{k}}a^{\dagger}_{\mathbf{k}}\right)\left[a^{\dagger}_{\mathbf{k}}\right]^{q-1}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle = (q)\left[a^{\dagger}_{\mathbf{k}}\right]^{q-1}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle \tag{E-21}$$

we have:

$$u_{\mathbf{k}}a_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle = u_{\mathbf{k}}\sum_{q=0}^{\infty}\frac{q}{q!}\left[-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right]^{q}\left[a^{\dagger}_{\mathbf{k}}\right]^{q-1}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle \tag{E-22}$$

or else, since $a^{\dagger}_{-\mathbf{k}}$ commutes with all the operators in this expression:

$$\begin{aligned} u_{\mathbf{k}}a_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle &= -v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\sum_{q'=0}^{\infty}\frac{1}{q'!}\left[-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right]^{q'}\left[a^{\dagger}_{\mathbf{k}}\right]^{q'}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q'}|0\rangle \\ &= -v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle \end{aligned} \tag{E-23}$$

where we have set $q' = q - 1$. This leads to:

$$\left(u_{\mathbf{k}}a_{\mathbf{k}} + v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right)|\varphi_{\mathbf{k}}\rangle = 0 \tag{E-24}$$

which clearly shows that $|\varphi_{\mathbf{k}}\rangle$ is an eigenvector of the operator $b_{\mathbf{k}}$ defined in (E-3):

$$b_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle = 0 \tag{E-25}$$

The same computation leads to:

$$u_{\mathbf{k}}a_{-\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle = -v_{\mathbf{k}}a^{\dagger}_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle \tag{E-26}$$

and hence to:

$$b_{-\mathbf{k}} |\varphi_{\mathbf{k}}\rangle = 0 \tag{E-27}$$

As for fermions, the two operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ have the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ as an eigenvector with an eigenvalue zero.

E-3. Basis of excited states, quasi-particles

For bosons as for fermions, we just saw that the new creation and annihilation operators introduced in (E-3) and (E-4) have the same properties as the usual creation and annihilation operators. In particular, the two operators:

$$\widehat{n}\,(b_{\mathbf{k}}) = b_{\mathbf{k}}^{\dagger} b_{\mathbf{k}} \qquad\qquad \widehat{n}\,\left(b_{-\mathbf{k}}\right) = b_{-\mathbf{k}}^{\dagger} b_{-\mathbf{k}} \tag{E-28}$$

have as eigenvalues all the positive or zero integers, in perfect analogy with the operators corresponding to the population of individual states. By analogy with (E-1), it is therefore natural to introduce the operator:

$$\widehat{H}_B = \sum_{\mathbf{k}\in D} \hbar\omega_k \left[b_{\mathbf{k}}^{\dagger} b_{\mathbf{k}} + b_{-\mathbf{k}}^{\dagger} b_{-\mathbf{k}} \right] \tag{E-29}$$

where, for the moment, the ω_k are free parameters, as are the parameters which define the paired state (they will be fixed later on, depending on the physical problem we study). In relation (E-29), the summation is limited, as above, to a momentum half-space, which avoids taking opposite momenta into account twice. The eigenvalues of $\widehat{H}_B$ are all of the form:

$$E_B = \sum_{\mathbf{k}\in D} \left[n\,(b_{\mathbf{k}}) + n\,\left(b_{-\mathbf{k}}\right) \right] \hbar\omega_k \tag{E-30}$$

where $n\,(b_{\mathbf{k}})$ and $n\,(b_{-\mathbf{k}})$ are any positive or zero integers for bosons, and restricted to 0 or 1 for fermions.

The ground state $|\Phi_0(b)\rangle$ of $\widehat{H}_B$ is an eigenvector of all the annihilation operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ with eigenvalues zero. Now we saw in (B-8) for bosons, and in (B-11) for fermions, that the paired state vector is a tensor product of states $|\varphi_{\mathbf{k}}\rangle$, which are precisely the eigenvectors, with zero eigenvalues, of these two operators. The paired state, $|\Psi_{\text{paired}}\rangle$ for bosons[9], or $|\Psi_{\text{BCS}}\rangle$ for fermions, is thus an eigenvector of $\widehat{H}_B$ with a zero eigenvalue (ground state).

One can then obtain the other eigenstates of $\widehat{H}_B$ (excited states) by the action of the creation operators $b_{\mathbf{k}}^{\dagger}$ and $b_{-\mathbf{k}}^{\dagger}$ on $|\Phi_0(b)\rangle$. For bosons, each of these two operators will be able to act any number of times. For fermions, on the other hand, we shall only get 3 excited states, by the action of either $b_{\mathbf{k}}^{\dagger}$, or $b_{-\mathbf{k}}^{\dagger}$, or their product; as these operators anticommute, any higher power of those operators' product will yield zero. We finally note that operator (E-29) shares many of the properties of the Hamiltonian of an ensemble of particles without mutual interactions. Just as the usual creation operators

[9]For bosons, in Complement B_{XVII}, we will associate to that paired state a coherent state $|\varphi_{\mathbf{0}}\rangle$ to obtain the state $|\Phi_B\rangle$. But, as none of the operators $b_{\mathbf{k}}$ or $b_{-\mathbf{k}}$ act in the Fock space associated with the individual state $\mathbf{k} = 0$, the conclusions will be unchanged for $|\Phi_B\rangle$.

can add particles in a system of free identical particles, the creation operators $b^\dagger_{\mathbf{k}}$ and $b^\dagger_{-\mathbf{k}}$ can be considered as the operators adding a supplementary "quasi-particle" into the physical system. These quasi-particles are not the same as particles in a system really without interactions, as illustrated by the expression of these creation operators. They yield, however, a basis of states in which we can reason as if there were no interactions, which is a very powerful framework for reasoning in many domains of physics.

For the previous considerations to be relevant from a physical point of view, we have yet to show that the Hamiltonian of the problem we study can be approximated by an operator $\widehat{H}_B$, provided we make a judicious choice of all the parameters $u_{\mathbf{k}}$, $v_{\mathbf{k}}$ and ω_k. This is not a priori easy: the Hamiltonian of an ensemble of particles includes, in general, two-body interaction terms, and those are expressed in terms of sums of products of two creation operators $a^\dagger_{\mathbf{k}}$ and two annihilation operators $a_{\mathbf{k}}$, hence of 4 operators. Now, if we insert definitions (E-3) and (E-4) into (E-29) to express $\widehat{H}_B$ as a function of the old creation and annihilation operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$, it is clear that we shall only obtain combinations of products of 2 operators. We shall need to make certain approximations to be able to consider $\widehat{H}_B$ as a physically pertinent approximate Hamiltonian. Examples of such situations will be given in the complements.

Conclusion

In conclusion, the paired states are a powerful tool for studying both fermions and bosons. They provide a systematic method allowing a certain flexibility in variational calculations in the presence of interactions. Furthermore, starting from a paired ground state, we were able to build a whole basis of excited states using creation and annihilation operators matching that ground state. In the complements of this chapter, we shall use the paired states to study different problems and compute the optimal parameters most relevant for each situation. The physical results will be quite different, depending on the cases, especially for fermions or for bosons; but the main point remains that the paired states offer a unified framework for obtaining all these different results.

COMPLEMENTS OF CHAPTER XVII, READER'S GUIDE

The first two complements provide more details about a number of results given in the chapter, concerning various properties of the pair operators and the paired states. The following three complements apply these concepts to physical phenomena involving fermions, and then bosons.

A_{XVII} : PAIR FIELD OPERATOR FOR IDENTICAL PARTICLES	The pair field operator is the analog, for a pair of particles, of the usual field operator for a single particle. It is a useful tool for computing average values in a paired state. The commutation relations of fermion pair operators are similar to those of bosons, except for an additional term due to the fermionic character of the pair constituents.
B_{XVII} : AVERAGE ENERGY IN A PAIRED STATE	This complement explains the computation of the average energy in a paired state. For bosons, we add to this paired state a condensate, described by a coherent state. The results of this complement are used in Complements C_{XVII} and E_{XVII} .
C_{XVII} : FERMION PAIRING, BCS THEORY	Even weak attractive interactions can greatly modify the ground state of a fermion system, via the BCS mechanism for pair formation. This complement discusses the theory of this phenomena, and its effect on the particle distribution and correlation functions, as well as its link to Bose-Einstein condensation of pairs of particles.
D_{XVII} : COOPER PAIRS	The simple Cooper model studies the bound states of two weakly attracted particles, in the presence of a Fermi sphere that prevents the particles from occupying states inside that sphere. Whereas, in general, a minimum depth of an attractive potential is required for two particles to form a bound state in 3-D, the presence of the Fermi sphere ensures the existence of a bound state, no matter how weak the attraction is. The Cooper model accounts in a somewhat intuitive way for a number of results of the BCS theory.
E_{XVII} : CONDENSED REPULSIVE BOSONS	For an ensemble of bosons, using paired states as variational states leads to the same results as the Bogolubov method based on operator transformations. We thus obtain the Bogolubov spectrum, compute the "quantum depletion" introduced by the interactions, etc.

Complement A_{XVII}

Pair field operator for identical particles

1 Pair creation and annihilation operators 1846

1-a Particles in the same spin state 1846

1-b Pairs in a singlet spin state 1849

2 Average values in a paired state 1851

2-a Average value of a field operator; pair wave function, and order parameter 1851

2-b Average value of a product of two field operators; factorization of the order parameter 1854

2-c Application to the computation of the correlation function (singlet pairs) 1858

3 Commutation relations of field operators 1861

3-a Particles in the same spin state 1861

3-b Singlet pairs 1866

In Chapter XVI we introduced a field operator $\Psi(\mathbf{r})$ acting in the state space of a system of identical particles. This operator was defined as a linear combination of annihilation operators associated with individual states having a given momentum. It proved to be a useful tool for various computations, and in particular for the determination of correlation functions. We then showed, in Chapter XVII, the relevance of paired states where, essentially, identical particles were grouped into pairs. We introduced creation and annihilation operators of pairs of particles in well defined momentum states, $A^{\dagger}_{\mathbf{K}}$ and $A_{\mathbf{K}}$. Consequently, it is natural to envisage the introduction of a field operator for pairs of particles, which will be the operator $\Phi_{\chi}(\mathbf{R})$ destroying a pair of particles whose center of mass is at point $\mathbf{R}$ and whose internal state is described by the wave function χ. Its adjoint, $\Phi^{\dagger}_{\chi}(\mathbf{R})$, creates a pair of particles in that same state. In this complement, we will define these operators and study some of their properties.

We start in § 1 by giving the expression of these field operators $\Phi_{\chi_n}(\mathbf{R})$ and $\Phi^{\dagger}_{\chi_n}(\mathbf{R})$ for pairs described by any orbital state χ_n. We consider the case where the particles are either in the same spin state, or in a spin singlet state. We then study, in § 2, the average values, in paired states, of pair field operators and of products of such operators. These average values have some very interesting properties leading us, in particular, to introduce a new wave function $\phi_{\text{pair}}(\mathbf{r})$, called the "pair wave function", which is not simply the two-particle wave function $\chi_{\text{pair}}(\mathbf{r})$ used to build the paired state. As we shall see in § 2-c, this new wave function explicitly appears in the binary correlation function of the particles' positions. Moreover, its norm is linked to the number of quanta present in the field of condensed pairs. The origin of this pair function is the fact that pairs can collectively contribute to the creation of a field whose average value is what we shall call an "order parameter". This non-zero order parameter indicates the existence of a macroscopic field associated with the pairs. We will show how it relates the "anomalous average values" (of operators that do not conserve particle number) to

the normal average values of a product of two field operators $\Phi_{\chi_n}(\mathbf{R})$ and $\Phi^\dagger_{\chi_n}(\mathbf{R})$, that does conserve particle number. In particular, we shall use, in § 2-c, the properties of the pair field operator to get the correlation functions in a paired BCS state, and to study the consequences of the existence of the macroscopic field associated with the pairs. We shall finally study, in § 3, the commutation properties of these operators; they will be found to be similar to those of bosons (whether the particles building the pair are bosons or fermions), but not completely identical as corrective terms must be added to the boson commutator. We shall see that, since the pairs are strongly bound and have a spatial extension much smaller than all the characteristic dimensions of the problem, the pairs can be assimilated to bosons; if, however, the pairs are weakly bound (as is the case, in particular, for the BCS mechanism we will discuss in Complement C_{XVII}, it is not possible to consider them as indivisible entities: the fermionic structure of their components plays an important role that cannot be ignored.

1. Pair creation and annihilation operators

By analogy with the field operator for the particles composing the pairs, we now introduce a field operator concerning the pairs themselves. The adjoint of this field operator allows the direct creation of a pair of particles at a given point, and with a given internal state; as for the operator itself, it annihilates that same pair.

1-a. Particles in the same spin state

We defined, in Chapter XVII, the operators $A^\dagger_{\mathbf{K}}$ and $A_{\mathbf{K}}$ for pairs of particles without spin (or in the same spin state), as:

$$A^\dagger_{\mathbf{K}} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}} g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}}$$
$$A_{\mathbf{K}} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}} g^*_{\mathbf{k}}\ a_{\frac{\mathbf{K}}{2}-\mathbf{k}} a_{\frac{\mathbf{K}}{2}+\mathbf{k}} \tag{1}$$

In this expression, $g_{\mathbf{k}}$ is the Fourier transform of the wave function $\chi(\mathbf{r})$ characterizing the pair:

$$g_{\mathbf{k}} = \frac{1}{L^{3/2}} \int_{L^3} d^3r\ e^{-i\mathbf{k}\cdot\mathbf{r}} \chi(\mathbf{r}) \tag{2}$$

and L is the edge length of a cube, of volume L^3 , which contains the physical system. We now generalize these definitions to the case where the pair is not necessarily in a given orbital state, but in any state belonging to an orthonormal basis of states $|\chi_n\rangle$, with the index n going from 1 to infinity; these states each have a wave function $\chi_n(\mathbf{r})$ whose Fourier transform is $g^n_{\mathbf{k}}$. We therefore simply add an index n to the previous definitions, as for example:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}} g^n_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{3}$$

We saw in Chapter XVII that the bosonic or fermionic character of the particles building the pair requires the functions $g^n_{\mathbf{k}}$ to have the parity η with respect to the

variable $\mathbf{k}$, which means $g^n_{-\mathbf{k}} = \eta\, g^n_{\mathbf{k}}$ with:

$$\eta = \begin{cases} +1 & \text{for bosons} \\ -1 & \text{for fermions} \end{cases} \tag{4}$$

Should $g^n_{\mathbf{k}}$ be of parity $-\eta$, it is easy to check by (anti)commutation of the operators that expression (3) yields zero. We can then only consider the case where the $g^n_{\mathbf{k}}$, and hence the corresponding wave functions $\chi_n(\mathbf{r})$, have the parity η. If, however, we need the basis of states $|\chi_n\rangle$ associated with these wave functions to be complete, we can include states of any parity, ± 1. We must then remember that the operators $A^\dagger_{\mathbf{K},n}$ are zero whenever the index n corresponds to a wave function of parity $-\eta$.

α. *Expression of $A^\dagger_{\mathbf{K},n}$ in terms of the particle field operator*

Relation (A-10) in Chapter XVI permits replacing the creation operators $a^\dagger_{\mathbf{k}}$ by:

$$a^\dagger_{\mathbf{k}} = \frac{1}{\sqrt{L^3}} \int \mathrm{d}^3 r\; e^{i\mathbf{k}\cdot\mathbf{r}}\; \Psi^\dagger(\mathbf{r}) \tag{5}$$

where $\Psi^\dagger(\mathbf{r})$ is the adjoint of the field operator associated with the elementary components of the pair (the "atoms" of each "molecule"). Using twice this relation in (3), we get:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2}L^3} \sum_{\mathbf{k}}\; g^n_{\mathbf{k}}\; \int \mathrm{d}^3 r \int \mathrm{d}^3 r'\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}'}\; \Psi^\dagger(\mathbf{r})\Psi^\dagger(\mathbf{r}') \tag{6}$$

or else, choosing as the integration variables $\mathbf{R} = (\mathbf{r}+\mathbf{r}')/2$ and $\mathbf{x} = \mathbf{r}-\mathbf{r}'$:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2}L^3} \int \mathrm{d}^3 R\; e^{i\mathbf{K}\cdot\mathbf{R}} \int \mathrm{d}^3 x \sum_{\mathbf{k}}\; g^n_{\mathbf{k}}\; e^{i\mathbf{k}\cdot\mathbf{x}}\; \Psi^\dagger(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{7}$$

The summation over $\mathbf{k}$ on the right-hand side leads to expression (A-2) of Chapter XVII for the wave function χ_n, and we can write:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2L^3}} \int \mathrm{d}^3 R\; e^{i\mathbf{K}\cdot\mathbf{R}} \int \mathrm{d}^3 x\; \chi_n(\mathbf{x})\; \Psi^\dagger(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{8}$$

This other form for the operator already introduced in (3) demonstrates the fact that it creates a pair of particles in a molecular state characterized, for its external variables, by a plane wave of wave vector $\mathbf{K}$, and for its internal variables, by the wave function χ_n .

β. *Pair field*

For each internal state $|\chi_n\rangle$ of the pair, we can introduce, using relation (A-3) of Chapter XVI, an operator $\Phi^\dagger_{\chi_n}(\mathbf{R})$ that creates a pair at point $\mathbf{R}$ and in the internal state χ_n:

$$\Phi^\dagger_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{L^3}} \sum_{\mathbf{K}}\; e^{-i\mathbf{K}\cdot\mathbf{R}}\; A^\dagger_{\mathbf{K},n} \tag{9}$$

Replacing in (8) the integral variable $\mathbf{R}$ by $\mathbf{R}'$, and using the result in equality (9), we get:

$$\Phi^{\dagger}_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{2}L^3}\sum_{\mathbf{K}}\int \mathrm{d}^3R'\ e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}\int \mathrm{d}^3x\ \chi_n(\mathbf{x})\ \Psi^{\dagger}(\mathbf{R}'+\frac{\mathbf{x}}{2})\Psi^{\dagger}(\mathbf{R}'-\frac{\mathbf{x}}{2}) \quad (10)$$

The sum over $\mathbf{K}$ of $e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}$ then yields $L^3\delta(\mathbf{R}-\mathbf{R}')$, which allows integrating over d^3R', and we obtain:

$$\Phi^{\dagger}_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{2}}\int \mathrm{d}^3x\ \chi_n(\mathbf{x})\ \Psi^{\dagger}(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^{\dagger}(\mathbf{R}-\frac{\mathbf{x}}{2}) \quad (11)$$

This operator is therefore a product of field operators creating successively each of the two elements of the pair, which is easy to understand from a physical point of view. Note, however, that the two elements are not created at the same point, but symmetrically with respect to point $\mathbf{R}$, and with a spatial distribution whose amplitude is given by the wave function $\chi(\mathbf{x})$ of the "molecule". The spatial zone involved in the process thus extends over a distance of the order of the range of this wave function.

As for the pair field operator itself, which annihilates a pair, it is defined by Hermitian conjugation of the previous relation:

$$\Phi_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{2}}\int \mathrm{d}^3x\ \chi^*_n(\mathbf{x})\ \Psi(\mathbf{R}-\frac{\mathbf{x}}{2})\Psi(\mathbf{R}+\frac{\mathbf{x}}{2}) \quad (12)$$

We now use relation (A-3) of Chapter XVI to come back to the annihilation operators in a basis of individual states with fixed momenta. Using (twice) this relation in (12), we get:

$$\Phi_{\chi_n}(\mathbf{R}) = \frac{1}{L^3\sqrt{2}}\int \mathrm{d}^3x\ \chi^*_n(\mathbf{x})\ \sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i\mathbf{k}_1\cdot\left(\mathbf{R}-\frac{\mathbf{x}}{2}\right)}e^{i\mathbf{k}_2\cdot\left(\mathbf{R}+\frac{\mathbf{x}}{2}\right)}a_{\mathbf{k}_1}a_{\mathbf{k}_2} \quad (13)$$

This relation will be useful in what follows.

γ. *Inversion; expression for the interaction energy*

We call $|\chi_n\rangle$ the individual states corresponding to the wave functions $\chi_n(\mathbf{r})$ and assume they form a complete basis. The closure relation on these states is written:

$$\sum_n \langle\mathbf{k}\,|\chi_n\rangle\,\langle\chi_n\,|\mathbf{k}'\rangle = \sum_n (g^n_{\mathbf{k}})\,(g^n_{\mathbf{k}'})^* = \delta_{\mathbf{k},\mathbf{k}'} \quad (14)$$

We now multiply relation (3) by $(g^n_{\mathbf{k}'})^*$, and sum over n to get:

$$\sum_n (g^n_{\mathbf{k}'})^*\,A^{\dagger}_{\mathbf{K},n} = \frac{1}{\sqrt{2}}\ a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k}'}a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k}'} \quad (15)$$

It is thus possible to invert relations (3) and express any two creation operators as a sum of pair creation operators, according to:

$$a^{\dagger}_{\mathbf{k}_1}a^{\dagger}_{\mathbf{k}_2} = \sqrt{2}\sum_n \left(g^n_{(\mathbf{k}_1-\mathbf{k}_2)/2}\right)^* A^{\dagger}_{\mathbf{k}_1+\mathbf{k}_2,n} \quad (16)$$

where we have replaced $\mathbf{K}$ par $\mathbf{k}_1+\mathbf{k}_2$ and $\mathbf{k}'$ by $(\mathbf{k}_1-\mathbf{k}_2)/2$. By Hermitian conjugation, we get a similar relation for any product $a_{\mathbf{k}_3}a_{\mathbf{k}_4}$ of annihilation operators.

Interaction Hamiltonian:

Any operator $\widehat{W}_{\rm int}$ for the binary interactions between particles can therefore be written as:

$$\widehat{W}_{\rm int} = \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4,n,n'} \langle 1:\mathbf{k}_1;2:\mathbf{k}_2|\, W_2\,(1,2)\,|1:\mathbf{k}_3;2:\mathbf{k}_4\rangle \left(g^n_{(\mathbf{k}_1-\mathbf{k}_2)/2}\right)^* \left(g^{n'}_{(\mathbf{k}_3-\mathbf{k}_4)/2}\right) A^\dagger_{\mathbf{K},n}\; A_{\mathbf{K},n'} \tag{17}$$

where $W_2\,(1,2)$ is the binary interaction between particles (as, for example, in Complement $\mathrm{E_{XV}}$; using momentum conservation, we have set:

$$\mathbf{K} = \mathbf{k}_1+\mathbf{k}_2 = \mathbf{k}_3+\mathbf{k}_4 \tag{18}$$

Written in terms of pair creation and annihilation operators, $\widehat{W}_{\rm int}$ is the sum of quadratic terms, and no longer of fourth degree terms as was the case with operators for individual particles. Note, however, that one must be careful when using relation (17) since, as we shall see in § 3, the pair creation and annihilation operators do not obey the usual commutation relations. The action of an operator $A_{\mathbf{K},n}$ on a paired state obtained by the action of $\left[A^\dagger_{\mathbf{K}',n}\right]^N$ on the vacuum, does not necessarily yield zero when $\mathbf{K}\neq\mathbf{K}'$. Pair creation operators are not as simple to handle as particle creation operators.

1-b. Pairs in a singlet spin state

For a pair of spin 1/2 particles in a singlet spin state, we use relation (A-23) of Chapter XVII and add an index n to represent the internal orbital state of the pair; this reads:

$$A^\dagger_{\mathbf{K},n} = \sum_{\mathbf{k}} g^n_{\mathbf{k}}\; a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} \tag{19}$$

The following computations apply directly to fermions in a singlet state, for which the functions $g^n_{\mathbf{k}}$ must be even with respect to the variable $\mathbf{k}$. We noted however in Chapter XVII, in comment (ii) just before § B, that they can also apply to fermions in a triplet spin state, when the function $g^n_{\mathbf{k}}$ is odd; even though this case can be included in the following discussion, for the sake of simplicity we will continue to talk about singlet pairs.

α. Expression of $A^\dagger_{\mathbf{K},n}$ in terms of the particle field operator

Relation (A-9) of Chapter XVI becomes here, taking into account the spin indices:

$$a^\dagger_{\mathbf{k},\nu} = \frac{1}{\sqrt{L^3}} \int \mathrm{d}^3r\; e^{i\mathbf{k}\cdot\mathbf{r}}\; \Psi^\dagger_\nu(\mathbf{r}) \tag{20}$$

Inserting this equality in (19) yields:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{L^3}\sum_{\mathbf{k}}\; g^n_{\mathbf{k}}\; \int \mathrm{d}^3r \int \mathrm{d}^3r'\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}'}\; \Psi^\dagger_\uparrow(\mathbf{r})\Psi^\dagger_\downarrow(\mathbf{r}') \tag{21}$$

As previously, the wave function $\chi_n(\mathbf{x})$ appears when we use as integral variables $\mathbf{R} = (\mathbf{r} + \mathbf{r}')/2$ and $\mathbf{x} = \mathbf{r} - \mathbf{r}'$, and we get:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{L^3}} \int \mathrm{d}^3R\, e^{i\mathbf{K}\cdot\mathbf{R}} \int \mathrm{d}^3x\, \chi_n(\mathbf{x})\ \Psi^\dagger_\uparrow(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger_\downarrow(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{22}$$

This yields the form of the operator creating a pair of particles in a singlet molecular state, characterized by a plane wave of wave vector $\mathbf{K}$ for its external variables, and by the wave function χ_n for its internal variables.

β. Pair field

We now insert relation (22) in (9); we get:

$$\Phi^\dagger_{\chi_n}(\mathbf{R}) = \frac{1}{L^3} \sum_{\mathbf{K}} \int \mathrm{d}^3R'\, e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})} \int \mathrm{d}^3x\, \chi_n(\mathbf{x})\ \Psi^\dagger_\uparrow(\mathbf{R}'+\frac{\mathbf{x}}{2})\Psi^\dagger_\downarrow(\mathbf{R}'-\frac{\mathbf{x}}{2}) \tag{23}$$

As before, the sum over $\mathbf{K}$ of $e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}$ yields $L^3\delta(\mathbf{R}-\mathbf{R}')$, and we get:

$$\Phi^\dagger_{\chi_n}(\mathbf{R}) = \int \mathrm{d}^3x\, \chi_n(\mathbf{x})\ \Psi^\dagger_\uparrow(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger_\downarrow(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{24}$$

The same comments as in § 1-a-β can be made: this operator successively creates the two elements of the pair at different points, with a probability amplitude given by the internal wave function $\chi_n(\mathbf{x})$ of the distance between these points. The field operator is obtained by Hermitian conjugation:

$$\Phi_{\chi_n}(\mathbf{R}) = \int \mathrm{d}^3x\, \chi^*_n(\mathbf{x})\ \Psi_\downarrow(\mathbf{R}-\frac{\mathbf{x}}{2})\Psi_\uparrow(\mathbf{R}+\frac{\mathbf{x}}{2}) \tag{25}$$

It will often be convenient to come back and use the annihilation operators in a basis of individual states of fixed momenta. Using (twice) relation (A-14) of Chapter XVI, we get:

$$\Phi_{\chi_n}(\mathbf{R}) = \frac{1}{L^3} \int \mathrm{d}^3x\, \chi^*_n(\mathbf{x}) \sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i\mathbf{k}_1\cdot(\mathbf{R}-\frac{\mathbf{x}}{2})} e^{i\mathbf{k}_2\cdot(\mathbf{R}+\frac{\mathbf{x}}{2})} a_{\mathbf{k}_1,\downarrow} a_{\mathbf{k}_2,\uparrow} \tag{26}$$

Comment:

For singlet pairs, we could invert those relations, as we did before, and express the interaction energy in terms of the pair creation and annihilation operators. It is, however, a bit more complicated in this case than when the pairs were in the same spin state: as we shall see in § 2-c-α, it would be necessary to involve another pair creation operator (in a triplet state). This would lead to cumbersome notation, and the computation will not be presented here.

2. Average values in a paired state

We now compute the average values of pair field operators, or of products of such operators, in the paired state we defined Chapter XVII. We shall use relations (13) or (26), depending on whether the particles are in the same spin state, or in a singlet spin state. In both cases, the computation of the average value of those operators in a paired state involves the computation of average values of products of annihilation operators – i.e. of "anomalous" average values as defined in Chapter XVII.

2-a. Average value of a field operator; pair wave function, and order parameter

Expressions for the paired kets were obtained in § B-2 of Chapter XVII as tensor products of states of pairs that are not eigenstates of the occupation number operators. These pairs all have a zero total momentum; we therefore assume, from now on, that $\mathbf{K} = 0$. The average value computation of a pair field operator in these states will lead to a new wave function, that we will call the "pair wave function".

α. Particles in the same spin state

Relations (B-8) and (B-9) of Chapter XVII give the expression of the paired state vector $|\Psi_{\text{paired}}\rangle$ for an ensemble of a large number of particles:

$$|\Psi_{\text{paired}}\rangle = \prod_{\otimes \mathbf{k}\in\mathbf{D}} \exp\left\{\sqrt{2}\; g_{\mathbf{k}}\; a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right\} |0\rangle \tag{27}$$

The function $g_{\mathbf{k}}$ used to build this paired state is a priori totally independent of the functions $g^n_{\mathbf{k}}$ defining the pair field operators. In such paired states, the populations of the states of the same pair are always equal; consequently, the only non-zero average values $\langle a_{\mathbf{k}_1} a_{\mathbf{k}_2}\rangle$ are those in which the two annihilation operators act on the two states of the same pair, which have opposite momenta. As the total momenta of each pair is zero, we can set $\mathbf{k}_1 = -\mathbf{k}_2$ in (13) and obtain:

$$\begin{aligned}\left\langle \Phi_{\chi_n}(\mathbf{R})\right\rangle &= \frac{1}{L^3\sqrt{2}} \int \mathrm{d}^3x\; \chi^*_n(\mathbf{x}) \sum_{\mathbf{k}_1} e^{-i\mathbf{k}_1\cdot\mathbf{x}} \langle a_{\mathbf{k}_1} a_{-\mathbf{k}_1}\rangle \\ &= \int \mathrm{d}^3x\; \chi^*_n(\mathbf{x})\; \phi_{\text{pair}}(\mathbf{x})\end{aligned} \tag{28}$$

where the (non normalized) "pair wave function" has already been defined in (D-6) of Chapter XVII:

$$\boxed{\phi_{\text{pair}}(\mathbf{x}) = \langle \mathbf{x}\,|\phi_{\text{pair}}\rangle = \frac{1}{L^3\sqrt{2}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \langle a_{\mathbf{k}} a_{-\mathbf{k}}\rangle} \tag{29}$$

Changing the sign of the summation variable $\mathbf{k}$, allows writing the pair wave function $\bar{\phi}_{\text{pair}}(\mathbf{k})$ in the momentum representation as:

$$\bar{\phi}_{\text{pair}}(\mathbf{k}) = \langle \mathbf{k}\,|\phi_{\text{pair}}\rangle = \frac{1}{L^{3/2}\sqrt{2}} \langle a_{-\mathbf{k}} a_{\mathbf{k}}\rangle \tag{30}$$

Note that because of the condition $\mathbf{k}_1 = -\mathbf{k}_2$ (the total momentum of each pair is zero), the average value $\left\langle \Phi_{\chi_n}(\mathbf{R}) \right\rangle$ no longer depends on $\mathbf{R}$.

The average value of the pair field operator is thus:

$$\left\langle \Phi_{\chi_n}(\mathbf{R}) \right\rangle = \left\langle \Phi_{\chi_n} \right\rangle = \langle \chi_n \, | \phi_{\text{pair}} \rangle \tag{31}$$

As expected from the translation invariance of the system, it is independent of $\mathbf{R}$. On the other hand, it depends on the internal state $|\chi_n\rangle$, and reaches a maximum when $|\chi_n\rangle$ is equal to the normalized state $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$ proportional to $|\phi_{\text{pair}}\rangle$:

$$\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle = \frac{|\phi_{\text{pair}}\rangle}{\sqrt{\langle \phi_{\text{pair}} \, | \phi_{\text{pair}} \rangle}} \tag{32}$$

This computation therefore leads to a new state $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$, different from the state $|\chi\rangle$ that was used in Chapter XVII to build the paired state $|\Psi_{\text{paired}}\rangle$. Choosing for the first vector of the basis $|\chi_1\rangle = \left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$, the average value of the field is given by:

$$\boxed{\left\langle \Phi_{\chi_1}(\mathbf{R}) \right\rangle = \left\langle \Phi_{\chi_1} \right\rangle = \sqrt{\langle \phi_{\text{pair}} \, | \phi_{\text{pair}} \rangle}} \tag{33}$$

This average value $\left\langle \Phi_{\chi_1}(\mathbf{R}) \right\rangle$ is often called the "order parameter of the pairs"; its non-zero value is important as it indicates the existence of a field constructed collectively by the pairs. In the present case, the average value of this field is independent of $\mathbf{R}$, as the paired state was built from pairs having a total momentum $\mathbf{K} = 0$ and whose center of mass has a constant wave function.

The average values $\langle a_{\mathbf{k}} a_{-\mathbf{k}} \rangle$, which according to (29) determine the pair wave function, have been called, in § C-3 of Chapter XVII, "anomalous average values", as they involve operators that do not conserve particle number. For bosons in the same spin state, relation (C-52) of that chapter indicates that:

$$\langle a_{\mathbf{k}} a_{-\mathbf{k}} \rangle = -e^{2i\zeta_{\mathbf{k}}} \, \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \tag{34}$$

One may wonder, of course, what the purpose of computing an anomalous average value is, as it can only be zero in a state with a fixed total particle number. We shall see, however, in § 2-b that these anomalous average values are a useful tool for computing average values of operators that do conserve the total number of particles and hence have a direct physical interpretation.

For bosons, operators $a_{\mathbf{k}}$ and $a_{-\mathbf{k}}$ commute, and hence the definition (29) shows that the wave function $\phi_{\text{pair}}(\mathbf{x})$ is even:

$$\phi_{\text{pair}}(-\mathbf{x}) = \phi_{\text{pair}}(\mathbf{x}) \tag{35}$$

The field mean value (31) is thus zero for any state χ_n of the basis whose wave function is odd: the postulate of symmetrization with respect to the pair components requires that pair to be in an even orbital state[1].

[1] If the particles composing the pair were fermions in the same spin state, the conclusions would be opposite. The wave function would be odd (because of the anticommutation of the operators $a_{\mathbf{k}}$ and $a_{-\mathbf{k}}$); the average values for even internal states χ_n would be zero.

β. Singlet pairs

For fermions in a singlet state, the paired state vector $|\Psi_{\rm BCS}\rangle$ is given by relations (B-11) and (B-12) of Chapter XVII:

$$|\Psi_{\rm BCS}\rangle = \prod_{\otimes \mathbf{k}} \exp\left[g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{36}$$

Following (26), we must add the spin index $\downarrow$ to the state $\mathbf{k}$, and the spin index $\uparrow$ to the state $-\mathbf{k}$. Here again, the average value of the product of annihilation operators is different from zero only if their wave vectors are opposite, and relation (26) then leads to:

$$\begin{aligned} \left\langle \Phi_{\chi_n}(\mathbf{R})\right\rangle &= \frac{1}{L^3}\int \mathrm{d}^3x\ \chi_n^*(\mathbf{x})\ \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \langle a_{\mathbf{k},\downarrow} a_{-\mathbf{k},\uparrow}\rangle \\ &= \langle \chi_n\ |\phi_{\rm pair}\rangle \end{aligned} \tag{37}$$

with the definition (D-14) of Chapter XVII of the state $|\phi_{\rm pair}\rangle$, associated with the (non-normalized) wave function:

$$\phi_{\rm pair}(\mathbf{x}) = \langle \mathbf{x}\ |\phi_{\rm pair}\rangle = \frac{1}{L^3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \langle a_{\mathbf{k},\downarrow} a_{-\mathbf{k},\uparrow}\rangle \tag{38}$$

In a similar way, the pair wave function in the momentum representation $\bar{\phi}_{\rm pair}(\mathbf{k})$ can be defined as:

$$\bar{\phi}_{\rm pair}(\mathbf{k}) = \frac{1}{L^{3/2}} \langle a_{-\mathbf{k},\downarrow} a_{\mathbf{k},\uparrow}\rangle \tag{39}$$

This wave function can be interpreted in the same way as the wave function defining the orbital variables of a pair in the singlet state. As in (32), we define the normalized ket $\left|\phi^{\rm norm}_{\rm pair}\right\rangle$. The field average value (37) is zero if $|\chi_n\rangle$ is orthogonal to $\left|\phi^{\rm norm}_{\rm pair}\right\rangle$ and reaches a maximum for $|\chi_n\rangle = |\chi_1\rangle = \left|\phi^{\rm norm}_{\rm pair}\right\rangle$; this maximum is equal to:

$$\boxed{\left\langle \Phi_{\chi_1}(\mathbf{R})\right\rangle = \left\langle \Phi_{\chi_1}\right\rangle = \sqrt{\langle \phi_{\rm pair}\ |\phi_{\rm pair}\rangle}} \tag{40}$$

and defines the order parameter of the physical system. It indicates the presence of a field created collectively by the pairs. As noted before, since the total momentum of each pair is zero, this average value does not depend on $\mathbf{R}$.

The average values that come into play in that definition are given by relation (C-42) of Chapter XVII:

$$\langle a_{\mathbf{k},\downarrow} a_{-\mathbf{k},\uparrow}\rangle = u^*_{\mathbf{k}} v_{\mathbf{k}} = \sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \tag{41}$$

(in the BCS state, the $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even functions of $\mathbf{k}$). We noted, at the end of § C-1-a of that chapter that, in the specific case where the $\theta_{\mathbf{k}}$ are either zero or equal to $\pi/2$, the paired ket is simply a Fock state of individual particles, hence a ket without pairing. Since (41) is then equal to zero, we see that the pair wave function is zero in the absence of pairing.

2-b. Average value of a product of two field operators; factorization of the order parameter

The field operators do not conserve particle number, as opposed to the usual operators such as the Hamiltonian, the total momentum, the double density, etc. On the other hand, the product of operators $\Phi^{\dagger}_{\chi_n}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')$ does conserve that number, and may help characterizing the properties of the pairs while being easier to interpret from a physical point of view.

α. Particles in the same spin state

Using relation (13) we get:

$$\left\langle \Phi^{\dagger}_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle = \frac{1}{2L^6}\int \mathrm{d}^3x\ \chi_n(\mathbf{x})\int \mathrm{d}^3x'\ \chi^*_{n'}(\mathbf{x}') \times \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} e^{-i\mathbf{k}_1\cdot\left(\mathbf{R}+\frac{\mathbf{x}}{2}\right)} e^{-i\mathbf{k}_2\cdot\left(\mathbf{R}-\frac{\mathbf{x}}{2}\right)} e^{i\mathbf{k}_3\cdot\left(\mathbf{R}'-\frac{\mathbf{x}'}{2}\right)} e^{i\mathbf{k}_4\cdot\left(\mathbf{R}'+\frac{\mathbf{x}'}{2}\right)} \left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle \tag{42}$$

The integrals over d^3x and d^3x' yield Fourier transforms $g^n_{\mathbf{k}}$ of the wave functions $\chi_n(\mathbf{x})$:

$$g^n_{\mathbf{k}} = \frac{1}{L^{3/2}}\int \mathrm{d}^3x\ e^{-i\mathbf{k}\cdot\mathbf{x}}\ \chi_n(\mathbf{x}) \tag{43}$$

and we get:

$$\left\langle \Phi^{\dagger}_{\chi_n}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle = \frac{1}{2L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} g_n\left(\frac{\mathbf{k}_1-\mathbf{k}_2}{2}\right)\ g^*_{n'}\left(\frac{\mathbf{k}_4-\mathbf{k}_3}{2}\right) \times\ e^{i\left[(\mathbf{k}_3+\mathbf{k}_4)\mathbf{R}'-(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{R}\right]\cdot} \left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle \tag{44}$$

where, to simplify the notation, we have written $g_n(\mathbf{k})$ the Fourier transform of $g^n_{\mathbf{k}}$.

Computation of the average value $\left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle$

This computation follows the same steps as the one in § D of Chapter XVII for the correlation between particles, as well as the one in § 3-a-β of Complement B_{XVII} for the interaction energy. Three cases must be distinguished:

– (**I**) The "forward scattering" terms are obtained either for $\mathbf{k}_4=\mathbf{k}_1$ and $\mathbf{k}_3=\mathbf{k}_2$ (direct terms), or $\mathbf{k}_3=\mathbf{k}_1$ and $\mathbf{k}_4=\mathbf{k}_2$ (exchange terms). We assume these forward scattering terms concern two different pairs, meaning $\mathbf{k}_1\neq\pm\mathbf{k}_2$. Since:

$$\left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_2} a_{\mathbf{k}_1}\right\rangle = \eta\left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_1} a_{\mathbf{k}_2}\right\rangle = \langle\widehat{n}_{\mathbf{k}_1}\rangle\,\langle\widehat{n}_{\mathbf{k}_2}\rangle \tag{45}$$

their sum yields the contribution:

$$\left\langle \Phi^{\dagger}_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} = \frac{1}{2L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2} g_n(\mathbf{k})\left[g^*_{n'}(\mathbf{k})+\eta g^*_{n'}(-\mathbf{k})\right]\ e^{i\mathbf{K}\cdot\left(\mathbf{R}'-\mathbf{R}\right)}\ \langle\widehat{n}_{\mathbf{k}_1}\rangle\,\langle\widehat{n}_{\mathbf{k}_2}\rangle \tag{46}$$

(as in § D-1-a of Chapter XVII, we may consider the summations over $\mathbf{k}_1$ and $\mathbf{k}_2$ as independent, since, for a large volume L^3, ignoring the constraint $\mathbf{k}_1 \neq \pm\mathbf{k}_2$ leads to a negligible error); we have used the notation:

$$\begin{aligned} \mathbf{K} &= \mathbf{k}_1 + \mathbf{k}_2 \\ \mathbf{k} &= \frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \end{aligned} \tag{47}$$

When the parity of the function $g_{n'}(\mathbf{k})$ is η, the two terms in the bracket of (46) are equal and we get the simpler relation:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{forward}} = \frac{1}{L^3} \sum_{\mathbf{k}_1,\mathbf{k}_2} g_n(\mathbf{k})\, g^*_{n'}(\mathbf{k})\ e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}\ \langle \widehat{n}_{\mathbf{k}_1} \rangle \langle \widehat{n}_{\mathbf{k}_2} \rangle \tag{48}$$

This result only depends on the difference $\mathbf{R}-\mathbf{R}'$ (translation invariance); it goes to zero when $|\mathbf{R}'-\mathbf{R}|$ becomes larger than the inverse of the momentum $\mathbf{K}$ distribution width of the function appearing on the right-hand side of (46), once it is summed over $\mathbf{k}$, the difference in momenta.

– (**II**) The terms corresponding to the annihilation-creation of different pairs are obtained for $\mathbf{k}_2 = -\mathbf{k}_1$ and $\mathbf{k}_4 = -\mathbf{k}_3$, with $\mathbf{k}_4 \neq \pm\mathbf{k}_2$. Their contribution is written:

$$\begin{aligned} &\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{paire-paire}} \\ &\quad = \frac{1}{2L^3} \sum_{\mathbf{k}_1} \left\langle a^\dagger_{\mathbf{k}_1} a^\dagger_{-\mathbf{k}_1} \right\rangle\ g_n(\mathbf{k}_1) \sum_{\mathbf{k}_4} \left\langle a_{-\mathbf{k}_4} a_{\mathbf{k}_4} \right\rangle\ g^*_{n'}(\mathbf{k}_4) \end{aligned} \tag{49}$$

Now, using (30) and the definition (2) of the Fourier components of each pair state $|\chi_n\rangle$, we have:

$$\frac{1}{L^{3/2}\sqrt{2}} \sum_{\mathbf{k}_4} \langle a_{-\mathbf{k}_4} a_{\mathbf{k}_4} \rangle\ g^*_{n'}(\mathbf{k}_4) = \sum_{\mathbf{k}_4} \langle \mathbf{k}_4\ |\phi_{\text{pair}}\rangle \langle \chi_{n'}\ |\mathbf{k}_4\rangle = \langle \chi_{n'}\ |\phi_{\text{pair}}\rangle \tag{50}$$

The summation over $\mathbf{k}_1$ is computed in a similar way, via a simple complex conjugation. We then get on the right-hand side of (49) two scalar products, which finally yields:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{pair-pair}} = \langle \phi_{\text{pair}}\ |\chi_n\rangle \langle \chi_{n'}\ |\phi_{\text{pair}}\rangle \tag{51}$$

Unlike the previous contribution, this one is independent of $\mathbf{R}-\mathbf{R}'$.

– (**III**) The terms corresponding to the annihilation-creation of the same pair are obtained for $\mathbf{k}_1 = -\mathbf{k}_2 = \pm\mathbf{k}_3 = \mp\mathbf{k}_4$, and yield the average values $\langle \widehat{n}_{\mathbf{k}_1} \rangle \langle \widehat{n}_{-\mathbf{k}_1} \rangle$ and $\eta \langle \widehat{n}_{\mathbf{k}_1} \rangle \langle \widehat{n}_{-\mathbf{k}_1} \rangle$ respectively. Those terms are just a particular case of the terms appearing in the summation (46) when $\mathbf{k}_1 = -\mathbf{k}_2$, and do not require a specific calculation. Finally, the terms $\mathbf{k}_1 = \mathbf{k}_2$ that we ignored in (**I**), and for which all the $\mathbf{k}$'s must be equal, contain only one summation over the wave vectors; consequently, they are negligible compared to (46), and will be omitted in this computation.

We are then left with the total (**I**) + (**II**), which yields:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\, \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle = \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{forward}} + \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{pair-pair}} \tag{52}$$

where only the second term on the right-hand side does not go to zero when $|\mathbf{R}' - \mathbf{R}|$ becomes large, which indicates a long-range non-diagonal order. According to (51), this second term reaches a maximum when the two internal states $|\chi_n\rangle$ and $|\chi_{n'}\rangle$ are equal to the state $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$ defined in (32). It indicates the existence of a cooperative field of pairs that have a total momentum $\mathbf{K} = 0$ as their external state, and $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$ as their internal state.

Comparing (31) and (51) shows that:

$$\left\langle \Phi_{\chi_n}^{\dagger}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{pair-pair}} = \left\langle \Phi_{\chi_n}(\mathbf{R})\right\rangle^{*} \times \left\langle \Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle \tag{53}$$

The pair-pair term of the two-point correlation function can thus be factored into a product of two one-point correlation functions; for $n = 1$, we get the same function we previously called the "order parameter". As already pointed out in § 2-a-α, it is because the pairs have a zero total momentum that any $\mathbf{R}$ and $\mathbf{R}'$ dependence has disappeared from both sides of (53), but this point is not essential. It is more important to note that introducing such an order parameter, a priori difficult to understand from a physical point of view as it is an average value that does not conserve particle number, is actually quite useful for computing other more physical parameters. We will make the connection between the factorization relation (53) and the Penrose-Onsager criterion for Bose-Einstein condensation in § 2-b-γ.

β. Singlet pairs

Using (26) instead of (13) now leads to a relation very similar to (42); the factor 1/2 is, however, missing, and we must make the substitution:

$$\left\langle a_{\mathbf{k}_1}^{\dagger} a_{\mathbf{k}_2}^{\dagger} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle \Rightarrow \left\langle a_{\mathbf{k}_1,\uparrow}^{\dagger} a_{\mathbf{k}_2,\downarrow}^{\dagger} a_{\mathbf{k}_3,\downarrow} a_{\mathbf{k}_4,\uparrow}\right\rangle \tag{54}$$

Relation (44) then becomes:

$$\begin{aligned}\left\langle \Phi_{\chi_n}^{\dagger}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle = \frac{1}{L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} & g_n\left(\frac{\mathbf{k}_1-\mathbf{k}_2}{2}\right)\; g_{n'}^{*}\left(\frac{\mathbf{k}_4-\mathbf{k}_3}{2}\right) \\ \times\; e^{i\left[(\mathbf{k}_3+\mathbf{k}_4)\cdot\mathbf{R}'-(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{R}\right]} & \left\langle a_{\mathbf{k}_1,\uparrow}^{\dagger} a_{\mathbf{k}_2,\downarrow}^{\dagger} a_{\mathbf{k}_3,\downarrow} a_{\mathbf{k}_4,\uparrow}\right\rangle\end{aligned} \tag{55}$$

The rest of the calculation is very similar to the one we just did, and involves the sum of several terms:

– (**I**) The forward scattering terms are obtained for $\mathbf{k}_4 = \mathbf{k}_1$ and $\mathbf{k}_3 = \mathbf{k}_2$. In two different pairs, a particle is destroyed and then created again in the same individual state (as we now have spin indices, there is no exchange term in this case). The computation is the same as the one that yielded (48) for spinless particles; with the notation (47) for the wave vectors, we get here:

$$\left\langle \Phi_{\chi_n}^{\dagger}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} = \frac{1}{L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2} g_n(\mathbf{k})\, g_{n'}^{*}(\mathbf{k})\;\; e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}\;\; \langle \widehat{n}_{\mathbf{k}_1,\uparrow}\rangle\, \langle \widehat{n}_{\mathbf{k}_2,\downarrow}\rangle \tag{56}$$

– (**II**) The terms corresponding to the annihilation-creation of different pairs are obtained for $\mathbf{k}_2 = -\mathbf{k}_1$ and $\mathbf{k}_4 = -\mathbf{k}_3$, with $\mathbf{k}_4 \neq \mathbf{k}_1$. The computation is now the same

as the one that yielded (51). The right-hand side of (55) becomes:

$$\frac{1}{L^3}\sum_{\mathbf{k}_1,\mathbf{k}_4} g_n(\mathbf{k}_1)\, g^*_{n'}(\mathbf{k}_4)\left\langle a^\dagger_{\mathbf{k}_1\uparrow} a^\dagger_{-\mathbf{k}_1\downarrow} a_{-\mathbf{k}_4\downarrow} a_{\mathbf{k}_4\uparrow}\right\rangle \tag{57}$$

Using the definition (38) for the pair wave function in the singlet case, we again obtain:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{pair-pair}} = \langle \phi_{\text{pair}}\,|\chi_n\rangle\,\langle\chi_{n'}\,|\phi_{\text{pair}}\rangle \tag{58}$$

As mentioned above, any $\mathbf{R}$ dependence has disappeared from this average value since the paired state was built from pairs having a zero total momentum.

– (**III**) The terms corresponding to the annihilation-creation of the same pair are obtained for $\mathbf{k}_2 = \mathbf{k}_3 = -\mathbf{k}_1 = -\mathbf{k}_4$; they are proportional to $\langle \widehat{n}_{\mathbf{k}_1,\uparrow}\rangle \left\langle \widehat{n}_{-\mathbf{k}_1,\downarrow}\right\rangle$ and already included in the terms (**I**). The terms where all the $\mathbf{k}$'s are equal are neglected for the same reason as above.

To sum up, we find as before:

$$\begin{aligned}\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle &= \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} + \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{pair-pair}} \\ &= \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} + \langle \Phi_{\chi_n}(\mathbf{R})\rangle^* \left\langle \Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle \end{aligned} \tag{59}$$

We arrive, finally, at the same results as for spinless bosons, with the same long-range non-diagonal order of the pairs, as well as the factorization (53) of the order parameters. We shall see in Complement C_{XVII} that this long-range order parameter is intimately linked to the nature of the BCS transition. Here again, the anomalous average values turn out to be useful tools for computing normal average values that conserve the particle number.

γ. *Link with Bose-Einstein condensation of pairs*

There is a close link between the order parameter of the pairs and the existence of Bose-Einstein condensation of those pairs. To show this, it is convenient to introduce the density operator of pairs, limiting ourselves, for the sake of simplicity, to the case of spinless particles. In Chapter XVII, the one-particle density operator ρ_I for identical particles was given, in terms of the field operator, by its matrix elements (B-26):

$$\langle \mathbf{r}', \nu'|\,\rho_I\,|\mathbf{r},\nu\rangle = \left\langle \Psi^\dagger_\nu(\mathbf{r})\Psi_{\nu'}(\mathbf{r}')\right\rangle \tag{60}$$

where $\mathbf{r}$ is the particle's position and ν its spin. For pairs, the corresponding relation is written[2]:

$$\langle \mathbf{R}', \chi_{n'}|\,\rho_I^{\text{pair}}\,|\mathbf{R},\chi_n\rangle = \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle \tag{61}$$

where $\mathbf{R}$ is the position of the center of mass, and χ_n and $\chi_{n'}$ define the internal state of the pair; the index n plays a role similar to that of a spin index for a single particle (even though it corresponds to an internal orbital state).

[2] As we shall see in § 3, the pair field operators do not exactly satisfy the boson commutation relations. Consequently, operator (61) is not, strictly speaking, a density operator; to underline this difference, ρ_I^{pair} is sometimes called a "density quasi-operator".

In the momentum representation, the diagonal matrix elements of this density operator are:

$$\langle \mathbf{K}', \chi_n | \, \rho_I^{\text{pair}} \, | \mathbf{K}', \chi_n \rangle = \frac{1}{L^3} \int \mathrm{d}^3 R \int \mathrm{d}^3 R' \; e^{i\mathbf{K}'\cdot(\mathbf{R}-\mathbf{R}')} \left\langle \Phi_{\chi_n}^\dagger(\mathbf{R}) \Phi_{\chi_n}(\mathbf{R}') \right\rangle \tag{62}$$

Since $\left\langle \Phi_{\chi_n}^\dagger(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle$ only depends on $\mathbf{R} - \mathbf{R}'$, we perform the change of variables $\mathbf{X} = \mathbf{R} - \mathbf{R}'$; the integral over $\mathrm{d}^3 R$ is then trivial and cancels the factor $1/L^3$; we therefore get:

$$\langle \mathbf{K}', \chi_n | \, \rho_I^{\text{pair}} \, | \mathbf{K}', \chi_n \rangle = \int \mathrm{d}^3 X \; e^{i\mathbf{K}'\cdot\mathbf{X}} \; \left\langle \Phi_{\chi_n}^\dagger(\mathbf{R}) \Phi_{\chi_n}(\mathbf{R} - \mathbf{X}) \right\rangle \tag{63}$$

Inserting relation (52) in this result, we get the sum of a contribution from the forward scattering term and from the pair-pair term.

(i) The first contribution comes from inserting (48) in (63). The integral over $\mathrm{d}^3 X$ yields a delta function $\delta_{\mathbf{K},\mathbf{K}'}$ and a factor L^3 that cancels the same factor in the denominator. As the sum $\mathbf{k}_1 + \mathbf{k}_2$ must now be equal to $\mathbf{K}'$, the double summation over $\mathbf{k}_1$ and $\mathbf{k}_2$ reduces to a summation over $\mathbf{k}$. We then get:

$$\sum_{\mathbf{k}} |g_n(\mathbf{k})|^2 \; \left\langle \widehat{n}_{\frac{\mathbf{K}'}{2}+\mathbf{k}} \right\rangle \left\langle \widehat{n}_{\frac{\mathbf{K}'}{2}-\mathbf{k}} \right\rangle \tag{64}$$

This result is a regular function of $\mathbf{K}'$, related to the wave number dependence of the occupation numbers.

(ii) The pair-pair term contains the integral of the function (53), which is a product of two constant order parameters; it therefore leads to:

$$\delta_{\mathbf{K},\mathbf{0}} \; L^3 \left| \left\langle \Phi_{\chi_n} \right\rangle \right|^2 \tag{65}$$

The presence of the delta function $\delta_{\mathbf{K},\mathbf{0}}$ shows that the $\mathbf{K} = 0$ level has an additional population (number of quanta of the pair field) that does not exist for any other value of the momentum $\mathbf{K}$; this population is simply the square of the order parameter, multiplied by the system's volume; it is thus an extensive quantity. It indicates that the pairs of the system undergo Bose-Einstein condensation. As the corresponding population is proportional to the square of the order parameter, this clearly shows the close link between the long-range non-diagonal order, the order parameter and the existence of condensation. The factorization appearing in (53) is often called the "Penrose-Onsager condensation criterion".

2-c. Application to the computation of the correlation function (singlet pairs)

The average values of products of pair field operators can also be used to get the correlation functions between particles. We are going to show, in particular, that the correlation function G_2 is the sum of an "incoherent" term, independent of the positions, and of a coherent term that involves the pair wave function defined previously. In order to keep the demonstration short, we shall limit the discussion to the case of fermions described by a paired state built from singlet pairs[3], but the transposition to spinless particles is fairly straightforward.

[3]Condensed bosons will be studied in Complement E_{XVII}. We will then show that the properties of the paired state, built from the $\mathbf{k} \neq 0$ states, are not determined by the interactions within this paired state, but rather by the interactions with a condensate $\mathbf{k} = 0$, external to the paired state. It is therefore a completely different case.

Relation (24) expresses the conjugate of the pair field operator as a function of products of creation operators for its constituent particles; we shall start by inverting this relation.

α. Inversion of the relation between fields

The closure relation for the orthonormal basis of the wave functions $\chi_p(\mathbf{r})$ with $p = 1, 2, ...$ is written:

$$\sum_n \chi_n^*(\mathbf{x}')\ \chi_n(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}') \tag{66}$$

This summation over n must include even orbital functions $\chi_n(\mathbf{x})$ (associated with a pair field Φ_{χ_n} describing paired fermions in a singlet state) as well as odd functions (associated with a pair field describing fermions paired in a triplet state). We then multiply (24) by $\chi_n^*(\mathbf{x}')$ and perform the summation over n. We recognize in the integral on the right-hand side the closure relation (66), which yields:

$$\sum_n \chi_n^*(\mathbf{x}')\ \Phi_{\chi_n}^\dagger(\mathbf{R}) = \Psi_\uparrow^\dagger(\mathbf{R} + \frac{\mathbf{x}'}{2})\Psi_\downarrow^\dagger(\mathbf{R} - \frac{\mathbf{x}'}{2}) \tag{67}$$

This leads to:

$$\Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2) = \sum_n \chi_n^*(\mathbf{r}_1 - \mathbf{r}_2)\ \Phi_{\chi_n}^\dagger\left(\frac{\mathbf{r}_1 + \mathbf{r}_2}{2}\right) \tag{68}$$

Creating two particles of opposite spins at points $\mathbf{r}_1$ and $\mathbf{r}_2$ thus amounts to creating a coherent superposition of pairs with a center of mass at $(\mathbf{r}_1 + \mathbf{r}_2)/2$, in a singlet or triplet spin state, and with coefficients equal to the wave functions χ_n^* taken at the position $\mathbf{r}_1 - \mathbf{r}_2$.

The average value of this expression can be computed in a paired state, using relation (37). This leads to:

$$\left\langle \Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2) \right\rangle = \sum_n \chi_n^*(\mathbf{r}_1 - \mathbf{r}_2)\ \ \langle \chi_n\, |\phi_{\text{pair}}\rangle \tag{69}$$

As ϕ_{pair} is an even function, it easily follows that only the even χ_n will contribute to this average value; the triplet pair fields have a zero average value in a singlet pair state.

As in § 2-a-β, we can choose for the $|\chi_n\rangle$ a basis whose first ket $|\chi_1\rangle$ coincides with the normalized pair ket $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$. We then get:

$$\left\langle \Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2) \right\rangle = \frac{\phi_{\text{pair}}^*(\mathbf{r}_1 - \mathbf{r}_2)}{\sqrt{\langle\phi_{\text{pair}}\,|\phi_{\text{pair}}\rangle}} \frac{\langle\phi_{\text{pair}}\,|\phi_{\text{pair}}\rangle}{\sqrt{\langle\phi_{\text{pair}}\,|\phi_{\text{pair}}\rangle}} = \phi_{\text{pair}}^*(\mathbf{r}_1 - \mathbf{r}_2) \tag{70}$$

β. 4-point correlation function

According to (68), the 4-point correlation function (for opposite spins) is written as:

$$\begin{aligned}&\left\langle \Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2)\Psi_\downarrow(\mathbf{r}_2')\Psi_\uparrow(\mathbf{r}_1') \right\rangle \\ &= \sum_{n,n'} \chi_n^*(\mathbf{r}_1 - \mathbf{r}_2)\,\chi_{n'}(\mathbf{r}_1' - \mathbf{r}_2') \left\langle \Phi_{\chi_n}^\dagger\left(\frac{\mathbf{r}_1 + \mathbf{r}_2}{2}\right)\Phi_{\chi_{n'}}\left(\frac{\mathbf{r}_1' + \mathbf{r}_2'}{2}\right)\right\rangle\end{aligned} \tag{71}$$

It is expressed in terms of the average values of products of pair creation and annihilation operators, hence in terms of the average values of products of fields for which the index n plays the role of an internal state of the molecule. We will show below that it can be expressed as:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}_2')\Psi_{\uparrow}(\mathbf{r}_1')\right\rangle = G_1\left(\mathbf{r}_1,\uparrow;\mathbf{r}_1',\uparrow\right)\times G_1\left(\mathbf{r}_2,\downarrow;\mathbf{r}_2',\downarrow\right) + \phi_{\text{pair}}^{*}\left(\mathbf{r}_1-\mathbf{r}_2\right)\times\phi_{\text{pair}}\left(\mathbf{r}_1'-\mathbf{r}_2'\right) \tag{72}$$

where $G_1\left(\mathbf{r},\uparrow;\mathbf{r}',\uparrow\right)$ is the non-diagonal one-particle correlation function, the Fourier transform of $\langle n_{\mathbf{k},\uparrow}\rangle$:

$$G_1\left(\mathbf{r},\uparrow;\mathbf{r}',\uparrow\right)=\frac{1}{L^3}\sum_{\mathbf{k}}e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)}\ \langle n_{\mathbf{k},\uparrow}\rangle \tag{73}$$

and with a similar definition for $G_1\left(\mathbf{r},\downarrow;\mathbf{r}',\downarrow\right)$, the occupation number $\langle n_{\mathbf{k},\uparrow}\rangle$ being simply replaced by $\langle n_{\mathbf{k},\downarrow}\rangle$; the pair wave function ϕ_{pair} has already been defined in (38).

The function $G_1\left(\mathbf{r},\uparrow;\mathbf{r}',\uparrow\right)$, being the Fourier transform of a regular function $\langle n_{\mathbf{k},\uparrow}\rangle$, tends toward zero when the difference $|\mathbf{r}-\mathbf{r}'|$ is larger than a certain (microscopic) limit; the only terms left are those on the second line of (72). Imagine then that positions $\mathbf{r}_1$ and $\mathbf{r}_2$ are close to each other, forming a first group, and that the same is true for positions $\mathbf{r}_1'$ and $\mathbf{r}_2'$, forming a second group, while these two groups are far from each other. The non-diagonal correlation function can then be factored into a product of functions ϕ_{pair}. This situation is reminiscent of the Penrose-Onsager criterion for Bose-Einstein condensation of bosons (Complement A_{XVI}, § 3-a), but it now concerns the 4-point (instead of 2-point) non-diagonal correlation function. As the norm of ϕ_{pair} is the order parameter, it again underlines the important role of this parameter.

An important particular case of the 4-point correlation function is the two-body (diagonal) correlation function for opposite spins:

$$G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right)=\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}_2)\Psi_{\uparrow}(\mathbf{r}_1)\right\rangle \tag{74}$$

The intensity of the pair field is therefore written:

$$\begin{aligned} G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right) &= \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2}\langle\widehat{n}_{\mathbf{k}_1,\uparrow}\rangle\langle\widehat{n}_{\mathbf{k}_2,\downarrow}\rangle+\left|\phi_{\text{pair}}\left(\mathbf{r}_1-\mathbf{r}_2\right)\right|^2 \\ &= \frac{\left\langle\widehat{N_{\uparrow}}\right\rangle\left\langle\widehat{N_{\downarrow}}\right\rangle}{L^6}+\left|\phi_{\text{pair}}\left(\mathbf{r}_1-\mathbf{r}_2\right)\right|^2 \end{aligned} \tag{75}$$

We find again relation (D-17) of Chapter XVII, but via another method. The two-body correlation function is the sum of a contribution independent of the positions (hence, with no correlations) and of the modulus squared of the pair wave function. This latter contribution comes from the term that, for pairs, indicates the existence of a long-range non-diagonal order (Bose-Einstein condensation). This is an important property, which is at the heart of the BCS mechanism, and which will be discussed in more detail in Complement B_{XVII}.

Demonstration

We insert in (71) relations (56) and (58). In the forward scattering term, we get the following expression:

$$\begin{aligned}
&\sum_{n,n'} \chi_n^*(\mathbf{r}_1-\mathbf{r}_2)\,\chi_{n'}\left(\mathbf{r}_1'-\mathbf{r}_2'\right) g_n(\mathbf{k})\, g_{n'}^*(\mathbf{k})\, e^{i\mathbf{K}\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \sum_{n,n'} \langle\chi_n\,|\mathbf{r}_1-\mathbf{r}_2\rangle\ \left\langle\mathbf{r}_1'-\mathbf{r}_2'\,|\chi_{n'}\right\rangle\ \langle\mathbf{k}\,|\chi_n\rangle\ \langle\chi_{n'}\,|\mathbf{k}\rangle\ e^{i\mathbf{K}\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \langle\mathbf{k}\,|\mathbf{r}_1-\mathbf{r}_2\rangle\left\langle\mathbf{r}_1'-\mathbf{r}_2'\,|\mathbf{k}\right\rangle\ e^{i\mathbf{K}\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \frac{1}{L^3}\ e^{i(\mathbf{k}_2-\mathbf{k}_1)\cdot\left(\mathbf{r}_1'-\mathbf{r}_2'-\mathbf{r}_1+\mathbf{r}_2\right)/2}\ e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \frac{1}{L^3}\ e^{i\mathbf{k}_2\cdot\left(\mathbf{r}_1'-\mathbf{r}_1\right)}\ e^{i\mathbf{k}_1\cdot\left(\mathbf{r}_2'-\mathbf{r}_2\right)}
\end{aligned} \tag{76}$$

where $\mathbf{k}$ and $\mathbf{K}$ were defined in (47). Inserting this result in (71), we get the first term of the right-hand side of (72)

As for the pair annihilation-creation term (58), it yields:

$$\begin{aligned}
&\left\langle\Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2)\Psi_\downarrow(\mathbf{r}_2')\Psi_\uparrow(\mathbf{r}_1')\right\rangle_{\text{pair-pair}} \\
&= \sum_{n,n'} \chi_n^*(\mathbf{r}_1-\mathbf{r}_2)\,\chi_{n'}\left(\mathbf{r}_1'-\mathbf{r}_2'\right)\langle\phi_{\text{pair}}\,|\chi_n\rangle\,\langle\chi_{n'}\,|\phi_{\text{pair}}\rangle \\
&= \sum_{n,n'} \left\langle\mathbf{r}_1'-\mathbf{r}_2'\,|\chi_{n'}\right\rangle\langle\chi_{n'}\,|\phi_{\text{pair}}\rangle\times\langle\phi_{\text{pair}}\,|\chi_n\rangle\,\langle\chi_n\,|\mathbf{r}_1-\mathbf{r}_2\rangle \\
&= \phi_{\text{pair}}^*(\mathbf{r}_1-\mathbf{r}_2)\times\phi_{\text{pair}}\left(\mathbf{r}_1'-\mathbf{r}_2'\right)
\end{aligned} \tag{77}$$

and we obtain the second term of the right-hand side of (72). Note that only the singlet pair fields (associated with the even function χ_n) contribute to this term.

3. Commutation relations of field operators

We now study the commutation relations between the pair field operators just defined. The "spin-statistics theorem" (Chapter XIV, § C-1) states that particles with integer spin are bosons, and particles with half-integer spin are fermions. If we consider two paired fermions, the rules for adding angular momenta (Chapter X) indicate that this composite system necessarily has an integer spin. Intuitively, one could thus expect two bound fermions to behave like a boson; this is the question we now discuss by examining the commutation relations between the operators $A_{\mathbf{K}}$ and $A_{\mathbf{K}}^\dagger$, and establishing the correction factors introduced by the underlying fermionic structure.

3-a. Particles in the same spin state

Starting with spinless particles, we shall explain in this simple case the main commutation properties of the pair operators. If the pairs created and annihilated by the operators $A_{\mathbf{K}}$ and $A_{\mathbf{K}'}$ and their Hermitian conjugates were really bosons, the commutator of these two operators should be equal to $\delta_{\mathbf{K}\mathbf{K}'}$. We are going to show that the commutator does contain such a term, but with several additional corrections.

α. Commutation relations of the $A_{\mathbf{K},n}$

Any product of two creation operators commute with any product of two creation operators (for fermions, two minus signs cancel each other when products of two operators cross each other); the same is true for two products of annihilation operators. We therefore have:

$$\begin{aligned}\left[A^\dagger_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right] &= 0\\ \left[A_{\mathbf{K},n}, A_{\mathbf{K}',n'}\right] &= 0\end{aligned} \tag{78}$$

The commutator of $A_{\mathbf{K},n}$ and $A^\dagger_{\mathbf{K}',n'}$ has yet to be computed:

$$\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right] = \frac{1}{2}\sum_{\mathbf{k}} \left(g^n_{\mathbf{k}}\right)^* \sum_{\mathbf{k}'} g^{n'}_{\mathbf{k}'} \left[a_{\frac{\mathbf{K}}{2}-\mathbf{k}} a_{\frac{\mathbf{K}}{2}+\mathbf{k}} \ , \ a^\dagger_{\frac{\mathbf{K}'}{2}+\mathbf{k}'} a^\dagger_{\frac{\mathbf{K}'}{2}-\mathbf{k}'}\right] \tag{79}$$

We will show below (§ 3-a-β) that:

$$\begin{aligned}\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right] &= \delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'} + 2\eta \sum_{\mathbf{k}} \left(g^n_{\mathbf{k}}\right)^*\ g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k}}\ a_{(\mathbf{K}/2)-\mathbf{k}}\\ &= \delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'} + 2\eta \sum_{\boldsymbol{\kappa}} \left(g^n_{\boldsymbol{\kappa}-\frac{\mathbf{K}}{2}}\right)^*\ g^{n'}_{\boldsymbol{\kappa}-\frac{\mathbf{K}'}{2}}\ a^\dagger_{\mathbf{K}'-\boldsymbol{\kappa}}\ a_{\mathbf{K}-\boldsymbol{\kappa}}\end{aligned} \tag{80}$$

(in the second line, we have set $\boldsymbol{\kappa} = \mathbf{k} + \mathbf{K}/2$ and used the parity η of the function g); if needed, we can get rid of the η coefficient on the right-hand side provided we change the sign of the subscript of g (or of g^*).

The first term $\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'}$, on the right-hand side of (80) , is exactly the commutator of two bosons with internal states n and n' (spin states for example): this term is different from zero only if both the external and internal variables are the same (in the present case, these internal states are actually orbital states). This first term is, however, followed by an additional term that shows that the fermionic structure of the pairs still plays a role. This latter term is a one-particle operator in the sense defined in § B of Chapter XV; relation (B-12) of that chapter permits computing the matrix elements of the corresponding operator $\widehat{f}$. This additional term contains creation and annihilation operators in normal order, which means that it will go to zero when the populations of the individual states tend to zero; in this limit, the pairs can be assimilated to bosons.

When $\mathbf{K}' = \mathbf{K}$ and $n = n'$ (pairs in the same internal and external states), we get the simpler relation:

$$\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K},n}\right] = 1 + 2\eta \sum_{\boldsymbol{\kappa}} \left|g^n_{\boldsymbol{\kappa}-\frac{\mathbf{K}}{2}}\right|^2\ \widehat{n}_{\mathbf{K}-\boldsymbol{\kappa}} \tag{81}$$

with the usual definition of the population operator $\widehat{n}_{\mathbf{k}}$:

$$\widehat{n}_{\mathbf{k}} = a^\dagger_{\mathbf{k}} a_{\mathbf{k}} \tag{82}$$

The corrections to a purely bosonic commutation are then proportional to the populations of the individual states of the particles forming the pair, hence confirming the fact that they become negligible when the sum of all these populations is small enough.

β. Demonstration

We start by computing the commutator:

$$\left[a_1 a_2\ ,\ a_3^\dagger a_4^\dagger\right] = a_1 a_2 a_3^\dagger a_4^\dagger - a_3^\dagger a_4^\dagger a_1 a_2 \tag{83}$$

where $1, 2, 3, 4$ are indices labeling any individual states. We can write:

$$\begin{aligned} a_1 a_2 a_3^\dagger a_4^\dagger &= \delta_{23}\ a_1 a_4^\dagger + \eta\ a_1 a_3^\dagger a_2 a_4^\dagger \\ &= \delta_{23}\ a_1 a_4^\dagger + \eta\delta_{13}\ a_2 a_4^\dagger + a_3^\dagger a_1 a_2 a_4^\dagger \\ &= \delta_{23}\ a_1 a_4^\dagger + \eta\delta_{13}\ a_2 a_4^\dagger + \delta_{24}\ a_3^\dagger a_1 + \eta\ a_3^\dagger a_1 a_4^\dagger a_2 \\ &= \delta_{23}\ a_1 a_4^\dagger + \eta\delta_{13}\ a_2 a_4^\dagger + \delta_{24}\ a_3^\dagger a_1 + \eta\ \delta_{14}\ a_3^\dagger a_2 + a_3^\dagger a_4^\dagger a_1 a_2 \end{aligned} \tag{84}$$

so that:

$$\left[a_1 a_2\ ,\ a_3^\dagger a_4^\dagger\right] = \delta_{23}\ a_1 a_4^\dagger + \delta_{24}\ a_3^\dagger a_1 + \eta\left[\delta_{13}\ a_2 a_4^\dagger + \delta_{14}\ a_3^\dagger a_2\right] \tag{85}$$

Putting all the operators in normal order, we get[4]:

$$\left[a_1 a_2\ ,\ a_3^\dagger a_4^\dagger\right] = \delta_{23}\delta_{14} + \eta\delta_{13}\delta_{24} + \delta_{24}\ a_3^\dagger a_1 + \delta_{13}\ a_4^\dagger a_2 + \eta\left[\delta_{23}\ a_4^\dagger a_1 + \delta_{14}\ a_3^\dagger a_2\right] \tag{86}$$

The commutator appearing on the right-hand side of (79) is therefore equal to:

$$\begin{aligned} &\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{\mathbf{k},\mathbf{k}'} + \eta\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{\mathbf{k},-\mathbf{k}'} + \Big[\delta_{(\mathbf{K}-\mathbf{K}')/2,-\mathbf{k}-\mathbf{k}'}\ a^\dagger_{(\mathbf{K}'/2)+\mathbf{k}'}\ a_{(\mathbf{K}/2)-\mathbf{k}} \\ &+ \delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}+\mathbf{k}'}\ a^\dagger_{(\mathbf{K}'/2)-\mathbf{k}'}\ a_{(\mathbf{K}/2)+\mathbf{k}} \\ &+\eta\delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}'-\mathbf{k}}\ a^\dagger_{(\mathbf{K}'/2)-\mathbf{k}'}\ a_{(\mathbf{K}/2)-\mathbf{k}} + \eta\delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}-\mathbf{k}'}\ a^\dagger_{(\mathbf{K}'/2)+\mathbf{k}'}\ a_{(\mathbf{K}/2)+\mathbf{k}}\Big] \end{aligned} \tag{87}$$

Inserting the first two terms back into (79), we get the following contribution to the commutator $\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right]$:

$$\frac{1}{2}\delta_{\mathbf{K},\mathbf{K}'}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ \left[g^{n'}_{\mathbf{k}} + \eta g^{n'}_{-\mathbf{k}}\right] = \delta_{\mathbf{K},\mathbf{K}'}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\, g^{n'}_{\mathbf{k}} = \delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'} \tag{88}$$

where we have taken into account the parity with respect to $\mathbf{k}$ of the the functions $g^n_{\mathbf{k}}$ – see relation (A-4) of Chapter XVII – and used the fact that the internal states are orthonormal. As mentioned above, this $\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'}$ is precisely what is expected for a boson commutation relation.

It is, however, followed in (87) by four other terms, which are written:

$$\begin{aligned} &\frac{1}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{-\frac{\mathbf{K}-\mathbf{K}'}{2}-\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k}} a_{(\mathbf{K}/2)-\mathbf{k}} \\ &\frac{1}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}-\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)+\mathbf{k}} a_{(\mathbf{K}/2)+\mathbf{k}} \\ &\frac{\eta}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k}} a_{(\mathbf{K}/2)-\mathbf{k}} \\ &\frac{\eta}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{-\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)+\mathbf{k}} a_{(\mathbf{K}/2)+\mathbf{k}} \end{aligned} \tag{89}$$

[4]Since $\left[a_i, a_j^\dagger\right]_{-\eta} = \delta_{ij}$, we have $a_i a_j^\dagger = \eta a_j^\dagger a_i + \delta_{ij}$.

In each of them, and without modifying the result, we can change the sign of the summation dummy index $\mathbf{k}$, or change the sign of the subscript of the functions g or g^* (provided we introduce a factor η). For example, in the second term, we can change the sign of the subscripts of the two functions g and g^* (two factors η then cancel each other), then change the summation index $\mathbf{k}$ into $-\mathbf{k}$: this second term then doubles the first one. As for the third term, we simply change the sign of the subscript $(\mathbf{K}-\mathbf{K}')/2+\mathbf{k}$ of the function g (which introduces a factor η canceling that same factor already present) and reproduce the first term. Finally, for the fourth term, a parity operation on the function g^* followed by a change of the summation index from $\mathbf{k}$ to $-\mathbf{k}$ makes it equal to the first. The four terms are therefore equal; choosing for example the expression of the third one, and replacing the summation index $\mathbf{k}$ by $\boldsymbol{\kappa}=\mathbf{k}+\mathbf{K}/2$, we get relation (80).

γ. Commutation relations of pair field operators

For the same reasons as explained above (commutation of any products of two annihilation operators), the operators $\Phi_{\chi_n}(\mathbf{R})$ all commute with each others; the same is true for the adjoint operators $\Phi^\dagger_{\chi_n}(\mathbf{R})$. We have yet to examine the relations between the $\Phi_{\chi_n}(\mathbf{R})$ and the $\Phi^\dagger_{\chi_{n'}}(\mathbf{R}')$. Relations (11) and (12) show that:

$$\begin{aligned}\left[\Phi_{\chi_n}(\mathbf{R}),\Phi^\dagger_{\chi_{n'}}(\mathbf{R}')\right] = \\ \frac{1}{2}\int \mathrm{d}^3x\ \chi^*_n(\mathbf{x})\int \mathrm{d}^3x'\ \chi_{n'}(\mathbf{x}') \\ \left[\Psi(\mathbf{R}-\frac{\mathbf{x}}{2})\Psi(\mathbf{R}+\frac{\mathbf{x}}{2}),\Psi^\dagger(\mathbf{R}'+\frac{\mathbf{x}'}{2})\Psi^\dagger(\mathbf{R}'-\frac{\mathbf{x}'}{2})\right]\end{aligned} \tag{90}$$

The computation will not be carried out explicitly (though it does not present any particular difficulty); it leads to:

$$\begin{aligned}&\left[\Phi_{\chi_n}(\mathbf{R}),\Phi^\dagger_{\chi_{n'}}(\mathbf{R}')\right] \\ &\quad=\delta(\mathbf{R}-\mathbf{R}')\ \delta_{n,n'} \\ &\qquad+16\int \mathrm{d}^3x\ \chi^*_n[\mathbf{x}]\,\chi_{n'}[2(\mathbf{R}'-\mathbf{R})-\mathbf{x}]\,\Psi^\dagger(2\mathbf{R}'-\mathbf{R}-\frac{\mathbf{x}}{2})\Psi(\mathbf{R}-\frac{\mathbf{x}}{2}) \\ &\quad=\delta(\mathbf{R}-\mathbf{R}')\ \delta_{n,n'} \\ &\qquad+16\int \mathrm{d}^3z\ \chi^*_n(\mathbf{R}'-\mathbf{R}-\mathbf{z})\,\chi_{n'}(\mathbf{R}'-\mathbf{R}+\mathbf{z}) \\ &\qquad\times\Psi^\dagger(\frac{\mathbf{R}+\mathbf{R}'+\mathbf{z}}{2}+\mathbf{R}'-\mathbf{R})\ \Psi(\frac{\mathbf{R}+\mathbf{R}'+\mathbf{z}}{2}+\mathbf{R}-\mathbf{R}')\end{aligned} \tag{91}$$

In the second more symmetrical form of this commutator, we have used the notation $\mathbf{z}=\mathbf{R}'-\mathbf{R}-\mathbf{x}$. These relations are the equivalent, in the position representation, of the commutation relations (80) in the momentum representation (as already mentioned, it is possible to make the factor η appear or disappear in front of the integral by changing the sign of the variable of one of the two functions χ^*_n or $\chi_{n'}$).

The commutator thus includes several terms. The first term in $\delta(\mathbf{R}-\mathbf{R}')\,\delta_{n,n'}$ corresponds to the commutation relation of a usual bosonic field (whether the pair constituents are bosons or fermions); the $\delta_{n,n'}$ reflects the commutation of field components

corresponding to different orbital internal states of the pairs. To this term must be added a correction that depends on the structure of the pair, characterized by the functions $\chi_n(\mathbf{r})$ and $\chi_{n'}(\mathbf{r})$. We again find a result similar to that obtained before: a first simple bosonic term, which only takes into account the simultaneous exchanges of the two constituents of a "molecule" with the two constituents of another one. This term is followed by a correction that comes from the possibility of exchanges other than those involving complete pairs. Note that this correction is expressed in terms of field operators of the elementary constituents themselves and not of the pairs, as was to be expected since it is the constituents themselves that are involved. This correction term is a one-particle operator, non-diagonal in the position representation, since it destroys a particle at point $\mathbf{r}$ and recreates another one at point $\mathbf{r} + \mathbf{2}(\mathbf{R}' - \mathbf{R})$, always at the same distance.

To keep things simple, let us assume the dimensions of the "molecules" that define the pair field for the two internal states n and n' we are concerned with, are both of the order of the same dimension a_0; this means that the wave functions $\chi_n(\mathbf{r})$ and $\chi_{n'}(\mathbf{r})$ go to zero when $r \gg a_0$. In relation (91), the values of $\mathbf{x}$ that contribute to the integral are those for which none of the two functions $\chi_n^*[\mathbf{x}]$ and $\chi_{n'}[2(\mathbf{R} - \mathbf{R}') - \mathbf{x}]$ takes a negligible value; this requires that neither $|\mathbf{x}|$, nor $|2(\mathbf{R} - \mathbf{R}') - \mathbf{x}|$, be large compared to a_0. This double condition imposes $|\mathbf{R} - \mathbf{R}'| \simeq a_0$, in which case there are values of $\mathbf{x}$ for which both functions take simultaneously large values and the correction to the commutator cannot be neglected. On the other hand, if $|\mathbf{R} - \mathbf{R}'| \gg a_0$, there is no common domain where both functions χ_n^* and $\chi_{n'}$ take on significant values and the integral over d^3z is practically negligible. In other words, the molecular wave function range a_0 also plays the role of the commutator correction range.

The limit $a_0 \to 0$ can be obtained by choosing functions χ proportional to a function δ^ε (Appendix II, § 1-b), whose width goes to zero as $\varepsilon \to 0$ and whose integral equals one (it takes values of the order of $1/\varepsilon^3$ in a domain of volume of the order of ε^3). For the sake of simplicity, we assume that $n' = n$; as it is the square of the function χ_n that is normalized to 1 (and not the function itself), we must choose:

$$\chi_n(\mathbf{r}) \simeq \varepsilon^{3/2}\ \delta^\varepsilon(\mathbf{r}) \tag{92}$$

Using this form for the functions χ, the integral over d^3z in (91) leads to the convolution of two delta functions, which yields a function $\delta(\mathbf{R} - \mathbf{R}')$ multiplied by the operator $\Psi^\dagger(\mathbf{R})\Psi(\mathbf{R})$; nevertheless, the coefficient ε^3 of this term yields zero in the limit where $\varepsilon \to 0$. Consequently, if the molecules' size is very small compared to all the characteristic lengths of the system (such as the distance between molecules), the commutation relations of the field operator are exactly the same as for fields associated with bosons.

In conclusion, when the "molecules" have no spatial overlap[5], the only relevant exchanges concern exchanges between both of their constituents. On the other hand, when the two molecules do overlap, individual exchanges between their constituents become possible. If the molecules are loosely bound, as in the example of the BCS fermion pairing mechanism (Complement $\mathrm{C_{XVII}}$), they cannot be treated as bosons without structure, and one must use the complete formula (91) for the commutator.

[5]This does not exclude the case where the distance between molecules is small or comparable to the de Broglie wavelength of their centers of mass: the gas of molecules may be degenerate.

3-b. Singlet pairs

We now study the case of particles in a singlet pair, as in § 1-b.

α. Commutation relations of the $A_{\mathbf{K},n}$

As before, any products of two creation operators commute with any products of two creation operators; the same is true for any products of two annihilation operators. Relations (78) are thus still valid. We now have to compute the commutator:

$$\left[A_{\mathbf{K},n}, A^{\dagger}_{\mathbf{K}',n'}\right] = \sum_{\mathbf{k}} \; (g^n_{\mathbf{k}})^* \sum_{\mathbf{k}'} \; g^{n'}_{\mathbf{k}'} \; \left[a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} a_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow} \; , \; a^{\dagger}_{\frac{\mathbf{K}'}{2}+\mathbf{k}',\uparrow} a^{\dagger}_{\frac{\mathbf{K}'}{2}-\mathbf{k}',\downarrow}\right] \tag{93}$$

We are going to show that:

$$\begin{aligned} &\left[A_{\mathbf{K},n}, A^{\dagger}_{\mathbf{K}',n'}\right] = \delta_{\mathbf{K}\mathbf{K}'} \; \delta_{n,n'} \\ &+\eta \textstyle\sum_{\mathbf{k}} (g^n_{\mathbf{k}})^* \; g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}} \left[a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k},\downarrow} a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} + a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k},\uparrow} \; a_{\frac{\mathbf{K}}{2}-\mathbf{k},\uparrow}\right] \\ &= \delta_{\mathbf{K}\mathbf{K}'} \; \delta_{n,n'} + \eta \textstyle\sum_{\kappa} \left(g^n_{\boldsymbol{\kappa}-\frac{\mathbf{K}}{2}}\right)^* \; g^{n'}_{\boldsymbol{\kappa}-\frac{\mathbf{K}'}{2}} \left[a^{\dagger}_{\mathbf{K}'-\boldsymbol{\kappa},\downarrow} a_{\mathbf{K}-\boldsymbol{\kappa},\downarrow} + a^{\dagger}_{\mathbf{K}'-\boldsymbol{\kappa},\uparrow} \; a_{\mathbf{K}-\boldsymbol{\kappa},\uparrow}\right] \end{aligned} \tag{94}$$

(with, in the last line, the notation $\boldsymbol{\kappa} = \mathbf{k}+\mathbf{K}/2$). Here again we find that the commutator of the two operators $A_{\mathbf{K},n}$ and $A^{\dagger}_{\mathbf{K}',n'}$ includes, to begin, a purely bosonic term, followed by corrections containing operators in normal order (which go to zero in the limit of low occupation numbers); a correction must be added for each of the two spin states.

Demonstration

To prove (94), we again use relation (86). As the indices 1, 2, 3 and 4 represent all the quantum numbers associated with an individual state, they must now contain the spin indices; these are added to the momentum indices, which play the same role as in the previous calculation. It then follows that the states 1 and 3 are always orthogonal, as are the states 2 and 4; the only terms remaining on the right-hand side of (86) are the terms in δ_{23} and δ_{14}, so that:

$$\begin{aligned} &\left[A_{\mathbf{K},n}, A^{\dagger}_{\mathbf{K}',n'}\right] = \textstyle\sum_{\mathbf{k}} (g^n_{\mathbf{k}})^* \sum_{\mathbf{k}'} g^{n'}_{\mathbf{k}'} \left[\delta_{\mathbf{K}\mathbf{K}'} \; \delta_{\mathbf{k}\mathbf{k}'}\right. \\ &\quad \left.+\eta \left(\delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}'-\mathbf{k}} \; a^{\dagger}_{\frac{\mathbf{K}'}{2}-\mathbf{k}',\downarrow} a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} + \delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}-\mathbf{k}'} \; a^{\dagger}_{\frac{\mathbf{K}'}{2}+\mathbf{k}',\uparrow} \; a_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow}\right)\right] \end{aligned} \tag{95}$$

or else (since the basis of the functions $g^n_{\mathbf{k}}$ is orthonormal):

$$\begin{aligned} &\delta_{\mathbf{K}\mathbf{K}'} \; \delta_{n,n'} \\ &+ \eta \sum_{\mathbf{k}} (g^n_{\mathbf{k}})^* \left[\; g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}} a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k},\downarrow} a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow}\right. \\ &\qquad\qquad \left. + g^{n'}_{-\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}} a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)+\mathbf{k},\uparrow} \; a_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow}\right] \end{aligned} \tag{96}$$

We now modify the second term in the bracket of the summation, to make it similar to the first one: as the functions $g^n_{\mathbf{k}}$ have a definite parity $(-\eta)$ with respect to $\mathbf{k}$, we can change the index signs of g and g^*, and the sign of the dummy index $\mathbf{k}$ of the summation. The only difference between the two terms is then the spin directions, and we therefore obtain (94).

β. Commutation relations of pair field operators

A similar calculation to the one that led to (91) permits obtaining:

$$\begin{aligned}\left[\Phi_{\chi}\left(\mathbf{R}\right),\Phi_{\chi}^{\dagger}\left(\mathbf{R}'\right)\right] &= \delta\left(\mathbf{R}-\mathbf{R}'\right)\ \delta_{n,n'} \\ &+ 8\eta\int \mathrm{d}^3x\ \chi_n^*\left[\mathbf{x}\right]\chi_{n'}\left[2\left(\mathbf{R}'-\mathbf{R}\right)-\mathbf{x}\right] \\ &\times\left[\Psi_{\downarrow}^{\dagger}(2\mathbf{R}'-\mathbf{R}-\frac{\mathbf{x}}{2})\Psi_{\downarrow}(\mathbf{R}-\frac{\mathbf{x}}{2})+\Psi_{\uparrow}^{\dagger}(2\mathbf{R}'-\mathbf{R}-\frac{\mathbf{x}}{2})\Psi_{\uparrow}(\mathbf{R}-\frac{\mathbf{x}}{2})\right] \qquad (97)\end{aligned}$$

(a more symmetrical form of the right-hand side can be obtained by again using the notation $\mathbf{z}=\mathbf{R}'-\mathbf{R}-\mathbf{x}$). The commutator is thus equal to that of elementary bosons plus a correction term. This latter term plays an important role over a distance $\mathbf{R}'-\mathbf{R}$, of the order of the range of the wave functions χ, and is the sum of contributions independent of the two spin states.

Conclusion

In conclusion, note that the pair field operator provides interesting insights concerning the physical properties of paired states in a many-body system. For a N-particle state, built from a two-particle wave function χ, it leads to a new wave function ϕ_{pair} when the particle indistinguishability is taken into account. In the framework of the BCS theory, we will see how this pair wave function allows characterizing the cooperative effects of pair interactions. Introducing an order parameter is also useful for showing the link between anomalous average values (which do not conserve particle number) and the normal average values. The results take on different forms for paired states of bosons or fermions. There is, however, a strong analogy between the two cases, which provides a unified framework for the study of different phenomena, such as Bose-Einstein condensation of particles or pairs.

Complement B_{XVII}

Average energy in a paired state

1 Using states that are not eigenstates of the total particle number 1869

1-a Computation of the average values 1870

1-b A good approximation 1870

2 Hamiltonian 1871

2-a Operator expression 1871

2-b Simplifications due to pairing 1874

3 Spin 1/2 fermions in a singlet state 1874

3-a Different contributions to the energy 1874

3-b Total energy 1880

4 Spinless bosons 1881

4-a Choice of the variational state 1881

4-b Different contributions to the energy 1882

4-c Total energy 1887

In Chapter XVII, the paired states were introduced in a general way, without specifying any particular form of the Hamiltonian. In order to use the paired states $|\Psi_{\text{paired}}\rangle$ in the framework of a variational method, i.e. to be able to minimize the average value of the energy of an N-particle system, we must compute the average value of the energy in these paired states; this is the purpose of this complement. We start (§ 1) by examining the consequences of the fact that these states are not eigenvectors of the total particle number operator $\widehat{N}$. In § 2, we clarify the notation and give the expression of the Hamiltonian $\widehat{H}$. We then deal successively with the fermion case (§ 3) and the boson case (§ 4). This second case is slightly more complicated since it requires the adjunction of a specific state to describe the condensate.

1. Using states that are not eigenstates of the total particle number

The paired states $|\Psi_{\text{paired}}\rangle$ are coherent superpositions of states containing different numbers of particles. One may wonder how the average values computed in such states can be relevant for a physical system where N has a fixed value. As we already mentioned in § D of Chapter XVII, this approach is correct for large values of the average particle number, provided the operators, whose average values we are computing, conserve the particle number (i.e. commute with the total particle number $\widehat{N}$, as is the case for the Hamiltonian operator $\widehat{H}$). We are going to show in more detail that when these conditions are met, the average values do not depend on the state vector's coherences between different N values; they can thus be obtained using the paired states.

1-a. Computation of the average values

The state $|\Psi_{\text{paired}}\rangle$ defined in (B-5) of Chapter XVII is a superposition of states $|\Psi_P\rangle$ where the particle number is exactly $N = 2P$:

$$|\Psi_{\text{paired}}\rangle = \sum_{P=0}^{\infty} \frac{1}{P!} |\Psi_P\rangle \tag{1}$$

As the matrix elements of the operator $\widehat{H}$ between eigenkets of $\widehat{N}$ corresponding to different eigenvalues are zero, we have:

$$\begin{aligned} \langle\Psi_{\text{paired}}| \widehat{H} |\Psi_{\text{paired}}\rangle &= \sum_{P=0}^{\infty} \left[\frac{1}{P!}\right]^2 \langle\Psi_P| \widehat{H} |\Psi_P\rangle \\ &= \sum_{P=0}^{\infty} \left[\frac{1}{P!}\right]^2 \langle\Psi_P |\Psi_P\rangle \ E_P \end{aligned} \tag{2}$$

where E_P is the energy average value in the state $|\Psi_P\rangle$:

$$E_P = \frac{\langle\Psi_P| \widehat{H} |\Psi_P\rangle}{\langle\Psi_P |\Psi_P\rangle} \tag{3}$$

Consequently, if we define the weight distribution $\mathcal{D}(P)$ as:

$$\mathcal{D}(P) = \left[\frac{1}{P!}\right]^2 \langle\Psi_P |\Psi_P\rangle \tag{4}$$

the diagonal element of $\widehat{H}$ in $|\Psi_{\text{paired}}\rangle$ is given by:

$$\langle\Psi_{\text{paired}}| \widehat{H} |\Psi_{\text{paired}}\rangle = \sum_{P=0}^{\infty} \mathcal{D}(P) \ E_P \tag{5}$$

The average value $\left\langle \widehat{H} \right\rangle$ is then obtained by dividing this expression by the square of the norm $\langle\Psi_{\text{paired}} |\Psi_{\text{paired}}\rangle$.

In a general way, the diagonal element in $|\Psi_{\text{paired}}\rangle$ of any operator $\widehat{A}$ that commutes with $\widehat{N}$ is given by a linear combination of the average values of this operator in the states $|\Psi_P\rangle$ with the weight distribution $\mathcal{D}(P)$. As an example, for any function $F\left(\widehat{N}\right)$ of the operator $\widehat{N}$, we can write:

$$\langle\Psi_{\text{app}}| F\left(\widehat{N}\right) |\Psi_{\text{app}}\rangle = \sum_{P=0}^{\infty} \mathcal{D}(P) \ F(2P) \tag{6}$$

1-b. A good approximation

For a system with a fixed $N = 2P$ particle number, we are trying to determine the eigenvalues E_P and the kets $|\Psi_P\rangle$; the most direct method would be to vary separately

each ket $|\Psi_P\rangle$ to optimize E_P. This would lead, however, to complicated calculations. It turns out to be much more practical to vary $|\Psi_{\text{paired}}\rangle$ and optimize the corresponding energy; this leads to nearly the same results for large particle numbers, as we now explain.

We saw in § C-2 of Chapter XVII that the particle number fluctuations in a state $|\Psi_{\text{paired}}\rangle$ are very small in relative value when N is large. This means that the distribution $\mathcal{D}(P)$ has a sharp peak around a certain value P_0 of P, which determines half the average value of the particle number. Now if the energies E_P are practically constant over the width of that distribution, the Hamiltonian diagonal matrix element (2) can be written:

$$\begin{aligned}\langle\Psi_{\text{paired}}|\,\widehat{H}\,|\Psi_{\text{paired}}\rangle &\simeq E_{P_0}\sum_{P=0}^{\infty}\left[\frac{1}{P!}\right]^2\langle\Psi_P\,|\Psi_P\rangle \\ &= E_{P_0}\,\langle\Psi_{\text{paired}}\,|\Psi_{\text{paired}}\rangle \end{aligned} \tag{7}$$

Making this diagonal matrix element stationary (keeping constant the norm of $|\Psi_{\text{paired}}\rangle$) is equivalent to making E_{P_0} stationary. The optimal value obtained for this matrix element, divided by the squared norm of $|\Psi_{\text{paired}}\rangle$, yields a good approximation of the energy E_{P_0} we are looking for. Once $|\Psi_{\text{paired}}\rangle$ has been optimized in this way, it can be projected onto the various subspaces with fixed particle numbers, and therefore obtain the $|\Psi_P\rangle$, corresponding to stationary states with fixed particle numbers. In the following complements, we shall use the paired states rather than the states $|\Psi_P\rangle$ with fixed particle number.

Comment:

In the following complements, rather than optimizing the average energy, it is the difference between this average energy and the average particle number multiplied by the chemical potential μ that we shall optimize. As the two operators $\widehat{H}$ and $\widehat{N}$ commute with the total particle number, the line of reasoning we just followed also applies to that case.

2. Hamiltonian

Consider a physical system composed of fermions or bosons, placed in a cubic box of edge length L.

2-a. Operator expression

The Hamiltonian $\widehat{H}$ is the same as the one used on several occasions, for example in Complement E_{XV} (but we assume here that there is no external potential):

$$\widehat{H} = \widehat{H}_0 + \widehat{W}_{\text{int}} \tag{8}$$

The operator $\widehat{H}_0$ is the sum of the kinetic energy operators $K_0(q)$ associated with each particle q:

$$\widehat{H}_0 = \sum_q K_0(q) = \sum_q \frac{\mathbf{P}^2(q)}{2m} \tag{9}$$

and $\widehat{W}_{\rm int}$ is the sum of the interaction energies between particles:

$$\widehat{W}_{\rm int} = \frac{1}{2}\sum_{q\neq q'=1}^{N} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{10}$$

where $W_2(\mathbf{R}_q, \mathbf{R}_{q'})$ only depends on the difference $\mathbf{R}_q - \mathbf{R}_{q'}$ (translation invariance) and does not act on the spins.

We now express $\widehat{H}$ in terms of creation and annihilation operators, according to formulas established in Chapter XV. We use the basis of individual states $|u_{\mathbf{k},\nu}\rangle$, where $\mathbf{k}$ labels the momentum $\hbar\mathbf{k}$ of a plane wave that satisfies the periodic conditions in the box; the index ν labels the spin state of the particles, but if they are all in the same spin state, it can be omitted in what follows. We get:

$$\begin{aligned}\widehat{H} = &\sum_{\mathbf{k},\nu} e_k\ a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k},\nu} + \\ &+\frac{1}{2}\sum_{\mathbf{k},\nu;\mathbf{k}',\nu';\mathbf{k}'';\mathbf{k}'''} \langle 1:\mathbf{k}'',\nu;\ 2:\mathbf{k}''',\nu'|\,\widehat{W}_2(\mathbf{R}_1-\mathbf{R}_2)\,|1:\ \mathbf{k},\nu;\ 2:\mathbf{k}',\nu'\rangle \\ &\qquad\qquad \times\ a^\dagger_{\mathbf{k}'',\nu} a^\dagger_{\mathbf{k}''',\nu'} a_{\mathbf{k}',\nu'} a_{\mathbf{k},\nu}\end{aligned} \tag{11}$$

with:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{12}$$

(since the interaction potential does not act on the spins, we were able to replace the spin index ν'' associated with $\mathbf{k}''$ by the index ν, as well as the index ν''' associated with $\mathbf{k}'''$ by the index ν'). The matrix elements of $\widehat{W}_2$ appearing in (11) can be written:

$$\int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2\ W_2(\mathbf{r}_1-\mathbf{r}_2)\frac{1}{L^6}\ e^{i(\mathbf{k}-\mathbf{k}'')\cdot\mathbf{r}_1} e^{i(\mathbf{k}'-\mathbf{k}''')\cdot\mathbf{r}_2} \tag{13}$$

We make the following change of variables: $\mathbf{R} = (\mathbf{r}_1+\mathbf{r}_2)/2$ and $\mathbf{r} = \mathbf{r}_1-\mathbf{r}_2$. The integral over d^3R of the exponential yields the Kronecker delta function:

$$\frac{1}{L^3}\int \mathrm{d}^3 R\ e^{i\left(\mathbf{k}+\mathbf{k}'-\mathbf{k}''-\mathbf{k}'''\right)\cdot\mathbf{R}} = \delta_{\mathbf{k}+\mathbf{k}',\mathbf{k}''+\mathbf{k}'''} \tag{14}$$

which enforces the conservation of the total momentum:

$$\mathbf{k}+\mathbf{k}' = \mathbf{k}''+\mathbf{k}''' \tag{15}$$

The integral over $\mathrm{d}^3 r$ introduces the Fourier transform $V_{\mathbf{q}}$ of the potential[1]:

$$V_{\mathbf{q}} = \frac{1}{L^3}\int \mathrm{d}^3 r\ e^{-i\mathbf{q}\cdot\mathbf{r}}\ W_2(\mathbf{r}) \tag{16}$$

with:

$$\mathbf{q} = \frac{(\mathbf{k}''-\mathbf{k}) - (\mathbf{k}'''-\mathbf{k}')}{2} \tag{17}$$

[1]The factor $1/L^3$ in (16) comes from the normalization of the plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}/L^{3/2}$ in a cube of edge length L; it ensures the potential V has the dimension of an energy.

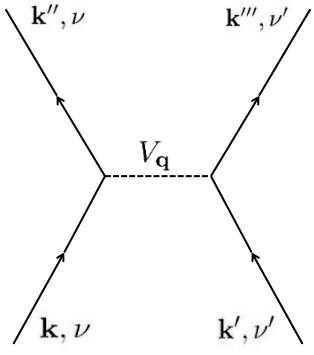


Figure 1: Symbolic plot of a general interaction process where two particles of momenta $\hbar\mathbf{k}$ and $\hbar\mathbf{k}'$ are replaced, as the result of their mutual interaction, by particles of momenta $\hbar\mathbf{k}''$ and $\hbar\mathbf{k}'''$. The indices ν and ν' label the spins, which are not modified by the interaction. The horizontal line represents the momentum transfer $\hbar\mathbf{q}$ whose value is given by (17) and (18).

or else, taking (15) into account:

$$\mathbf{q} = \mathbf{k}'' - \mathbf{k} = \mathbf{k}' - \mathbf{k}''' \tag{18}$$

The momentum transfer $\mathbf{q}$ gives the momentum variation of particle 1, as well as the opposite of the momentum variation of particle 2. Since $W_2(\mathbf{r}_1 - \mathbf{r}_2)$ is symmetric with respect to the exchange of the variables $\mathbf{r}_1$ and $\mathbf{r}_2$, the functions $W_2(\mathbf{r})$ and $V_{\mathbf{q}}$ are both even and real.

The matrix element of the interaction potential is then:

$$\langle 1 : \mathbf{k}'', \nu;\ 2 : \mathbf{k}''', \nu' | \, \widehat{W}_2(\mathbf{R}_1, \mathbf{R}_2) \, | 1 :\ \mathbf{k}, \nu;\ 2 : \mathbf{k}', \nu' \rangle = \delta_{\mathbf{k}+\mathbf{k}', \mathbf{k}''+\mathbf{k}'''}\ V_{\mathbf{q}} \tag{19}$$

and is schematized in Figure 1, where the horizontal line represents the momentum transfer $\hbar\mathbf{q}$ resulting from the interaction between the ingoing and outgoing particles. The interaction potential operator can thus be written:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''}{}' V_{\mathbf{q}} \sum_{\nu, \nu'} a^{\dagger}_{\mathbf{k}'', \nu} a^{\dagger}_{\mathbf{k}''', \nu'} a_{\mathbf{k}', \nu'} a_{\mathbf{k}, \nu} \tag{20}$$

where the summation over the $\mathbf{k}$ actually concerns only three wave vectors, since $\mathbf{k}''' = \mathbf{k} + \mathbf{k}' - \mathbf{k}''$.

In a frequently used approximation, one assumes the interaction potential range to be very small compared with the de Broglie wavelengths of all the particles involved

(contact potential). The variations with $\mathbf{k}$ of $\overline{W}_2(\mathbf{k})$ can then be neglected, and all the matrix elements of the interaction potential are equal to a given constant $V_\mathbf{0}$ (provided they conserve the total momentum; otherwise, they are obviously zero):

$$V_\mathbf{0} = \frac{1}{L^3}\int \mathrm{d}^3 r\; W_2(\mathbf{r}) \tag{21}$$

2-b. Simplifications due to pairing

In general, the computation of the average value of the operator (11) is very complex, due to the large number of possible interaction terms. However, as we already saw in § D-1-a of Chapter XVII, some simplifications occur for a paired state. The main reason is that in the various components of a paired state on Fock states, all the paired individual states have the same population. If the population of an individual state $\mathbf{k}$ changes, the population of the individual state $-\mathbf{k}$ must change by the same quantity, otherwise the average value of the operator is zero. To get a non-zero average value in a paired state, the combination of creation and annihilation operators in the considered interaction term must respect this parity condition.

Now the interaction operator (20) is a sum of terms containing two annihilation operators on the right, and two creation operators on the left. Only two possibilities exist for the population balance of all the pairs to be conserved upon the action of these four operators: either the two creation operators re-establish the initial populations of the two states that were depopulated by the annihilation operators (in which case none of the populations are changed); or else, the two annihilation operators destroy particles in the same pair of states, and the creation operators produce another pair (in which case the population of the first pair[2] is lowered by 2, and the population of the second increased by 2). The two possibilities are combined in the particular case where the creation operators restore precisely the pair of particles destroyed by the annihilation operators. We are then led to the different cases examined in detail in § D-1-a of Chapter XVII: Case **I** (direct and exchange forward scattering terms), Case **II** (pair annihilation-creation terms) and Case **III** (combination of the two previous terms, yielding a negligible contribution).

3. Spin 1/2 fermions in a singlet state

We now compute the average value $\langle H\rangle$ of the operator $\widehat{H}$, written in (11), in the state $|\Psi_{\mathrm{BCS}}\rangle$ defined in § B-2-b of Chapter XVII. As far as the interaction energy is concerned, we will show that the terms associated with Case **I** only yield the usual mean field contributions, already discussed in the previous chapters. On the other hand, the terms associated with Case **II** are a direct consequence of the pairing, and are therefore totally new; they play a leading role in the BCS theory. The terms associated with Case **III**, being a particular case of the other two cases, generally play a negligible role.

3-a. Different contributions to the energy

The different contributions to the energy will be computed successively, starting with the kinetic energy.

[2] We defined in § C-2 of Chapter XVII, the pair population operator $\hat{n}_{\mathrm{pair}}$ as the sum of the population operators of each of the two individual states forming the pair.

α. Kinetic energy

The first term (kinetic energy) is, as for the particle number, the sum of the contributions of the pairs of states, labeled by $\mathbf{k}$ (each of the two states having the same kinetic energy):

$$\begin{aligned}\langle H_0\rangle &= \sum_{\mathbf{k}} e_k \ \langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{(\text{pair }\mathbf{k})}\,|\overline{\varphi}_{\mathbf{k}}\rangle = 2\sum_{\mathbf{k}} e_k \ \left|v\left(\mathbf{k}\right)\right|^2 \\ &= 2\sum_{\mathbf{k}} e_k \ \sin^2\theta_{\mathbf{k}} \end{aligned} \tag{22}$$

β. Interaction energy

The average of the interaction potential energy is the sum of the averages of the terms on the right-hand side of (20), i.e. of terms that belong to one of the three possibilities **I**, **II** and **III** cited above; we study them successively.

– **Case I (the creation operators repopulate the states depopulated by the annihilation operators)**

For such terms, the occupation numbers of each individual state remain unchanged in the course of the interaction process. They are "diagonal terms" (sometimes called "mean field terms"). Two cases may arise, depending on whether the spin index ν is the same as, or different from ν'; we examine each of these possibilities in turn.

(i) If $\nu' = -\nu$, as the interaction potential does not act on the spin, we can trace each particle using its spin direction; it is as if the particles were distinguishable. If the creation operators repopulate exactly the same individual states depopulated by the annihilation operators, the only possible interaction is schematized in Figure 2, and corresponds to a *forward scattering*. As the momentum transfer $\mathbf{q}$ is zero, the potential term includes the constant $V_{\mathbf{0}}$, and we get the following contribution to the average energy:

$$\frac{V_{\mathbf{0}}}{2}\sum_{\mathbf{k}\neq-\mathbf{k}',\nu}\left\langle\overline{\Psi}_{\text{BCS}}\right|\left(a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',-\nu}a_{\mathbf{k}',-\nu}a_{\mathbf{k},\nu}\right)\left|\overline{\Psi}_{\text{BCS}}\right\rangle \tag{23}$$

(the condition $\mathbf{k}\neq-\mathbf{k}'$ comes from the fact that the pairs are different, each pair being labeled by the value of $\mathbf{k}$ associated with the spin $+$). Two anticommutations permit bringing the last operator $a_{\mathbf{k},\nu}$ right after the first one $a^{\dagger}_{\mathbf{k},\nu}$ (with two sign changes that cancel each other). If we now sum all the contributions from $\nu=+$ and from $\nu=-$, we get:

$$\begin{aligned}\frac{V_{\mathbf{0}}}{2}\sum_{\mathbf{k}\neq-\mathbf{k}'}&\left[\langle\overline{\varphi}_{\mathbf{k}}|\,a^{\dagger}_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}\,|\overline{\varphi}_{\mathbf{k}}\rangle\left\langle\overline{\varphi}_{-\mathbf{k}'}\right|a^{\dagger}_{\mathbf{k}',\downarrow}a_{\mathbf{k}',\downarrow}\left|\overline{\varphi}_{-\mathbf{k}'}\right\rangle\right.\\ &\left.+\left\langle\overline{\varphi}_{-\mathbf{k}}\right|a^{\dagger}_{\mathbf{k},\downarrow}a_{\mathbf{k},\downarrow}\left|\overline{\varphi}_{-\mathbf{k}}\right\rangle\langle\overline{\varphi}_{\mathbf{k}'}|\,a^{\dagger}_{\mathbf{k}',\uparrow}a_{\mathbf{k}',\uparrow}\,|\overline{\varphi}_{\mathbf{k}'}\rangle\right]\end{aligned} \tag{24}$$

We can show that the two terms inside the brackets yield the same contribution by interchanging the two dummy indices $\mathbf{k}$ and $\mathbf{k}'$ in the summation. We thus double the first term, and after changing the sign of $\mathbf{k}'$, we get:

$$\begin{aligned}V_{\mathbf{0}}\sum_{\mathbf{k}\neq\mathbf{k}'}\langle\overline{\varphi}_{\mathbf{k}}|\,a^{\dagger}_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}\,|\overline{\varphi}_{\mathbf{k}}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a^{\dagger}_{-\mathbf{k}',\downarrow}a_{-\mathbf{k}',\downarrow}\,|\overline{\varphi}_{\mathbf{k}'}\rangle &= V_{\mathbf{0}}\sum_{\mathbf{k}\neq\mathbf{k}'}\left|v\left(\mathbf{k}\right)\right|^2\left|v\left(\mathbf{k}'\right)\right|^2\\ &= V_{\mathbf{0}}\sum_{\mathbf{k}\neq\mathbf{k}'}\sin^2\theta_{\mathbf{k}}\ \sin^2\theta_{\mathbf{k}'}\end{aligned} \tag{25}$$

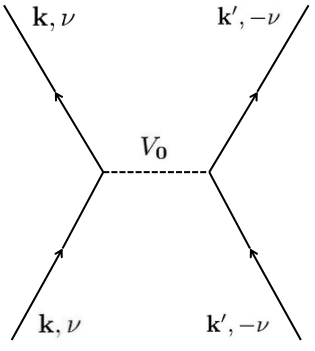


Figure 2: Schematic plot of the interaction between particles of opposite spins, which do not belong to the same pair (forward scattering). This diagram contributes to the particles' mean field.

When the particles are distributed in a large number of individual states, the value of the summation in the above expression is barely changed if we ignore the constraint $\mathbf{k} \neq \mathbf{k}'$. If we now use for $\langle N \rangle$ the expression (C-19) of Chapter XVII, we can write this contribution as:

$$\frac{V_0}{4} \langle N \rangle^2 \tag{26}$$

According to relation (21), the constant V_0 is proportional to the inverse of the volume L^3. This term can be interpreted as a mean field term, where $\langle N \rangle /2$ particles with a spin + interact with $\langle N \rangle /2$ particles having a spin −; a particle with a given spin direction feels the mean field exerted by all the particles with opposite spin, whose numerical density is $\langle N \rangle /2L^3$.

(ii) if $\nu' = \nu$, it is no longer possible to distinguish the particles by the direction of their spin, and the indistinguishability effects play their full role. Two cases must be distinguished for these "diagonal terms": either $\mathbf{k}'' = \mathbf{k}$ and $\mathbf{k}''' = \mathbf{k}'$, which yields a *direct term*; or $\mathbf{k}''' = \mathbf{k}$ and $\mathbf{k}'' = \mathbf{k}'$, which yields an *exchange term.* In both cases, the individual states populated in the bra and the ket are the same, and we are dealing with "diagonal processes" that can be called "mean field terms".

For the direct term, no particle changes its momentum, which again corresponds to a "forward scattering" (left-hand side of Figure 3), and the potential term again includes the constant V_0 written in (21). The average value of this direct term is:

$$\frac{V_0}{2} \sum_{\mathbf{k} \neq \mathbf{k}', \nu} \left\langle \overline{\Psi}_{\mathrm{BCS}} \right| a^\dagger_{\mathbf{k},\nu} a^\dagger_{\mathbf{k}',\nu} a_{\mathbf{k}',\nu} a_{\mathbf{k},\nu} \left| \overline{\Psi}_{\mathrm{BCS}} \right\rangle \tag{27}$$

Here again, since $\mathbf{k} \neq \mathbf{k}'$ (otherwise we would have the square of an annihilation operator, which is zero), two anticommutations let us bring the operator $a_{\mathbf{k},\nu}$ to the second position,

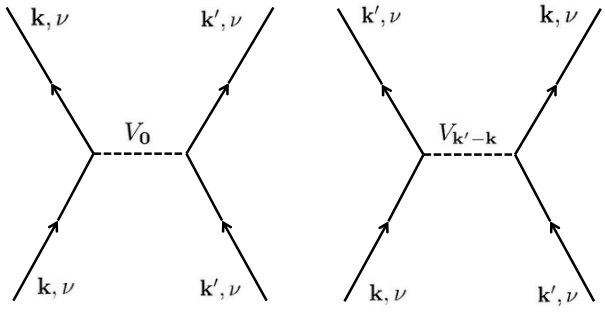


Figure 3: Interaction between particles having the same spin; the direct term (forward scattering) is schematized on the left, and the exchange term on the right. These two diagrams add their contributions to the diagram of Figure 2 to build the particles' mean field.

and we get:

$$\frac{V_0}{2}\sum_{\mathbf{k}\neq\mathbf{k}',\nu}\langle\overline{\varphi}_{\mathbf{k}}|\,a^{\dagger}_{\mathbf{k},\nu}a_{\mathbf{k},\nu}\,|\overline{\varphi}_{\mathbf{k}}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a^{\dagger}_{\mathbf{k}',\nu}a_{\mathbf{k}',\nu}\,|\overline{\varphi}_{\mathbf{k}'}\rangle = V_0\sum_{\mathbf{k}\neq\mathbf{k}'}|v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2$$
$$= V_0\sum_{\mathbf{k}\neq\mathbf{k}'}\sin^2\theta_{\mathbf{k}}\;\sin^2\theta_{\mathbf{k}'} \tag{28}$$

(the two values of ν yield the same contribution, hence the disappearance of the factor $1/2$ on the right-hand side). As for the exchange term, we have $\mathbf{k}''' = \mathbf{k}$ and $\mathbf{k}'' = \mathbf{k}'$ (right-hand side of Figure 3); for such a momentum exchange, the transfer $\mathbf{q}$ is no longer zero, but equal to:

$$\mathbf{q} = \mathbf{k}'-\mathbf{k} \tag{29}$$

and the potential term now includes $V_{\mathbf{k}'-\mathbf{k}}$ obtained by inserting $\mathbf{q} = \mathbf{k}'-\mathbf{k}$ in (16). Furthermore, when $\mathbf{k} \neq \mathbf{k}'$:

$$a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',\nu}a_{\mathbf{k},\nu}a_{\mathbf{k}',\nu} = -a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',\nu}a_{\mathbf{k}',\nu}a_{\mathbf{k},\nu} \tag{30}$$

Apart from this sign change, the computation is the same as for the direct term. The sum of the two direct plus the exchange contributions finally yields:

$$\sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_0 - V_{\mathbf{k}'-\mathbf{k}}\right]|v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 = \sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_0 - V_{\mathbf{k}'-\mathbf{k}}\right]\;\sin^2\theta_{\mathbf{k}}\;\sin^2\theta_{\mathbf{k}'} \tag{31}$$

In the short-range potential approximation where $V_{\mathbf{k}'-\mathbf{k}} = V_0$, this sum is zero: the Pauli exclusion principle prevents particles having the same spin components from interacting via a contact potential.

– **Case II (particles annihilated in a pair of states and restored in another pair)**

Considering the nature of the creation and annihilation operators it contains, this process may be called "pair annihilation-creation". It plays an essential role in the BCS pairing, as we shall see in Complement C_{XVII}; the corresponding term in the Hamiltonian is thus often called the "pairing term".

We then have, on one side $\mathbf{k}' = -\mathbf{k}$ and $\nu' = -\nu$, and on the other, $\mathbf{k}''' = -\mathbf{k}''$, so that, according to its definition (17), the momentum transfer is $\mathbf{q} = \mathbf{k}'' - \mathbf{k}$; the corresponding diagram is shown in Figure 4. We are going to show below that its contribution to the energy can be written as:

$$\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}} \ \sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}} \ \sin\theta_{\mathbf{k}''}\cos\theta_{\mathbf{k}''} \ e^{2i(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}''})} \tag{32}$$

This term is new, in the sense that it is not a mean field term, like the previous ones, but that its existence is due to the pairing process. We will show in Complement C_{XVII} that its contribution to the average energy plays an essential role in the BCS theory.

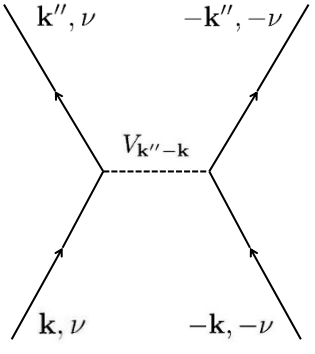


Figure 4: Interaction process between two particles in the same pair, which, in their final states, end up in another pair. In terms of creation and annihilation operators, this process is a "pair annihilation-creation" (two particles of the same pair are annihilated, while two particles are created in another pair). As opposed to the terms introduced by the other interaction processes, this term's contribution to the energy depends on the pairing; it is sometimes called the "pairing term", and is responsible for the energy gain in the BCS theory (Complement C_{XVII}).

Demonstration:

If $\nu = +$, the contribution contains a product of "anomalous" average values:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \left\langle \overline{\Psi}_{\rm BCS}\right| a^{\dagger}_{\mathbf{k}'',\uparrow}a^{\dagger}_{-\mathbf{k}'',\downarrow}a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}\left|\overline{\Psi}_{\rm BCS}\right\rangle$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \langle\overline{\varphi}_{\mathbf{k}''}|\, a^{\dagger}_{\mathbf{k}'',\uparrow}a^{\dagger}_{-\mathbf{k}'',\downarrow}\,|\overline{\varphi}_{\mathbf{k}''}\rangle\,\langle\overline{\varphi}_{\mathbf{k}}|\, a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}\,|\overline{\varphi}_{\mathbf{k}}\rangle \tag{33}$$

that is, using (C-42) and (C-44) of Chapter XVII:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ u_{\mathbf{k}''}v^{*}_{\mathbf{k}''}u^{*}_{\mathbf{k}}v_{\mathbf{k}}$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\ \ \sin\theta_{\mathbf{k}''}\cos\theta_{\mathbf{k}''}\ e^{2i(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}''})} \tag{34}$$

If $\nu = -$, it is now the kets $\left|\overline{\varphi}_{-\mathbf{k}}\right\rangle$ and $\left|\overline{\varphi}_{-\mathbf{k}''}\right\rangle$ that come into play, and we obtain another product of anomalous average values for which we must use (C-43) and (C-45) of Chapter XVII (as well as the fact that the functions of $\mathbf{k}$ are even, as indicated in that chapter):

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \left\langle \overline{\Psi}_{\rm BCS}\right| a^{\dagger}_{\mathbf{k}'',\downarrow}a^{\dagger}_{-\mathbf{k}'',\uparrow}a_{-\mathbf{k},\uparrow}a_{\mathbf{k},\downarrow}\left|\overline{\Psi}_{\rm BCS}\right\rangle$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ v^{*}_{-\mathbf{k}''}u_{-\mathbf{k}''}v_{-\mathbf{k}}u^{*}_{-\mathbf{k}} \tag{35}$$

This expression is the same as the previous one, since it only differs by the sign of the summation dummy indices $\mathbf{k}$ and $\mathbf{k}''$ (remember that $V_{\mathbf{q}}$ is even). We therefore remove the factor 1/2 in (34) and get (32).

– **Case III (particles annihilated in a pair of states, then restored in the same pair)**

We then have again $\mathbf{k}' = -\,\mathbf{k}$ and $\nu' = -\nu$, but in addition $\mathbf{k}'' = \mathbf{k}$ (and hence necessarily $\mathbf{k}''' = \mathbf{k}'$), as shown in Figure 5; this is another case of forward scattering.

We now check that this term can be neglected. Its contribution to the energy is:

$$\frac{V_0}{2}\sum_{\mathbf{k},\nu}\ \left\langle \overline{\Psi}_{\rm BCS}\right| a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{-\mathbf{k},-\nu}a_{-\mathbf{k},-\nu}a_{\mathbf{k},\nu}\left|\overline{\Psi}_{\rm BCS}\right\rangle \tag{36}$$

If $\nu = +$, we get (after two operator anticommutations):

$$\frac{V_0}{2}\sum_{\mathbf{k}}\langle\overline{\varphi}_{\mathbf{k}}|\left(a^{\dagger}_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}\right)\left(a^{\dagger}_{-\mathbf{k},\downarrow}a_{-\mathbf{k},\downarrow}\right)|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{V_0}{2}\sum_{\mathbf{k}}|v_{\mathbf{k}}|^2 \tag{37}$$

and if $\nu = -$:

$$\frac{V_0}{2}\sum_{\mathbf{k}}\left\langle\overline{\varphi}_{-\mathbf{k}}\right|\left(a^{\dagger}_{\mathbf{k},\downarrow}a_{\mathbf{k},\downarrow}\right)\left(a^{\dagger}_{-\mathbf{k},\uparrow}a_{-\mathbf{k},\uparrow}\right)\left|\overline{\varphi}_{-\mathbf{k}}\right\rangle = \frac{V_0}{2}\sum_{\mathbf{k}}|v_{-\mathbf{k}}|^2 \tag{38}$$

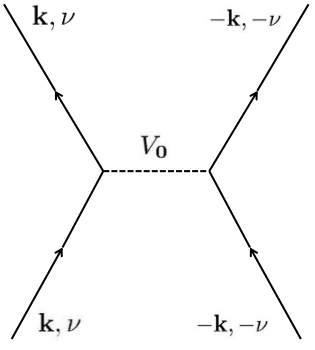


Figure 5: Interaction process where two particles of the same pair are scattered in the forward direction.

This term is the same as the previous one, as it only differs by the sign of the summation dummy index $\mathbf{k}$. Taking into account expression (C-19) of Chapter XVII for $\left\langle \hat{N} \right\rangle$, we can write the total contribution as:

$$V_0 \sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = \frac{V_0}{2} \left\langle \hat{N} \right\rangle \tag{39}$$

This contribution is interpreted as the average attraction energy in an ensemble of $\left\langle \hat{N} \right\rangle /2$ pairs. When the average particle number is large, we can neglect (39) compared to (26). Consequently, the pairing effects we are going to discuss cannot be simply interpreted as an attraction among an ensemble of $N/2$ pairs.

3-b. Total energy

Finally, adding the terms (22), (31), (26) and the double of (34), we get the average energy[3]:

$$\langle H \rangle = 2 \sum_{\mathbf{k}} e_k \ \sin^2 \theta_{\mathbf{k}} + \frac{V_0}{4} \langle N \rangle^2 + \sum_{\mathbf{k} \neq \mathbf{k}'} [V_0 - V_{\mathbf{k}'-\mathbf{k}}] \sin^2 \theta_{\mathbf{k}} \sin^2 \theta_{\mathbf{k}'} + \sum_{\mathbf{k},\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}} \ \sin \theta_{\mathbf{k}} \sin \theta_{\mathbf{k}''} \cos \theta_{\mathbf{k}} \cos \theta_{\mathbf{k}''} e^{2i(\zeta_{\mathbf{k}} - \zeta_{\mathbf{k}''})} \tag{40}$$

[3]The summations over $\mathbf{k}$ have no restrictions, contrary to the tensor product appearing in relation (B-8) of Chapter XVII, where the summation is limited to a half-space to avoid redundancy.

The first term on the right-hand side corresponds to the kinetic energy, the second to the mean field for particles of opposite spins, the third one is the analogous term for particles having the same spin (it goes to zero for a short-range potential); these three terms were already present in the Hartree-Fock theory. The fourth term, however, is new: it corresponds to the pair annihilation-creation (pairing term) whose average value is non-zero only in a paired state. It is the only one that depends on the phases $\zeta_{\mathbf{k}}$, which will prove to be essential in the BCS theory (Complement C_{XVII}).

4. Spinless bosons

For bosons, we must take into account the Bose-Einstein condensation phenomenon (Complements B_{XV}, C_{XV} and F_{XV}): in the ground state, a large fraction of the particles can occupy a single quantum state, the state $\mathbf{k} = 0$. This is not the case for a paired state; we must therefore choose a variational state permitting such a condensation.

We assume the interactions to be repulsive, in order to avoid the instabilities occurring for a system of attractive bosons (Complement H_{XV}, § 4-b).

4-a. Choice of the variational state

In Complement C_{XV}, we used the Gross-Pitaevskii approximation to treat, in the simplest way, Bose-Einstein condensation: the system of N bosons is supposed to be, at a given instant, in a state that is the product of identical individual states, generally chosen as the zero momentum state, $\mathbf{k} = 0$; the system state is thus written as:

$$|\Phi\rangle = \left[a_0^\dagger\right]^N |0\rangle \tag{41}$$

($a_0^\dagger$ is the creation operator in the individual state $\mathbf{k} = 0$). However, whereas such a state is suitable for an ideal gas ground state, it can only be an approximation for a gas of interacting particles: it is an eigenvector of the kinetic energy, but not of the operator associated with the interaction energy. The interaction potential actually couples this state to all the states where two particles are transferred from the individual state $\mathbf{k} = 0$ toward any two individual states of opposite momenta $\hbar\mathbf{k}$ and $-\hbar\mathbf{k}$ (because of momentum conservation), such as, for example, the state:

$$|\Phi'\rangle = \left[a_{\mathbf{k}}^\dagger\right]\left[a_{-\mathbf{k}}^\dagger\right]\left[a_0^\dagger\right]^{N-2} |0\rangle \tag{42}$$

where two states of a pair are occupied. This suggests using a state $|\Psi_{\text{paired}}\rangle$ as a variational ket[4] for describing the components of the system state vector associated with all the individual states $\mathbf{k} \neq 0$. We must also include the components corresponding to the individual state $\mathbf{k} = 0$; those will be described[5] by a "coherent state" (Complement G_V).

[4] The interaction potential also couples a state such as (42) to numerous states of the form $\left[a_{\mathbf{k}+\mathbf{q}}^\dagger\right]\left[a_{-\mathbf{q}}^\dagger\right]\left[a_{-\mathbf{k}}^\dagger\right]\left[a_0^\dagger\right]^{N-3}|0\rangle$, where $\mathbf{q}$ can take on any value. An exact theory would require taking those "unpaired" states into account, but leads to complex calculations. This is why we limit ourselves to a variational method in the framework of an approximation where the states $\mathbf{k} \neq 0$ are only accessible to pairs (we assume $N - N_0 \ll N$).

[5] This individual state must be treated separately, as applying the general formula (B-9) of Chapter XVII, used when $\mathbf{k} \neq \mathbf{0}$, to obtain $|\varphi_{\mathbf{k}=\mathbf{0}}\rangle$ would involve the exponential of the square of the operator

We therefore choose a variational state vector of the form:

$$|\Phi_B\rangle = |\varphi_{\mathbf{0}}\rangle \otimes |\Psi_{\text{paired}}\rangle = |\varphi_{\mathbf{0}}\rangle \otimes \prod_{\otimes \mathbf{k}\in D} |\overline{\varphi}_{\mathbf{k}}\rangle \tag{43}$$

(the notation B refers to the name Bogolubov).

In this expression, $|\Psi_{\text{paired}}\rangle$ is the paired state for spinless particles (B-8) of Chapter XVII, a tensor product of the normalized states $|\overline{\varphi}_{\mathbf{k}}\rangle$ defined in (B-9) and (C-13). The domain D of the tensor product in (43) is half the $\mathbf{k}$-space to avoid (as seen previously) a double appearance of each state $|\overline{\varphi}_{\mathbf{k}}\rangle$; the origin $\mathbf{k} = 0$ is excluded from D. This domain could eventually have an upper bound for $\mathbf{k}$.

As for $|\varphi_{\mathbf{0}}\rangle$, it is the coherent state whose expression can be found, for example, in Complement G_V, whose relations (65) and (66) provide[6]:

$$|\varphi_{\mathbf{0}}\rangle = e^{-N_0/2}\ e^{\alpha_0 a_0^\dagger}\,|n_0 = 0\rangle \tag{44}$$

This state depends on a complex parameter α_0, characterized by its modulus $\sqrt{N_0}$ and its phase ζ_0:

$$\alpha_0 = \sqrt{N_0} e^{i\zeta_0} \tag{45}$$

It is a normalized eigenvector of the operator a_0 with the eigenvalue α_0:

$$a_0\,|\varphi_{\mathbf{0}}\rangle = \alpha_0\,|\varphi_{\mathbf{0}}\rangle \tag{46}$$

The average particle number in the state $\mathbf{k} = 0$ is thus:

$$\langle\varphi_{\mathbf{0}}|\,a_0^\dagger a_0\,|\varphi_0\rangle = \alpha_0^* \alpha_0\,\langle\varphi_{\mathbf{0}}\,|\varphi_{\mathbf{0}}\rangle = N_0 \tag{47}$$

The width of the corresponding distribution is $\sqrt{N_0}$ (Complement G_V), hence negligible compared to N_0 (supposed to be a large number).

The variational variables contained in the trial ket (43) are thus the set of $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$, as well as N_0 and ζ_0.

4-b. Different contributions to the energy

We now compute the average energy, in the variational state $|\Phi_B\rangle$ written in (43), of the Hamiltonian operator given by (11).

$a^\dagger_{\mathbf{k}=\mathbf{0}}$, leading to large fluctuations of the particle number in the state $\mathbf{k} = 0$ (condensed particles). This would necessarily yield large fluctuations of the total particle number, as well as of the average repulsive energy, whereas, as we saw in § 3-b-β of Complement G_{XV}, those fluctuations are not possible precisely because of this repulsion.

[6] One shold be careful about the change of notation: in Chapter V and its complements, $|\varphi_0\rangle$ denotes the ground state of the one particle harmonic oscillator, which here corresponds to the vacuum $|n_0 = 0\rangle$. In the present complement, $|\varphi_{\mathbf{0}}\rangle$ is the ket associated with a large number of particles occupying the same individual state, as is also the case of the wave function $\varphi(\mathbf{r})$ of the Gross-Pitaevskii equation (Complement C_{XV}); with the notation of Complement G_V, this ket would rather correspond to a $|\alpha_0\rangle$ state.

α. Kinetic energy

The kinetic energy term is the sum of the contributions from the different individual states $\mathbf{k}$, with no contribution from the $\mathbf{k} = 0$ state (since $e_{k=0} = 0$). Each term of the summation contains the operator $\widehat{n}_{\mathbf{k}}$, whose average value in the factored (over the $\mathbf{k}$) state (43) is given by the average value $\langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle$ in the state $|\overline{\varphi}_{\mathbf{k}}\rangle$. This average value is given by relation (C-33) of Chapter XVII as $\sinh^2\theta_{\mathbf{k}}$, which yields, for the average value E_c of the kinetic energy in the state $|\Phi_B\rangle$ the expression:

$$E_c = \sum_{\mathbf{k}\neq 0} e_k \;\langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = \sum_{\mathbf{k}\neq 0} e_k \,\sinh^2\theta_{\mathbf{k}} \tag{48}$$

with:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{49}$$

where m is the particle mass.

β. Interaction energy

The average value of the interaction potential energy is a sum over four indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ of the potential matrix elements described in (19). As opposed to what happened for the kinetic energy, these elements have no particular reason to cancel out if one (or several) of their indices is zero. We shall therefore compute the different contributions, arranging them in decreasing order of the number of their zero indices.

A noticeable simplification of the computation occurs with the choice of the trial vector, as the coherent state $|\varphi_{\mathbf{0}}\rangle$ is one of its factors. Any time one of the four indices in the potential energy term is zero, the corresponding annihilation operator may be replaced by the complex number α_0. This is because the trial ket $|\Phi_B\rangle$ is an eigenvector of the operator a_0 with eigenvalue α_0 - see relation (46). In the same way, each time one of the two indices $\mathbf{k}$ or $\mathbf{k}'$ is zero, the creation operators on the left of the product, and hence acting on the bra $\langle\Phi_B|$, can be replaced by α_0^*, since the Hermitian conjugate of relation (46) is:

$$\langle\varphi_{\mathbf{0}}|\,a_0^\dagger = \alpha_0^*\,\langle\varphi_{\mathbf{0}}| \tag{50}$$

These two operators are therefore simply replaced by numbers. Let us examine in turn all the possible cases.

(i) If the four indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ are zero, the interaction potential contributes via the constant $V_{\mathbf{0}}$ (forward scattering term), defined in (16) as the integral of the interaction potential $W_2(\mathbf{r})$; this contribution is written as:

$$\left\langle\widehat{W}_{\text{forward}}\right\rangle_{0,0} = \frac{V_{\mathbf{0}}}{2}\,\langle\varphi_{\mathbf{0}}|\,a_0^\dagger a_0^\dagger a_0 a_0\,|\varphi_{\mathbf{0}}\rangle = V_{\mathbf{0}}\frac{(N_0)^2}{2} \tag{51}$$

The corresponding term is represented in Figure 6.

There is no contribution from terms where three (and only three) indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ are zero: total momentum conservation would require the fourth index to also be zero.

(ii) If among the four indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ two are zero, one concerning an annihilation operator, the other a creation operator, the momentum conservation requires the

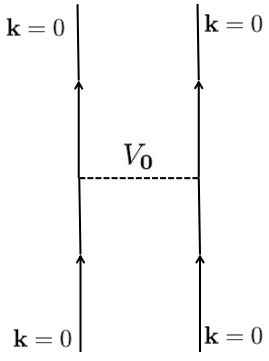


Figure 6: Diagram symbolizing the interactions between particles in the individual state $\mathbf{k}=0$*, and which remain in that state after the interaction (forward scattering term that yields the internal mean field of the condensate).*

other two operators to be $a_{\mathbf{k}}^{\dagger}$ and $a_{\mathbf{k}}$, with the same index $\mathbf{k}$. This can yield either a direct term, or an exchange term.

- the direct terms contain the average value of either the product $a_{\mathbf{k}}^{\dagger}a_0^{\dagger}a_0a_{\mathbf{k}}$, or of the product $a_0^{\dagger}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}a_0$; it is again a forward scattering process and the potential appears via the constant $V_{\mathbf{0}}$. The two average values can be factored into two terms $\langle\varphi_{\mathbf{0}}|\, a_0^{\dagger}a_0\,|\varphi_{\mathbf{0}}\rangle = |\alpha_0|^2$ and $\langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = \sinh^2\theta_{\mathbf{k}}$; they are thus equal and the corresponding contribution is written as:

$$\left\langle \widehat{W}_{\text{direct}} \right\rangle_{0,e} = V_{\mathbf{0}}N_0 \sum_{\mathbf{k}\neq 0} \sinh^2\theta_{\mathbf{k}} \tag{52}$$

where the subscript e symbolizes the ensemble of the "excited" states, i.e. those with momentum $\hbar\mathbf{k} \neq 0$ that have a non-zero kinetic energy. Introducing the average total number of particles N_e in these excited states:

$$N_e = \sum_{\mathbf{k}\neq 0} \left\langle a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}} \right\rangle = \sum_{\mathbf{k}\neq 0} \sinh^2\theta_{\mathbf{k}} \tag{53}$$

we can write:

$$\left\langle \widehat{W}_{\text{direct}} \right\rangle_{0,e} = V_{\mathbf{0}}N_0N_e \tag{54}$$

This term is simply interpreted as coming from the interaction between N_0 particles in the condensed state $\mathbf{k}=0$ and N_e particles in the other individual states.

- the exchange terms contain $a_{\mathbf{k}}^{\dagger}a_0^{\dagger}a_{\mathbf{k}}a_0$ and $a_0^{\dagger}a_{\mathbf{k}}^{\dagger}a_0a_{\mathbf{k}}$. We are now dealing with a momentum transfer process, and the potential now appears via the constant $V_{\mathbf{k}}$ obtained by inserting $\mathbf{q}=\mathbf{k}$ in (16). Otherwise, the computation is the same as for the direct

terms: the two average values can be factored, and the corresponding contribution is written:

$$\left\langle \widehat{W}_{\text{ex}} \right\rangle_{0,e} = N_0 \sum_{\mathbf{k}\neq 0} V_{\mathbf{k}} \, \sinh^2 \theta_{\mathbf{k}} \tag{55}$$

The two terms (54) and (55) correspond to mean field contributions associated with the interaction between $\mathbf{k} = 0$ particles and $\mathbf{k} \neq 0$ particles, taking into account the indistinguishability of the particles that led to the exchange term.

(iii) if the product of operators contains the annihilation operator in the state $\mathbf{k} = 0$ twice, the momentum conservation requires the product to be of the form $a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a_0 a_0$. We are now dealing with a process where two particles in the state $\mathbf{k} = 0$ are replaced by a pair $(\mathbf{k}, -\mathbf{k})$, which amounts to creating a pair from particles initially in the condensate, as shown on the left-hand side of Figure 7; here again, the potential appears via the constant $V_{\mathbf{k}}$. The two annihilation operators introduce the factor $[\alpha_0]^2 = N_0 e^{2i\zeta_0}$ and the other two operators, an "anomalous average value" in a state $|\overline{\varphi}_{\mathbf{k}}\rangle$, which we already computed in (C-51) of Chapter XVII. We therefore get:

$$-\frac{N_0}{2} \sum_{\mathbf{k}\neq 0} V_{\mathbf{k}} \, \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, e^{2i(\zeta_0 - \zeta_{\mathbf{k}})} \tag{56}$$

If the product of operators contains the creation operator in the state $\mathbf{k} = 0$ twice, it must necessarily be of the form $a^\dagger_0 a^\dagger_0 a_{\mathbf{k}} a_{-\mathbf{k}}$, which corresponds to the annihilation of a pair whose particles are transferred to the state $\mathbf{k} = 0$ (right-hand side of Figure 7). This product is the Hermitian conjugate of the previous one, and its average value is the complex conjugate of the previous result. The sum of these two terms is the contribution of the processes of creation and annihilation of pairs from the condensate:

$$-N_0 \sum_{\mathbf{k}\neq 0} V_{\mathbf{k}} \, \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\,(\zeta_0 - \zeta_{\mathbf{k}}) \tag{57}$$

These terms come from the pairing of particles, as opposed to other terms that are related to the mean field. We shall see in Complement E_{XVII} the essential role they play in the Bose-Einstein condensation of an ensemble of bosons.

(iv) There are matrix elements of the interaction potential involving a single particle in the $\mathbf{k} = 0$ individual state, and three other particles in the $\mathbf{k} \neq 0$ states. The corresponding terms have a zero average value in the state $|\Phi_B\rangle$ because of the structure of its component $|\Psi_{\text{paired}}\rangle$, where the occupation numbers of two paired states must always vary together.

(v) We have yet to compute the contribution of terms where none of the wave vectors are zero. The computation is very similar to that of § 3, and we shall again distinguish three cases:

– **Case I**

Terms containing interactions where particles are created in the states from which they were destroyed: a direct term in $a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{k}'} a_{\mathbf{k}'} a_{\mathbf{k}}$, and an exchange term in $a^\dagger_{\mathbf{k}'} a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} a_{\mathbf{k}}$. The computation is the same as in § 3, except for the fact that no minus sign occurs in

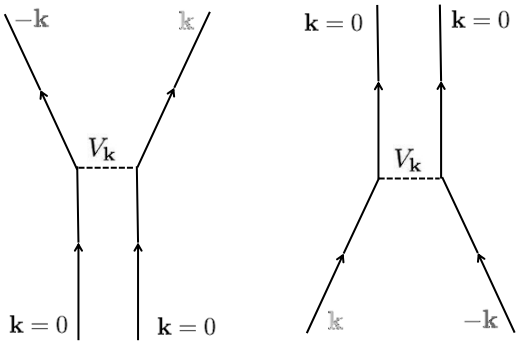


Figure 7: The diagram on the left represents a process where two particles, initially in the $\mathbf{k} = 0$ *individual state, interact and end up in states of opposite momenta. The diagram on the right represents the inverse process, where two particles of opposite momenta collide and end up in the* $\mathbf{k} = 0$ *individual state (i.e. in the condensate). As opposed to the previous terms, corresponding to the mean field of interacting particles, the terms corresponding to this diagram are introduced by the pairing process: they play a central role in the Bogolubov theory (Complement* E_{XVII}*).*

the exchange term. Result (31) therefore becomes, for bosons[7]:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_{\mathbf{0}}+V_{\mathbf{k}'-\mathbf{k}}\right]\left|v_{\mathbf{k}}\right|^2\left|v_{\mathbf{k}'}\right|^2=\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_{\mathbf{0}}+V_{\mathbf{k}'-\mathbf{k}}\right]\ \sinh^2\theta_{\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}'}$$
$$\simeq\frac{V_{\mathbf{0}}}{2}\left(N_e\right)^2+\frac{1}{2}\sum_{\mathbf{k},\mathbf{k}'}V_{\mathbf{k}'-\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}'}\qquad(58)$$

The first term is the direct term that, when $N_e \gg 1$, is interpreted as the effect of the interaction mean field between the $N_e\left(N_e-1\right)/2$ different pairs of particles (when $N_e \gg 1$). It is corrected by a second exchange term, which expresses the increased interaction between particles due to the boson bunching effect.

– **Case II**

The "pair annihilation-creation" term in $a^\dagger_{\mathbf{k}}a^\dagger_{-\mathbf{k}}a_{\mathbf{k}'}a_{-\mathbf{k}'}$, which yields here, taking

[7]As opposed to the fermion case, the contribution of the terms $\mathbf{k}=\mathbf{k}'$ is not zero but involves the average value of the operator $\widehat{n}_{\mathbf{k}}\left(\widehat{n}_{\mathbf{k}}-1\right)$ in the state $|\varphi_{\mathbf{k}}\rangle$, see § C-2-b of Chapter XVII. However, for a large system, the number of individual states $\mathbf{k}$ is very important, and this contribution is totally negligible compared to (58). This is why we neglected this term.

As for fermions, the summations over $\mathbf{k}$ do not have any restriction (no limitation to half the reciprocal space).

(C-51) and (C-52) of Chapter XVII into account:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \langle\overline{\varphi}_{\mathbf{k}''}|\, a^{\dagger}_{\mathbf{k}''}a^{\dagger}_{-\mathbf{k}''}\,|\overline{\varphi}_{\mathbf{k}''}\rangle\,\langle\overline{\varphi}_{\mathbf{k}}|\, a_{-\mathbf{k}}a_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}\ \sinh\theta_{\mathbf{k}''}\cosh\theta_{\mathbf{k}''}\ e^{2i(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}''})} \tag{59}$$

– **Case III**

Finally, the case where only one pair is involved leads, as in the fermion case, to a term proportional to $V_{\mathbf{0}}\langle N\rangle$, negligible compared to the term in $V_{\mathbf{0}}\langle N\rangle^2$ of (58) when the average particle number is large. We shall therefore neglect it.

4-c. Total energy

Regrouping the terms in $V_{\mathbf{0}}$ of (51), (54) and of (58), we get a total mean field term:

$$\left\langle \widehat{W}_{\text{mean field}}\right\rangle = \frac{V_{\mathbf{0}}}{2}\,(N_0+N_e)^2 \tag{60}$$

From a physical point of view, it is natural that this term be proportional to the square of the total particle number divided by 2, that is to the number of ways N particles can be associated by pairs (when $N \gg 1$). If we now include (55), (57) and (59), we get for the total energy:

$$\left\langle \widehat{H}\right\rangle_B = \sum_{\mathbf{k}\neq 0} e_k\ \sinh^2\theta_{\mathbf{k}} + \frac{V_{\mathbf{0}}}{2}\,(N_0+N_e)^2$$
$$+ N_0\sum_{\mathbf{k}\neq 0} V_{\mathbf{k}}\left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\ \cosh\theta_{\mathbf{k}}\ \cos 2\,(\zeta_0-\zeta_{\mathbf{k}})\right]$$
$$+\frac{1}{2}\sum_{\mathbf{k},\mathbf{k}'\neq 0} V_{\mathbf{k}-\mathbf{k}'}\left[\sinh^2\theta_{\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}'} + \sinh\theta_{\mathbf{k}}\ \sinh\theta_{\mathbf{k}'}\ \cosh\theta_{\mathbf{k}}\ \cosh\theta_{\mathbf{k}'}\cos 2\,(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}'})\right] \tag{61}$$

The second summation on the right-hand side describes the effect of momentum transfers between $\mathbf{k}\neq 0$ particles and the condensate, as well as the processes of annihilation and creation of pairs from the condensate (this last process depends on the relative phases $\zeta_0-\zeta_{\mathbf{k}}$ and, as already mentioned, arises from the pairing of particles). The last terms on the right-hand side, included in the double summation over $\mathbf{k}$ and $\mathbf{k}'$, correspond to interactions between particles in the $\mathbf{k}\neq 0$ states. Since the number of individual states is very large, we have ignored the constraint $\mathbf{k}\neq\mathbf{k}'$ of relation (59), which has a negligible effect; furthermore, as we noted in Chapter XVII, it is justified to replace the imaginary exponential by a cosine.

We have shown, in this complement, that the paired states are a useful tool for computing the average energy of an ensemble of interacting particles. In the following complements, we shall use these results successively for fermions and bosons.

Complement C_{XVII}

Fermion pairing, BCS theory

1 Optimization of the energy 1890
1-a Function to be optimized 1891
1-b Cancelling the total variation 1893
1-c Short-range potential, study of the gap 1896
2 Distribution functions, correlations 1899
2-a One-particle distribution 1899
2-b Two-particle distribution, internal pair wave function 1901
2-c Properties of the pair wave function, coherence length 1909
3 Physical discussion 1914
3-a Modification of the Fermi surface and phase locking 1914
3-b Gain in energy 1917
3-c Non-perturbative character of the BCS theory 1918
4 Excited states 1919
4-a Bogolubov-Valatin transformation 1919
4-b Broken pairs and excited pairs 1920
4-c Stationarity of the energies 1920
4-d Excitation energies 1922

We present in this complement the BCS mechanism for the pairing of fermions through attractive interactions. The three letters BCS refer to the names of J. Bardeen, L.N. Cooper and J.R. Schrieffer who proposed in 1957 [9] a theory for a physical phenomenon already observed in 1911 by H. Kamerlingh Onnes in Leiden, but as yet unexplained. This latter scientist observed that, below a certain temperature, the electrical resistivity of certain metals (mercury in his case) abruptly goes to zero as a phase transition occurs toward a so-called "superconducting" state. Along with this transition, many other spectacular effects occur, such as the expulsion of magnetic fields from the material. In this complement, we shall be concerned, with the general pairing mechanism of attractive fermions in the framework of BCS theory. We shall not, however, give any detail about the theory of metals, simply accepting the existence of an attraction between the fermions, without justifying its precise origin. In metals, this effective attraction comes from a coupling between electrons and phonons, and is therefore indirect, introducing an additional complexity to the problem. Furthermore, we shall not present any calculation of electrical resistivity, and hence not show that it can go to zero.

The BCS theory is a mean field theory, of the same type as the Hartree-Fock theory (Complements D_{XV} and E_{XV}). In this latter theory, particles are assumed to independently propagate in the mean field created by all the others; the system is described by an N-particle Fock state. Here, we shall assume that the particles form pairs, and this hypothesis will lead us to use, as a variation trial ket, the ket $|\Psi_{BCS}\rangle$ introduced in Chapter XVII; this complement is a direct application of the results of that chapter. The state

we will choose does indeed mathematically resemble a Fock state of "molecules", each composed of two particles. It should not be concluded, however, that this approximation reduces to a theory where each molecule is considered as an identifiable object moving in the mean field created by all the others. This naive picture is correct in the limit where the molecules are very strongly bound, but we shall see that it is totally inappropriate for loosely bound pairs such as those in the BCS theory. As already underlined in the introduction to Chapter XVII, the use of paired states brings a lot of flexibility to the mean field approach, as it allows modulating the binary correlation function between particles, and then to adapt it to interactions.

We start (§ 1) by minimizing the energy to determine the optimal quantum state in the family considered. In § 2, we discuss some physical properties of the optimized BCS wave function, mainly in terms of one- or two-particle correlation functions, but also in terms of what is called "non-diagonal order" (Complements A_{XVI} and A_{XVII}). Finally, in § 3, we shall study in more detail the physical content of the BCS pairing mechanism allowing the optimization of the energy of a fermion system, and in particular the role of phase locking (spontaneous symmetry breaking). For the sake of simplicity, we assume throughout this complement that the temperature is zero, but the BCS method can also be extended to the study of non-zero temperatures. This will lead to the study of excited states (§ 4), as will be briefly mentioned in § 4-d).

Shortly before the BCS theory was established, Cooper proposed a model including two attractive fermions. He showed that the exclusion of their wave functions from the interior of a Fermi sphere led to a bound state having certain properties similar to those described later by the complete BCS theory. This theory can be considered to be a generalization to N particles of the Cooper model, highlighting the collective effects leading to the properties of the BCS ground state. The Cooper model will be studied in Complement D_{XVII}, and its analogies with the N-particle theory will be underlined. In the present complement, we present the BCS theory, starting directly from the general results of Chapitre XVII; we shall also use the average energy values calculated in § 3 of Complement B_{XVII}.

We obviously cannot give here a detailed account of superconductivity theory and its various resulting effects, which would require an entire book. Limiting a large part of the computations to zero temperature situations already implies that numerous phenomena are outside the scope of this complement. To learn more about the subject, the reader can consult reference [8].

1. Optimization of the energy

Relation (B-11) of Chapter XVII yields the expression of the paired state[1] $|\Psi_{BCS}\rangle$:

$$|\Psi_{BCS}\rangle = \exp\left\{\sum_{\mathbf{k}} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right\}|0\rangle = \prod_{\otimes \mathbf{k}} |\varphi_{\mathbf{k}}\rangle \tag{1}$$

[1] This state is a superposition of components containing different numbers of particles. As already mentioned in Chapter XVII, one could also choose a variational state where the particle number is perfectly determined ([8], § 5-4 and Appendix C of Chap 5), but that would make the calculations a bit more complex.

The ket $|\Psi_{\rm BCS}\rangle$ was then normalized by separately normalizing each ket $|\varphi_{\mathbf{k}}\rangle$, which became the kets $|\overline{\varphi}_{\mathbf{k}}\rangle$:

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{2}$$

where the two functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are related by:

$$v_{\mathbf{k}} = u_{\mathbf{k}}\ g_{\mathbf{k}} \tag{3}$$

and satisfy:

$$|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1 \tag{4}$$

In that chapter, $g_{\mathbf{k}}$ was introduced as the Fourier transform of the wave function $\chi(\mathbf{r})$ of the "diatomic molecule" used to build the paired state; until now, this state was not specified. Here, we shall consider the $g_{\mathbf{k}}$ as variational parameters. Choosing $g_{\mathbf{k}} = 0$ leads to $v_{\mathbf{k}} = 0$ and $|u_{\mathbf{k}}| = 1$: in that case, the two individual states $\mathbf{k},\uparrow$ and $-\mathbf{k},\downarrow$ are neither occupied nor paired. They will be, however, if $g_{\mathbf{k}}$ is not zero. In general, the number of non-zero $g_{\mathbf{k}}$ is, a priori, arbitrary (finite or infinite). We can, for example, limit their number by setting a maximum value k_c for the modulus of $\mathbf{k}$, and consider this maximum value as a supplementary variational parameter defining the trial ket.

We were led, in that same Chapter XVII, to set $u_{\mathbf{k}} = \cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}}$ and $v_{\mathbf{k}} = \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}$, relations that imply that $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ have opposite phases (a situation always possible to obtain by changing the global phase of the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$, which has no physical consequences). In the present complement, it will be more convenient to assume that the phase of $|\overline{\varphi}_{\mathbf{k}}\rangle$ is chosen in order to make $u_{\mathbf{k}}$ real and positive, and we will set:

$$\begin{aligned} u_{\mathbf{k}} &= \cos\theta_{\mathbf{k}} \\ v_{\mathbf{k}} &= \sin\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \end{aligned} \tag{5}$$

Relation (C-19) of Chapter XVII yields the particle number in the state $\left|\overline{\Psi}_{\rm BCS}\right\rangle$:

$$\langle N\rangle = 2\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = 2\sum_{\mathbf{k}} \sin^2\theta_{\mathbf{k}} \tag{6}$$

1-a. Function to be optimized

The average particle number in the state $|\Psi_{\rm BCS}\rangle$ may be changed by varying the $\mathbf{k}$ dependence of the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$: as an example, choosing $|u_{\mathbf{k}}| = 1$ and $v_{\mathbf{k}} = 0$ for any value of $\mathbf{k}$, the average number $\langle N\rangle$ will be zero; on the other hand, if $|u_{\mathbf{k}}|$ is very small and $|v_{\mathbf{k}}|$ equals 1 for a great number of $\mathbf{k}$ values, the average total particle number $\langle N\rangle$ can attain arbitrarily large values. As the energy minimization operation makes sense only for a fixed value of $\langle N\rangle$, we shall determine that value with the Lagrangian multiplier μ (chemical potential; see Appendix VI, § 1-c). We will optimize the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ choices by introducing the variations $\mathrm{d}u_{\mathbf{k}}$ and $\mathrm{d}v_{\mathbf{k}}$ and cancelling the variation of the average value $A = \langle H\rangle - \mu\langle N\rangle$. The volume L^3 of the physical system and its chemical potential μ are supposed to be fixed; we can choose one of two equivalent sets of variables to be determined, either the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$, or the $\theta_{\mathbf{k}}$ and the $\zeta_{\mathbf{k}}$.

Relation (40) of Complement B_{XVII} yields $\langle H \rangle$, whereas $\langle N \rangle$ is given by (6). We then have:

$$\begin{aligned} A &= \langle H \rangle - \mu \langle N \rangle \\ &= 2 \sum_{\mathbf{k}} (e_k - \mu) \ \ |v_{\mathbf{k}}|^2 + \frac{V_{\mathbf{0}}}{4} \langle N \rangle^2 + \sum_{\mathbf{k} \neq \mathbf{k}'} [V_{\mathbf{0}} - V_{\mathbf{k}'-\mathbf{k}}] \, |v_{\mathbf{k}}|^2 \, |v_{\mathbf{k}'}|^2 \\ &\qquad + \sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \ v_{\mathbf{k}}^* \ u_{\mathbf{k}} \ v_{\mathbf{k}'} \ u_{\mathbf{k}'} \end{aligned} \tag{7}$$

with, according to (5):

$$|v_{\mathbf{k}}|^2 = \sin^2 \theta_{\mathbf{k}} \qquad ; \qquad v_{\mathbf{k}}^* \ u_{\mathbf{k}} = \sin \theta_{\mathbf{k}} \ \cos \theta_{\mathbf{k}} \ e^{-2i\zeta_{\mathbf{k}}} \tag{8}$$

In the above expression for A, e_k is the kinetic energy of a free particle with momentum $\hbar\mathbf{k}$:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{9}$$

and the $V_{\mathbf{k}}$ are the Fourier transforms of the interaction potential $W_2(\mathbf{r})$:

$$V_{\mathbf{k}} = \frac{1}{L^3} \int \mathrm{d}^3 r \ e^{-i\mathbf{k}\cdot\mathbf{r}} \ W_2(\mathbf{r}) \tag{10}$$

As this potential is rotationally invariant, the function $V_{\mathbf{k}}$ only depends on the modulus k of $\mathbf{k}$, and it must be real (Appendix I, § 2-e); as the potential is attractive, we can assume all the $V_{\mathbf{k}}$ to be negative.

We saw in Chapter XVII that the first term of (7) corresponds to the kinetic energy, the second to the mean field (diagonal term) for particles with opposite spins, the third one to the similar term for particles with identical spins (in which the direct and exchange terms cancel each other for a short-range potential). Finally, the fourth term, on the second line (which contains $V_{\mathbf{k}-\mathbf{k}'}$) plays a particularly important role in what follows; it comes from the pair annihilation-creation diagram schematized in Figure 4 of Complement B_{XVII}. It is often called the "pairing term".

We use relation (6) to write::

$$\mathrm{d}\left[\langle N \rangle\right]^2 = 2 \langle N \rangle \ \mathrm{d} \langle N \rangle = 4 \langle N \rangle \sum_{\mathbf{k}} \mathrm{d} \, |v_{\mathbf{k}}|^2 \tag{11}$$

We then get:

$$\begin{aligned} \mathrm{d}A &= 2 \sum_{\mathbf{k}} \xi_k \ \mathrm{d} \, |v_{\mathbf{k}}|^2 \\ &\qquad + \sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \left\{ v_{\mathbf{k}'} u_{\mathbf{k}'} \ \mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}^*\right] + v_{\mathbf{k}'}^* u_{\mathbf{k}'} \ \mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}\right] \right\} \end{aligned} \tag{12}$$

where the variable ξ_k is the kinetic energy with respect to the chemical potential, and corrected by the interaction effects [2]:

$$\xi_k = e_k - \mu + \frac{V_{\mathbf{0}}}{2} \langle N \rangle + \sum_{\mathbf{k}'} (V_{\mathbf{0}} - V_{\mathbf{k}'-\mathbf{k}}) \ \ |v_{\mathbf{k}'}|^2 \tag{13}$$

[2] A factor 2 appears in front of the summation over $\mathbf{k}'$ as the variations of $|v_{\mathbf{k}}|^2$ and $|v_{\mathbf{k}'}|^2$ in (7) must be added, but is included in the factor 2 in front of the summation containing ξ_k.

Comment:

For the applications of the BCS theory, the choice of the interaction potential to be used in the equations is not necessarily self-evident.

This is especially true in the superconductivity theory of metals, where the fermions involved are electrons which, isolated, interact via a repulsive Coulomb potential. In a metal, however, the direct repulsive interaction between electrons is mostly screened and they interact indirectly via the crystalline network deformations (phonons, see Complement J_V). This phenomenon leads to a long-range attractive component in their effective interaction, and explains why pairing between electrons is possible. This effective interaction depends on the phonon characteristics, and in particular on the Debye frequency of the solid under study.

This is also true in the theory of an ultra-cold diluted fermionic gas, where we do not use directly the interatomic potential in the equations. This interatomic potential contains, at short-range, a strongly repulsive part (often assimilated to a "hard core") and, at an intermediate distance, a strongly attractive well, permitting the formation of a large number of molecular states. Now when the gas under study is very dilute, the three-body collisions leading to these molecules are very rare, meaning these molecular states play practically no role; only the long distance effects of the potential have a real importance. In other words, the essential role is played by the asymptotic properties of the stationary collision states, as described by the scattering amplitude (Chapter VIII, relation B-9) and the associated phase shifts (Chapitre VIII, § C). The potential used in the BCS computations will therefore be an "effective potential". Furthermore, as the collisions occur at very low energy, this effective potential only depends on the phase shift δ_0 associated with $l = 0$. This phase shift is generally characterized by a "scattering length" a_0 defined as $\delta_0 \simeq -ka_0$ when $k \longrightarrow 0$; the effective potential will be attractive if this scattering length is negative.

As this complement deals mainly with the quantum mechanism for BCS pairing, and not with the determination of a valid potential, we shall not examine this point further and assume a pertinent choice of the interaction potential has been made.

1-b. Cancelling the total variation

It is obvious in (7) that the first three terms on the right-hand side depend only on the moduli of the $v_{\mathbf{k}}$; only the last term (annihilation-creation) depends on the phases $\zeta_{\mathbf{k}}$. Now the function A must be minimal if we vary the $\zeta_{\mathbf{k}}$ without changing the $\theta_{\mathbf{k}}$, i.e. when we vary only this last term:

$$\sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \ \sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}} \ \sin\theta_{\mathbf{k}'} \cos\theta_{\mathbf{k}'} \ e^{2i(\zeta_{\mathbf{k}'}-\zeta_{\mathbf{k}})} \tag{14}$$

We assumed that all the potential matrix elements were negative, whereas relation (C-5) of Chapter XVII shows that the products $\sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}}$ are positive. The lowest value of this sum will be obtained when all the terms in the summation over $\mathbf{k}$ and $\mathbf{k}'$ have the same phase in order to add coherently. This condition is called the "phase locking condition", and will be discussed in more detail in § 3-a-β. The minimum obtained does not depend on the absolute phase of the $v_{\mathbf{k}}$, but only on their relative phases. One can then simply choose all the $\zeta_{\mathbf{k}}$ to be equal to zero, which means that all the $v_{\mathbf{k}}$ are real and positive. This is the choice we shall adopt from now on.

In relation (12), the terms in $\mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}^{*}\right]$ and $\mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}\right]$ now become equal (the $\mathbf{k}$ and $\mathbf{k}'$ are dummy indices), and we have:

$$\mathrm{d}A = 2\sum_{\mathbf{k}} \xi_k \ \mathrm{d}\left(v_{\mathbf{k}}\right)^2 + 2\sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \ v_{\mathbf{k}'} u_{\mathbf{k}'} \ \mathrm{d}\left(u_{\mathbf{k}} v_{\mathbf{k}}\right) \tag{15}$$

We must now vary the $\theta_{\mathbf{k}}$. For that purpose, we introduce the quantities $\Delta_{\mathbf{k}}$, having the dimension of an energy, as:

$$\Delta_{\mathbf{k}} = \sum_{\mathbf{k}'} \left(-V_{\mathbf{k}-\mathbf{k}'}\right) \ u_{\mathbf{k}'} \ v_{\mathbf{k}'} \tag{16}$$

The $\Delta_{\mathbf{k}}$ are real since the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are real, and positive since we assumed the interaction potential matrix elements to be negative. They are called "gaps" and play an important role in the BCS theory. It will be easier to discuss this role in the case of a very short-range potential; this will be done in § 1-c. The choice of the word "gap" will also be explained later (in § 4, see in particular Figure 7). The variation of A can now be written as:

$$\mathrm{d}A = 4\sum_{\mathbf{k}} \xi_k \ v_{\mathbf{k}} \ \mathrm{d}v_{\mathbf{k}} - 2\sum_{\mathbf{k}} \Delta_{\mathbf{k}} \ \left[u_{\mathbf{k}} \mathrm{d}v_{\mathbf{k}} + v_{\mathbf{k}} \mathrm{d}u_{\mathbf{k}}\right] \tag{17}$$

The variations of $\mathrm{d}u_{\mathbf{k}}$ and $\mathrm{d}v_{\mathbf{k}}$ are, however, not independent since relation (4) requires (for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ real) that:

$$2u_{\mathbf{k}} \ \mathrm{d}u_{\mathbf{k}} + 2v_{\mathbf{k}} \ \mathrm{d}v_{\mathbf{k}} = 0 \tag{18}$$

This means we can replace $\mathrm{d}u_{\mathbf{k}}$ by $-v_{\mathbf{k}}\mathrm{d}v_{\mathbf{k}}/u_{\mathbf{k}}$. The right-hand side of (17) then becomes:

$$4\sum_{\mathbf{k}} \xi_k \ v_{\mathbf{k}} \ \mathrm{d}v_{\mathbf{k}} - 2\sum_{\mathbf{k}} \Delta_{\mathbf{k}} \ \left[u_{\mathbf{k}} - \frac{\left(v_{\mathbf{k}}\right)^2}{u_{\mathbf{k}}}\right] \mathrm{d}v_{\mathbf{k}} \tag{19}$$

Cancelling the variation of A with respect to all the $v_{\mathbf{k}}$ leads to:

$$2\xi_k v_{\mathbf{k}} - \Delta_{\mathbf{k}} \ \left[u_{\mathbf{k}} - \frac{\left(v_{\mathbf{k}}\right)^2}{u_{\mathbf{k}}}\right] = 0 \tag{20}$$

Multiplying by $u_{\mathbf{k}}$, we get:

$$2\xi_k \ u_{\mathbf{k}} v_{\mathbf{k}} = \Delta_{\mathbf{k}} \left[\left(u_{\mathbf{k}}\right)^2 - \left(v_{\mathbf{k}}\right)^2\right] \tag{21}$$

or else:

$$\xi_k \ \sin 2\theta_{\mathbf{k}} = \Delta_{\mathbf{k}} \ \cos 2\theta_{\mathbf{k}} \tag{22}$$

One can then compute the sine and the cosine, since:

$$\left[\cos 2\theta_{\mathbf{k}}\right]^2 = \frac{\left[\cos 2\theta_{\mathbf{k}}\right]^2}{\left[\cos 2\theta_{\mathbf{k}}\right]^2 + \left[\sin 2\theta_{\mathbf{k}}\right]^2} = \frac{\left(\xi_k\right)^2}{\left(\xi_k\right)^2 + \left(\Delta_{\mathbf{k}}\right)^2} \tag{23}$$

and we obtain:

$$\cos 2\theta_{\mathbf{k}} = \frac{\pm\xi_k}{\sqrt{(\xi_k)^2 + (\Delta_{\mathbf{k}})^2}} = \pm\frac{\xi_k}{E_k}$$
$$\sin 2\theta_{\mathbf{k}} = \frac{\pm\Delta_{\mathbf{k}}}{\sqrt{(\xi_k)^2 + (\Delta_{\mathbf{k}})^2}} = \pm\frac{\Delta_{\mathbf{k}}}{E_k} \tag{24}$$

where we have set:

$$E_k = \sqrt{(\xi_k)^2 + (\Delta_{\mathbf{k}})^2} \tag{25}$$

We finally obtain:

$$[u_{\mathbf{k}}]^2 = \frac{1}{2}\left[1 + \cos 2\theta_{\mathbf{k}}\right] = \frac{1}{2}\left[1 \pm \frac{\xi_k}{E_k}\right]$$
$$[v_{\mathbf{k}}]^2 = \frac{1}{2}\left[1 - \cos 2\theta_{\mathbf{k}}\right] = \frac{1}{2}\left[1 \mp \frac{\xi_k}{E_k}\right] \tag{26}$$

They are many possibilities for rendering stationary the difference of average values, $A = \langle H\rangle - \mu\langle N\rangle$, depending on the sign $\pm$ chosen in each equation, and for each value of $\mathbf{k}$. This multiplicity of solutions is not surprising since the stationarity is obtained not only for the ground state, but also for all the possible excited states of the physical system; those will be discussed in § 4, and we will see that the stationarity conditions (26) include the possibility of "excited pairs" (§ 4-c-β). For the moment, we shall concentrate on the search for the ground state and look for the absolute minimum of the average value (7).

Let us examine which signs must be chosen in (26) to get the ground state, i.e. the lowest possible value of A in expression (7). The chemical potential of an ideal fermion gas in its ground state is positive, equal to the Fermi level, and proportional to the particle density to the power 2/3 (Complement C_{XIV}, § 1-a). In the presence of a weak attractive potential, the factor $(e_k - \mu)$ in the first term on the right-hand side of (7) is negative when the modulus of $\mathbf{k}$ is small, and positive when $e_k \gg \mu$. In the first case, to minimize A, it is better to choose values of $[v_{\mathbf{k}}]^2$ as large as possible, and hence the $-$ sign in the second equation (26) since $\xi_{\mathbf{k}}$ is negative in this case. On the other hand, when $e_k \gg \mu$, it is better to choose values of $[v_{\mathbf{k}}]^2$ as small as possible, and it is again the $-$ sign in the second equality that must be chosen. As a result, it is the $-$ sign that must be taken in the second equality (26), and hence the $+$ sign in the first one.

As we know that $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are positive, we finally obtain for the ground state:

$$u_{\mathbf{k}} = \sqrt{\frac{1}{2}\left(1 + \frac{\xi_k}{E_k}\right)}$$
$$v_{\mathbf{k}} = \sqrt{\frac{1}{2}\left(1 - \frac{\xi_k}{E_k}\right)} \tag{27}$$

Inserting these results in (21), we verify that the stationarity relations are fulfilled, independently of the sign of ξ_k. They apply as long as the self-consistent condition derived

from the $\Delta_{\mathbf{k}}$ definition (16) is satisfied:

$$\Delta_{\mathbf{k}} = -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sqrt{1-\left(\frac{\xi_{k'}}{E_{k'}}\right)^2} = -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\frac{\Delta_{\mathbf{k}'}}{\sqrt{(\xi_{k'})^2+(\Delta_{\mathbf{k}'})^2}} \tag{28}$$

As we now show, this condition takes on a simpler form for a very short-range interaction potential.

The above computation shows that starting from any function $\chi(\mathbf{r})$, or (which amounts to the same thing) from functions $g_{\mathbf{k}}$ considered to be entirely free variables, the optimization procedure yields values for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$; this in turn fixes the optimal $g_{\mathbf{k}}$ and determines the function $\chi(\mathbf{r})$ for building the paired state described in (1).

1-c. Short-range potential, study of the gap

The matrix elements of the interaction potential V_k were defined in Complement B_{XVII} as the Fourier transforms of the potential – see relations (16) and (18) of that complement. For a regular potential of range b, the matrix element V_k necessarily varies when k changes by a quantity of the order of $1/b$; in particular, $V_k \to 0$ when $k \gg 1/b$. However, in many physical applications of the BCS theory, the k wave vectors involved remain very small compared to $1/b$, and a useful approximation is to ignore the variations of the V_k. We therefore consider them all equal to the same constant $-V$:

$$V_{\mathbf{k}} = -V \tag{29}$$

The minus sign was introduced to make V a positive number for an attractive potential; this number is inversely proportional to the volume $\mathcal{V}$, as shown by relation (10). Definition (13) of $\xi_{\mathbf{k}}$ now takes on a simplified form:

$$\xi_k = e_k - \mu - \frac{V}{2}\langle N\rangle \tag{30}$$

that can be inserted in the relations (27) to get the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$.

Relation (16) also has a simpler version; all the $\Delta_{\mathbf{k}}$ take on the same value Δ:

$$\Delta = V\sum_{\mathbf{k}'} u_{\mathbf{k}'}v_{\mathbf{k}'} \tag{31}$$

In this case, there exists only one value of the gap, and since the $u_{\mathbf{k}'}$ and $v_{\mathbf{k}'}$ are real, this value is also real. We shall see in what follows that Δ plays a particularly important role, especially in the dispersion curve characterizing the system excitations (see for example Figure 7 which shows the existence of an energy minimum equal to Δ).

All the previous formulas apply, provided we replace the $\Delta_{\mathbf{k}}$ by Δ. Relations (24) then become, with the sign choice leading to the ground state:

$$\begin{aligned}\cos 2\theta_{\mathbf{k}} &= \frac{\xi_k}{\sqrt{(\xi_k)^2+\Delta^2}} = \frac{\xi_k}{E_k}\\ \sin 2\theta_{\mathbf{k}} &= \frac{\Delta}{\sqrt{(\xi_k)^2+\Delta^2}} = \frac{\Delta}{E_k}\end{aligned} \tag{32}$$

Equalities (27) are unchanged. When $k \to \infty$ the second relation (32) shows that, taking (25) and (30) into account, the value of $\theta_{\mathbf{k}}$ goes to zero as:

$$\theta_{\mathbf{k}} \sim \Delta/2E_k \sim \Delta/2\xi_k \tag{33}$$

It will be useful in what follows to know the asymptotic behaviors of the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, whose values as a function of $\theta_{\mathbf{k}}$ are given by (5). When $k \to \infty$, relation (33) shows that $u_{\mathbf{k}}$ goes to 1 whereas $v_{\mathbf{k}}$ goes to zero as:

$$v_{\mathbf{k}} \underset{k\to\infty}{\sim} \frac{\Delta}{2\xi_k} + 0(\frac{\Delta^2}{\xi_k^2}) \sim \frac{1}{k^2} + 0(\frac{1}{k^4}) \tag{34}$$

This ensures the convergence of the summations (C-19) and (C-26) of Chapter XVII giving the average values of $\widehat{N}$ and of its square.

α. Self-consistency condition and divergences

The self-consistent condition (28) now takes on the simpler form:

$$\Delta = \frac{V}{2}\sum_{\mathbf{k}} \sqrt{1 - \left(\frac{\xi_k}{E_k}\right)^2} = \frac{V}{2}\sum_{\mathbf{k}} \frac{\Delta}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{35}$$

that is:

$$1 = \frac{V}{2}\sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{36}$$

which is an implicit equation expressing the gap Δ in terms of μ (as this latter parameter appears in the definition of ξ_k).

Choosing a large volume L^3, we can replace the discrete summation over $\mathbf{k}$ by an integral. We shall assume, as mentioned in § 1, that the $g_{\mathbf{k}}$ in the variational ket (1) are zero when the modulus of $\mathbf{k}$ is larger than a given cutoff value k_c. Under these conditions, the implicit equation for the gap becomes:

$$1 = \frac{V}{2}\frac{L^3}{(2\pi)^3}\int_0^{k_c} \mathrm{d}^3k \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{37}$$

where k_c is the upper limit of the wave vectors introduced in § 1; remember that V is inversely proportional to the volume, and hence the right-hand side of this equation does not depend on that volume. The integral will diverge if k_c is infinite, since when $k \to \infty$, the function to be integrated behaves as $1/\xi_k \sim 1/e_k \sim 1/k^2$. The value obtained for Δ therefore depends on the value chosen for k_c; this upper limit then plays an important role.

β. Calculation of the gap

We note e_c the equivalent, in terms of energy, of the cutoff frequency k_c:

$$e_c = \frac{\hbar^2 k_c^2}{2m} \tag{38}$$

We now choose the energy e_k as the integration variable. We then have to consider the density of states $D(e_k)$, obtained[3] by taking the differential of the definition of e_k:

$$D(e_k) = \frac{L^3}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{e_k} \tag{39}$$

Relation (37) then becomes:

$$1 = \frac{V}{2}\int_0^{e_c} D(e_k)\ \mathrm{d}e_k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}} \tag{40}$$

where, to simplify relation (30), we introduced a chemical potential $\overline{\mu}$ relative to the mean field energy[4]:

$$\overline{\mu} = \mu + \frac{V}{2}\langle N\rangle \tag{41}$$

In relation (40), the function to be integrated over e_k contains a fraction that is maximum for $e_k = \overline{\mu}$; this fraction takes on significant values in an energy band of width Δ centered, in $\mathbf{k}$ space, on the surface of the "Fermi sphere" (see Complement C_{XIV}) whose radius k_F obeys:

$$\frac{\hbar^2 k_F^2}{2m} = \mu \tag{42}$$

As for the density of states $D(e_k)$, it takes on low values in the vicinity of the center of that sphere, but increasingly larger ones outside. The inside of the sphere barely contributes to the summation, the main contribution coming from the outside, in between the Fermi surface and the cutoff energy e_{k_c}. We can then find in this region an intermediate value D_0 for the density of states that can replace $D(e_k)$ without changing the integral, with:

$$D(\mu) \leq D_0 \leq D(e_c) \tag{43}$$

The density of states can be removed from the integral, and we get:

$$1 = \frac{VD_0}{2}\int_0^{e_c} \mathrm{d}e_k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}} \tag{44}$$

As V is inversely proportional to the volume, whereas, according to (39), the density of states is proportional to it, this relation is independent of the volume.

In the physics of superconducting metals, the attractive interaction between the electrons is mediated by the motions of the crystal's ions, i.e. by the phonons of the network. A cutoff energy naturally appears in the matrix elements of the interaction potential, the Debye energy $\hbar\omega_D$ of the phonons. One often uses a simple model where,

[3] This density of states is defined in Complement C_{XIV}, and given by formula (8) of that complement. In our case, as we do not have to take into account the two spin states, the density of states is half the one computed in that complement.

[4] With the sign convention we chose for V in (29), this mean field energy is equal to $-V\langle N\rangle/2$ per particle.

in (28), the potential matrix elements $V_{\mathbf{k}-\mathbf{k}'}$ are zero as soon as the difference in energy $|e_k - e_{k'}|$ is larger than $\hbar\omega_D$, supposed to be much smaller than $\overline{\mu}$; otherwise, they are all equal to a constant V. The same computations as those that led to (44), yield in this case the gap equation[5] for levels close to the Fermi surface:

$$1 \simeq \frac{VD_F}{2}\int_{\overline{\mu}-\hbar\omega_D}^{\overline{\mu}+\hbar\omega_D} \mathrm{d}e_k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}} = VD_F \ \mathrm{arsinh}\left(\frac{\hbar\omega_D}{\Delta}\right) \tag{45}$$

where D_F is the density of states on the Fermi surface. If, furthermore, we make the approximation $VD_F \ll 1$, we get:

$$\Delta = \frac{\hbar\omega_D}{\sinh\left(1/VD_F\right)} \simeq 2\hbar\omega_D \exp\left(-\frac{1}{VD_F}\right) \tag{46}$$

This important equation is called the "BCS gap equation".

It is worth noting that this expression cannot be expanded in a power series of V when the interaction goes to zero: all the derivatives of Δ with respect to V are zero for $V = 0$. Consequently, this expression cannot be obtained in the framework of a perturbation theory in powers of the interaction (this point will be discussed in § 3-c).

2. Distribution functions, correlations

Inserting expressions (27) in the trial ket, we obtain the optimal state vector $|\Psi_{\mathrm{BCS}}\rangle$ that best describes the ground state. We now examine the physical implications of that optimized quantum state, concerning the properties of the one- or two-particle distribution functions. These properties will be used later on in this complement to understand the origin of the energy lowering due to condensation into pairs of particles.

2-a. One-particle distribution

As we are going to show, the properties of the one-particle distribution are fairly close to those of an ideal gas.

α. Momentum space

Once the gap Δ is obtained, we can use relations (30) and (32) to determine the values of $\theta_{\mathbf{k}}$ for each value of $\mathbf{k}$; relation (C-16) of Chapter XVII then yields the average number of particles in each pair of states. As the two states composing the pair play the same role, the average number of particles in each of the states is simply half that number, that is $\sin^2\theta_{\mathbf{k}}$. Figure 1 plots, as a function of e_k, the variation of the distribution function $\langle n_{\mathbf{k},\nu}\rangle$ obtained, which is the momentum distribution function of a particle, once the variables $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ have been optimized. For an ideal gas, we saw in Complement $\mathrm{B_{XV}}$ that it is a Fermi-Dirac distribution; at zero temperature (as is the case here, since we are studying the ground state), this distribution is a "step function" equal to 1 for $e_k < \mu$ and to zero for $e_k > \mu$ (dotted curve). In our case, the transition

[5] There are two equal contributions to the integral on the right-hand side, one from the values of e_k above $\overline{\mu}$, the others from those below; this is why the factor 1/2 has disappeared from the second equality.

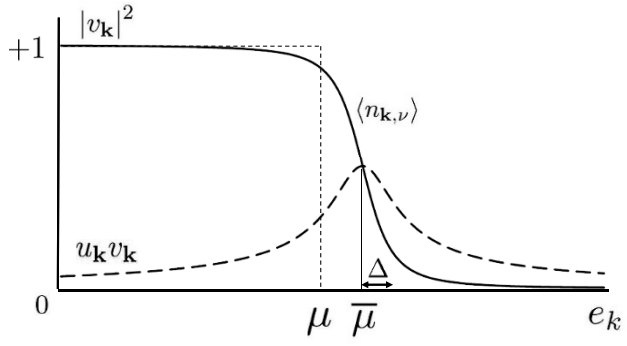


Figure 1: Plots of the one-particle distribution function $\langle n_{\mathbf{k},\nu} \rangle = |v_{\mathbf{k}}|^2$ *in the BCS state, as a function of the energy* e_k*. In the absence of interactions, this function is equal to* 1 *for* $e_k < \mu$*, and zero for* $e_k > \mu$ *(dotted line step function). In the presence of interactions, due to the pairing of fermions the curve is rounded off over an energy domain of the order of the gap* Δ *(double arrow on the horizontal axis), and the variations occur around the value* $\overline{\mu}$ *(value of* μ *shifted by the mean field effect). The dashed curve plots the product* $u_{\mathbf{k}} v_{\mathbf{k}}$ *as a function of the same variable* e_k*. This function is largest around* $e_k = \overline{\mu}$*, in a domain spreading over a few* Δ*.*

between 0 and 1 occurs around $\overline{\mu}$, hence for a value of chemical potential shifted by the mean field effect as indicated by relation (41); this energy shift due to the mean field is natural. What is more striking is that the curve no longer presents a discontinuous step, but varies progressively over an energy domain whose width is of the order of the gap Δ. The interaction effect depopulates certain pairs of states in favor of other pairs having higher kinetic energies. Certain fermions are promoted from the inside of the Fermi surface towards the outside, this effect occurring over a depth of the order of Δ. In $\mathbf{k}$ space, the perturbation introduced by the attractions is localized in the neighborhood of that surface; the fermions situated close to the center of the Fermi sphere are not concerned, whereas those close to that surface gain an energy of the order of the gap.

β. Position space

Relations (B-22) and (B-23) of Chapter XVI yield the one-particle correlation function in position space:

$$G_1(\mathbf{r}, \nu; \mathbf{r}', \nu') = \left\langle \Psi_\nu^\dagger(\mathbf{r}) \Psi_{\nu'}(\mathbf{r}') \right\rangle = \langle \mathbf{r}', \nu' | \, \rho_I \, | \mathbf{r}, \nu \rangle \tag{47}$$

where ρ_I is the one-particle density operator. Taking into account formula (A-14) of Chapter XVI, applied to normalized plane waves, we can write:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \frac{1}{L^3}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'-\mathbf{k}\cdot\mathbf{r}\right)} \left\langle a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k}',\nu'} \right\rangle$$
$$= \frac{1}{L^3}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'-\mathbf{k}\cdot\mathbf{r}\right)} \ \langle\Psi_{\mathrm{BCS}}|\, a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k}',\nu'}\, |\Psi_{\mathrm{BCS}}\rangle \tag{48}$$

where $a^\dagger_{\mathbf{k},\nu}$ is the creation operator in the individual state of momentum $\hbar\mathbf{k}$ and spin ν, and $a_{\mathbf{k}',\nu'}$ the annihilation operator in the state of momentum $\hbar\mathbf{k}'$ and spin ν'. Now, in the state $|\Psi_{\mathrm{BCS}}\rangle$, the occupation number of each momentum pair are either 0 or 2, which means that the average values of the product of these operators is zero whenever each of them concerns a different pair, or if the two individual states are in the same pair, but are different from each other. We now have:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \frac{1}{L^3}\delta_{\nu\nu'}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)} \ \langle\Psi_{\mathrm{BCS}}|\, a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k},\nu}\, |\Psi_{\mathrm{BCS}}\rangle$$
$$= \frac{1}{L^3}\delta_{\nu\nu'}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)} \langle\hat{n}_{\mathbf{k},\nu}\rangle \tag{49}$$

with:

$$\langle n_{\mathbf{k},\nu}\rangle = \langle\Psi_{\mathrm{BCS}}|\, a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k},\nu}\, |\Psi_{\mathrm{BCS}}\rangle = |v_{\mathbf{k}}|^2 \tag{50}$$

The function G_1 is proportional to the Fourier transform of the average population $\langle n_{\mathbf{k},\nu}\rangle$ of the individual state $\mathbf{k},\nu$. As this average population is a function whose width is of the order of the Fermi wave vector k_F, the function G_1 goes to zero when $|\mathbf{r}-\mathbf{r}'| \gg 1/k_F$, i.e. as soon as the difference in positions is no longer microscopic: in this system, there exists no "long-range non-diagonal order" of the one-particle correlation function. For $\mathbf{r}=\mathbf{r}'$, we get:

$$G_1(\mathbf{r},\nu;\mathbf{r},\nu') = \frac{\delta_{\nu\nu'}}{2L^3}\left\langle\hat{N}\right\rangle = \delta_{\nu\nu'}\frac{n_{\mathrm{BCS}}}{2} \tag{51}$$

where n_{BCS} is the numerical particle density:

$$n_{\mathrm{BCS}} = \frac{\left\langle\hat{N}\right\rangle}{L^3} \tag{52}$$

The function $G_1(\mathbf{r},\nu;\mathbf{r},\nu)$ has no spatial dependence; the factor $1/2$ reflects the fact that the total density n_{BCS} is shared equally among the two spin states. The results are the same as for an ideal gas.

2-b. Two-particle distribution, internal pair wave function

As opposed to the one-particle distribution, the two-particle distribution is strongly affected by the BCS mechanism, which is to be expected since it is a pairing process.

α. Momentum space, peak in the distribution

Relation (C-19) of Chapter XV yields the expression for the matrix elements of the two-particle density operator ρ_{II} in an unspecified basis. In the momentum representation, they are written:

$$\langle 1:\mathbf{k}'',\nu_3;2:\mathbf{k}''',\nu_4|\,\rho_{II}\,|1:\mathbf{k},\nu_1;2:\mathbf{k}',\nu_2\rangle = \left\langle a^{\dagger}_{\mathbf{k},\nu_1}a^{\dagger}_{\mathbf{k}',\nu_2}a_{\mathbf{k}''',\nu_4}a_{\mathbf{k}'',\nu_3}\right\rangle \tag{53}$$

We shall mainly consider the diagonal elements, characterizing the correlations between the momenta of two particles:

$$\langle 1:\mathbf{k},\nu;2:\mathbf{k}',\nu'|\,\rho_{II}\,|1:\mathbf{k},\nu;2:\mathbf{k}',\nu'\rangle = \left\langle a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',\nu'}a_{\mathbf{k}',\nu'}a_{\mathbf{k},\nu}\right\rangle \tag{54}$$

In this expression, the creation operators repopulate precisely the same states as those that have been depopulated by the annihilation operators.

For $\nu'=\nu$, in order to obtain a non-zero result, we must have $\mathbf{k}'\neq\mathbf{k}$ (otherwise we get the square of a fermionic operator, which is zero). We get:

$$\langle 1:\mathbf{k},\nu;2:\mathbf{k}',\nu|\,\rho_{II}\,|1:\mathbf{k},\nu;2:\mathbf{k}',\nu\rangle = |v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 \qquad (\text{if } \mathbf{k}'\neq\mathbf{k}) \tag{55}$$

which means there exists no correlations between the momenta.

For $\nu'=-\nu$, when $\mathbf{k}'$ is different from $-\mathbf{k}$, two different pairs are involved, and we again get a product[6]:

$$\langle 1:\mathbf{k},\nu;2:\mathbf{k}',-\nu|\,\rho_{II}\,|1:\mathbf{k},\nu;2:\mathbf{k}',-\nu\rangle = |v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 \qquad (\text{if } \mathbf{k}'\neq-\mathbf{k}) \tag{56}$$

On the other hand, if $\mathbf{k}'=-\mathbf{k}$, only one pair is concerned, destroyed and then reconstructed by the operators (as before, this is a contribution from the diagram in Figure 4 of Complement B_{XVII}; the computation then involves a single state $|\overline{\varphi}_{\mathbf{k}}\rangle$, and we get:

$$\langle 1:\mathbf{k},\nu;2:-\mathbf{k},-\nu|\,\rho_{II}\,|1:\mathbf{k},\nu;2:-\mathbf{k},-\nu\rangle = |v_{\mathbf{k}}|^2 \tag{57}$$

This result is not the limit of the previous one when $\mathbf{k}'\to-\mathbf{k}$, which would be $|v_{\mathbf{k}}|^4$; the value we obtain is larger, since $|v_{\mathbf{k}}|^2\leq 1$.

This shows that for all the k values that do not correspond to a pair of opposite momenta, the density operator is simply a product, involving no correlations between the particles' momenta. This confirms what we found in the study of the one-particle density operator: all the $\mathbf{k}$ states having a momentum smaller than that of the Fermi level are populated, with a rounding off of the functions due to the pairing phenomenon. On the other hand, for opposite values of both the momenta and the spin values, as is the case for a pair, we observe a discontinuity of the diagonal correlation function: it jumps from $|v_{\mathbf{k}}|^4$ to the larger value $|v_{\mathbf{k}}|^2$. The corresponding discontinuity $\delta n^{(2)}_{\mathbf{k}}$ can be written as:

$$\begin{aligned}\delta n^{(2)}_{\mathbf{k}} &= |v_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^4 = |v_{\mathbf{k}}|^2\left[1-|v_{\mathbf{k}}|^2\right]\\ &= |u_{\mathbf{k}}v_{\mathbf{k}}|^2\end{aligned} \tag{58}$$

[6] If $\nu=-$, to get this result we use the fact that the function $v(\mathbf{k})$ is even. Now if $\mathbf{k}'=\mathbf{k}$, two pairs are still involved, labeled by opposite values of momentum (remember that we chose the convention where each pair is labeled by the momentum of the spin + particle); but, here again, the parity of $\nu(\mathbf{k})$ leads to $|v(\mathbf{k})|^4$, in agreement with (56).

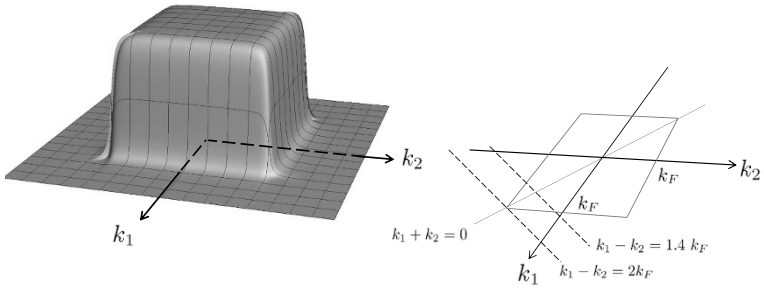


Figure 2: The left part of the figure shows the distribution function of the momenta of two particles, assuming that the two momenta $\hbar\mathbf{k}_1$ et $\hbar\mathbf{k}_2$ are parallel (or antiparallel); a value $\Delta/E_F = 1/10$ has been chosen. In the descending part of the surface, and in the two corners where $\mathbf{k}_1 + \mathbf{k}_2$ vanishes, one can distinguish a small crest indicating a partial Bose-Einstein condensation. In order to see this effect more clearly, we cut the surface along vertical planes whose trace is indicated by the dashed lines in the right part of the figure. This leads to the curves of Figure 3.

We shall see below (§ 2-b-β) that $\delta n_{\mathbf{k}}^{(2)}$ is none other than the square of the $\mathbf{k}$ component of the pair wave function. This discontinuity is significant for the values of $\mathbf{k}$ for which the product $u_{\mathbf{k}} v_{\mathbf{k}}$ takes on its largest values; Figure 1 shows that it corresponds to a region around the Fermi surface, with an energy bandwidth of the order of the gap Δ.

The momentum distribution function depends on the 6 components of the two momenta, which does not allow a simple graphic representation. To simplify, we are going to assume the two particles' momenta $\hbar\mathbf{k}_1$ and $\hbar\mathbf{k}_2$ are parallel (or antiparallel), so that the distribution we wish to represent becomes a surface in three-dimensional space: we plot k_1 along one axis, k_2 along the second, and the probability along the third perpendicular axis. We then obtain the surface shown in the left part of Figure 2, where it has been assumed that $\Delta/E_F = 1/10$.

To explore this surfact in more detail, we plot in Figure 3 the curves we obtain by cutting this surface by vertical planes parallel to the first bisector of the k_1 and k_2 axes; we assume the difference $k_1 - k_2$ to be constant (dashed lines in the right part of Figure 2) and use, as the variable, the sum of these two momenta. The horizontal axis in Figure 3 then represents the dimensionless variable q:

$$q = \frac{k_1 + k_2}{2k_F} \tag{59}$$

As q varies, the corresponding point in the plane k_1, k_2 moves along a straight line.

For a fixed value of $\kappa = k_1 - k_2$, we must set $q = 0$ for the wave vector components to have opposite values: $k_1 = \kappa/2$ and $k_2 = -\kappa/2$. On the left-hand side of Figure 3, the difference κ is chosen equal to 1.4 k_F; we obtain an almost square curve, rounded off by the fact that Δ is not zero (as was the case in Figure 1), and which presents

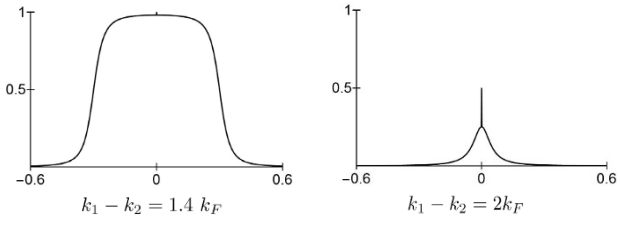


Figure 3: Plots of the distribution function for two particles of parallel (or antiparallel) momenta $\mathbf{k}_1$ *and* $\mathbf{k}_2$*, as a function of the dimensionless variable* $q = (k_1 + k_2)/2k_F$*; the figure was plotted with the choice* $\Delta/E_F = 1/10$*.*
For the curve on the left-hand side, the difference $(k_1 - k_2)$ *is chosen equal to* $1.4\ k_F$*. The curve looks like a bell shaped function, practically constant for small values of q, and decreasing for larger values following a rounded slope similar to the one in Figure 1 (and all the more steep as the* Δ *value is chosen smaller). No singularity of the distribution is clearly visible (except for a minuscule peak at the origin).*
For the curve on the right-hand side, the difference in momenta is chosen equal to $2\hbar k_F$*; when q is close to zero, the two momenta take on values that both fall into the rounded part of the one-particle distribution. A singularity at* $q = 0$ *is now clearly visible, signaling an accumulation of "molecules" in a state where their center of mass does not move. The height of the central peak corresponds to the population of the discrete level having a zero total momentum, and its width reduces to zero as it is a discrete level.*

a barely visible peak. But, as we saw before, the effects of the pairing are important when $k_1 = -k_2 \simeq \pm k_F$, i.e. when $\kappa \simeq \pm 2k_F$. On the right-hand side of Figure 3, the momentum difference is chosen equal to $2\hbar k_F$, so that the momenta can both fall in the rounded part of the distributions. We observe, superposed on a "pedestal", a narrow peak indicating an additional population in the level having a zero total momentum. The value of that population is given by the height of the peak, whose width is strictly zero for discrete levels. The singularity of this momentum distribution is then clearly visible.

Therefore, a singularity appears in the momentum distribution of particle pairs, whose centers of mass present a condensation in momentum space. This is, however, a partial condensation: as opposed to the boson case, the condensation peak appears on a pedestal due the presence of a majority of non-condensed pairs. Actually, the only pairs involved are those whose two components have energies falling around the Fermi level E_F, in a bandwidth of the order of the gap Δ. Despite these restrictions, it is nevertheless true that the condensation phenomenon into attractive BCS pairs has properties related to Bose-Einstein condensation for repulsive bosons. The link between that condensation and the appearance of an order parameter for the pair field is discussed in § 2-b-γ of Complement A_{XVII}.

β. Position space, correlations described by the pair wave functions

We did not find any effects of the interactions on G_1. But, as pointed out before, since the BCS theory relies on pairing, one expects to find more interesting properties concerning the two-particle correlation functions. They will be studied now, limiting ourselves to the "diagonal" correlation function, as defined by relation (B-33) of Chapter XVI, including the spin variables as in (B-36). This function is written as:

$$\begin{aligned} G_2(\mathbf{r},\nu;\mathbf{r}',\nu') &= \left\langle \Psi_\nu^\dagger(\mathbf{r})\Psi_{\nu'}^\dagger(\mathbf{r}')\Psi_{\nu'}(\mathbf{r}')\Psi_\nu(\mathbf{r})\right\rangle \\ &= \langle 1:\mathbf{r},\nu;2:\mathbf{r}',\nu'|\ \rho_{II}\,|1:\mathbf{r},\nu;2:\mathbf{r}',\nu'\rangle \end{aligned} \tag{60}$$

(ρ_{II} is the two-particle density operator), or else, as before:

$$G_2(\mathbf{r},\nu;\mathbf{r}',\nu') = \frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''} e^{i\left[(\mathbf{k}-\mathbf{k}''')\cdot\mathbf{r}+(\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{r}'\right]}\left\langle a_{\mathbf{k}''',\nu}^\dagger a_{\mathbf{k}'',\nu'}^\dagger a_{\mathbf{k}',\nu'}a_{\mathbf{k},\nu}\right\rangle \tag{61}$$

This expression includes the average values of the product of four operators, whose computation is similar to the one explained in § 3 of Complement $\mathrm{B_{XVII}}$ for the average interaction energy, except that, in our case, the spin indices are fixed rather than appearing as summation indices. Figure 4 schematizes with a diagram each term of (61): the incoming arrows represent particles that disappear (annihilation operator action), the outgoing arrows represent those that will appear (creation operator action); each value of $\mathbf{k}$ is associated with a position value $\mathbf{r}$, via an exponential $e^{i\mathbf{k}\cdot\mathbf{r}}$ for the incoming arrows, or $e^{-i\mathbf{k}\cdot\mathbf{r}}$ for the outgoing ones, as well as with a value of the spin ν.

Parallel spins: if $\nu=\nu'$, the two destruction operators necessarily concern pairs with different $\mathbf{k}$. To restore the populations of these two couples of states to an even value, the only possibility is to again give each one its initial value; otherwise the result will be zero. We must have, either $\mathbf{k}=\mathbf{k}'''$ and $\mathbf{k}'=\mathbf{k}''$ (direct term), or $\mathbf{k}=\mathbf{k}''$ and $\mathbf{k}'=\mathbf{k}'''$ (exchange term). In the first case, we obtain (after two anticommutations whose sign changes cancel each other) a result[7] independent of the position variables:

$$\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}\langle\overline{\varphi}_\mathbf{k}|\,a_{\mathbf{k},\nu}^\dagger a_{\mathbf{k},\nu}\,|\overline{\varphi}_\mathbf{k}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a_{\mathbf{k}',\nu}^\dagger a_{\mathbf{k}',\nu}\,|\overline{\varphi}_{\mathbf{k}'}\rangle = \frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}|v_\mathbf{k}|^2\,|v_{\mathbf{k}'}|^2 \tag{62}$$

and in the second case (after only one anticommutation):

$$\begin{aligned} -\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}e^{i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{r}-\mathbf{r}')}\,&\langle\overline{\varphi}_\mathbf{k}|\,a_{\mathbf{k},\nu}^\dagger a_{\mathbf{k},\nu}\,|\overline{\varphi}_\mathbf{k}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a_{\mathbf{k}',\nu}^\dagger a_{\mathbf{k}',\nu}\,|\overline{\varphi}_{\mathbf{k}'}\rangle \\ &= -\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}e^{i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{r}-\mathbf{r}')}\,|v_\mathbf{k}|^2\,|v_{\mathbf{k}'}|^2 \end{aligned} \tag{63}$$

Regrouping these two contributions, and using (6), we get:

$$G_2(\mathbf{r},\nu;\mathbf{r}',\nu) = \left(\frac{\langle N\rangle}{2L^3}\right)^2\left[1-\left[F(\mathbf{r}-\mathbf{r}')\right]^2\right] \tag{64}$$

[7] If $\nu=-$, we must change the sign of $\mathbf{k}$ and $\mathbf{k}'$ in $v_\mathbf{k}$ and $v_{\mathbf{k}'}$ but, as before, it does not change the result since we can change the sign of summation variables.

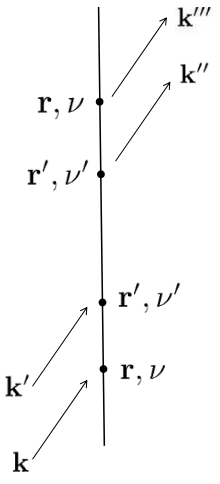


Figure 4: This diagram symbolizes each term involved in the binary correlation function. The two incoming arrows on the bottom represent particles that will disappear during the interaction under the action of the two annihilation operators; the two outgoing arrows on the top represent particles that appeared during the interaction under the action of the two creation operators. All the arrows are associated with an imaginary exponential of the position, with a positive argument for the incoming arrows, and a negative one for the outgoing arrows. The indices ν label the spins.

with an exchange term containing the (real) function:

$$F(\mathbf{r}) = \frac{2}{\langle N \rangle} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \left| v_{\mathbf{k}} \right|^2 \tag{65}$$

This result has the same form as relation (22) of Complement A_{XVI}, taking into account the fact that the population of each spin state is half of $\langle N \rangle$. It shows that the correlation function for two parallel spins exhibits an "exchange hole" very similar to the one plotted in Figure 2 of that complement, but with a slightly different shape, since here the functions $\left| v_{\mathbf{k}} \right|^2$ are no longer exactly discontinuous step functions. The width of this exchange hole is of the order of the inverse of k_F, the Fermi wave number related to the Fermi energy E_F by $E_F = \hbar^2 k_F^2 / 2m$.

Opposite spins: If $\nu \neq \nu'$, it is possible for the two annihilation or creation operators to act on the same pair of states; we are then dealing with a pair annihilation-creation term (term of type **II** according to the classification presented in § D-1-a of Chapter XVII). Figure 5 symbolizes the three types of diagrams playing a role in the computation of the correlation function for opposite spins: **I** (forward scattering), **II** (pair-pair) and **III** (special cases). The computation of their sum has been performed in § D-2 of that same chapter, and led to the following result:

$$G_2(\mathbf{r}, \nu; \mathbf{r}', -\nu) \simeq \left[\frac{\langle N \rangle}{2L^3} \right]^2 + \left| \phi_{\text{pair}}(\mathbf{r} - \mathbf{r}') \right|^2 \tag{66}$$

We have used the following definition of the (non-normalized) "pair wave function"[8] $\phi_{\text{pair}}(\mathbf{r}-\mathbf{r}')$:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}} = \frac{\Delta}{2L^3}\sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2+\Delta^2}} e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{67}$$

which is simply the pair wave function already introduced in relation (D-14) of Chapter XVII. We find again relations (38) and (39) of Complement A_{XVII}, where this wave function was obtained by a different method involving the two-particle field operator. The presence of the second term on the right-hand side of (66) is thus due to the non-zero average value of the pair field, introduced in that complement (non-zero order parameter).

We have just shown that, contrary to what happens in an ideal gas (Complement A_{XVI}, § 2-b), two particles with opposite spins may be spatially correlated. This correlation is described by the modulus squared of the wave function $\phi_{\text{pair}}(\mathbf{r}-\mathbf{r}')$, defined by its spatial Fourier transform $u_{\mathbf{k}} v_{\mathbf{k}}$. This new wave function, different from the one we used at the beginning to build the N-particle trial wave function, was introduced in § D-2 of Chapter XVII, as well as in Complement A_{XVII}, starting from the pair field operator. The spatial correlation it characterizes is purely dynamic, as it does not exist in the absence of interactions. Its physical consequences, in terms of potential energy, will be discussed in § 3-a-α.

Physical discussion: on the right-hand side of (66), the first term does not contain any spatial dependence; it simply corresponds to the correlation function of an ensemble of totally independent particles. The second term, on the other hand, depends on the position differences $\mathbf{r}-\mathbf{r}'$; we now discuss its physical origin in terms of quantum interference.

This second term comes from the contribution of the pair annihilation-creation terms for which we have, in relation (61), $k=-k'$ and $k''=-k'''$. Let us show that "cutting them in half", they look like interference terms. They include average values of operator products that, when $\nu'=-\nu$, can be written:

$$\begin{aligned}\left\langle a^{\dagger}_{\mathbf{k}''',\nu} a^{\dagger}_{-\mathbf{k}''',-\nu} a_{-\mathbf{k},-\nu} a_{\mathbf{k},\nu}\right\rangle &= \langle\Psi_{\text{BCS}}|\, a^{\dagger}_{\mathbf{k}''',\nu} a^{\dagger}_{-\mathbf{k}''',-\nu} a_{-\mathbf{k},-\nu} a_{\mathbf{k},\nu}\,|\Psi_{\text{BCS}}\rangle \\ &= \left\langle \widetilde{\Psi}(\mathbf{k}''')\,\middle|\,\widetilde{\Psi}(\mathbf{k})\right\rangle \end{aligned} \tag{68}$$

where $\left|\widetilde{\Psi}(\mathbf{k})\right\rangle$ is defined as:

$$\left|\widetilde{\Psi}(\mathbf{k})\right\rangle = a_{-\mathbf{k},-\nu} a_{\mathbf{k},\nu}\,|\Psi_{\text{BCS}}\rangle \tag{69}$$

Relation (66) then becomes:

$$G_2(\mathbf{r},\nu;\mathbf{r}',-\nu) \simeq \left[\frac{\langle N\rangle}{2L^3}\right]^2 + \frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'''} e^{i(\mathbf{k}-\mathbf{k}''')\cdot(\mathbf{r}-\mathbf{r}')} \left\langle \widetilde{\Psi}(\mathbf{k}''')\,\middle|\,\widetilde{\Psi}(\mathbf{k})\right\rangle \tag{70}$$

[8] The factor $1/L^3$ appearing in (67) permits defining a pair wave function independent of the dimension L of the physical system, in the limit of large L where the sum over the $\mathbf{k}$ becomes an integral over d^3k multiplied by $(L/2\pi)^3$. As a result, the square of that wave function is not homogeneous to the inverse of a volume, as is generally the case for a particle wave function, but to $1/L^6$. Actually, it should be considered as a two-particle wave function, product of a constant wave function $1/L^{3/2}$ of the center of mass of the pair (assumed to have a zero momentum) and a wave function describing its relative position variable.

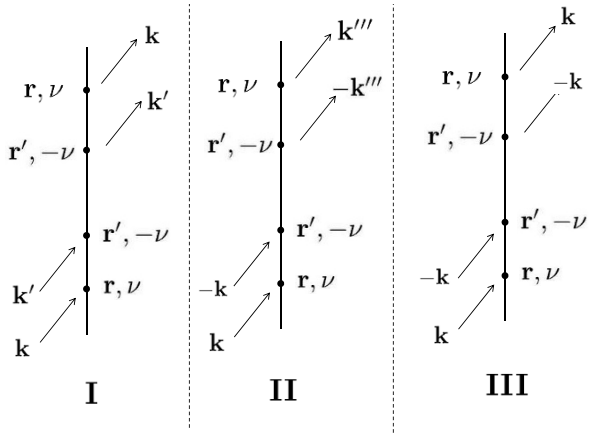


*Figure 5: Diagrams symbolizing various contributions to the binary correlation function for opposite spins ($\nu \neq \nu'$). The diagram of type **I** corresponds to a process where two particles with opposite spins are destroyed and then re-created in exactly the same individual states (forward scattering). The type **II** diagram corresponds to the case where two particles of the same pair are destroyed, and then two particles are created in the states of another pair (pair annihilation-creation process). Finally, type **III** diagram is a special case of the previous diagrams, and yields a negligible contribution. It is the type **II** diagram that introduces the spatial dependance of the correlation function.*

The position dependent term in the correlation function can be interpreted as resulting from the interference between a process where two particles of the same pair $(\mathbf{k}, -\mathbf{k})$ are annihilated, and a second process where the two annihilated particles are from another pair $(\mathbf{k}''', -\mathbf{k}''')$; these two processes are schematized in Figure 6.

According to (1) and (69), we have:

$$\left|\widetilde{\Psi}\left(\mathbf{k}\right)\right\rangle = v_{\mathbf{k}}\left|n_{\mathbf{k}}=0; n_{-\mathbf{k}}=0\right\rangle \prod_{\otimes \mathbf{k}' \neq \mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}'}\right\rangle \tag{71}$$

If $\mathbf{k} \neq \mathbf{k}'''$, the two states $\left|\widetilde{\Psi}\left(\mathbf{k}\right)\right\rangle$ and $\left|\widetilde{\Psi}\left(\mathbf{k}'''\right)\right\rangle$ are neither identical, nor orthogonal; they actually have identical components on all the pairs of states different from $(\mathbf{k}, -\mathbf{k})$ and $(\mathbf{k}''', -\mathbf{k}''')$, but these two pairs have the same component only for states where the 4 populations are zero. We can then write:

$$\left\langle\widetilde{\Psi}\left(\mathbf{k}'''\right)\right|\left.\widetilde{\Psi}\left(\mathbf{k}\right)\right\rangle = v^{*}_{\mathbf{k}'''} u_{\mathbf{k}'''} u^{*}_{\mathbf{k}} v_{\mathbf{k}} \tag{72}$$

Inserting this result in (70) yields relation (66), whose spatial dependence does come from the interference between the two processes schematized in Figure 6.

One could also use relation (68) of Complement A_{XVII} to express the product $\Psi_{\nu'}(\mathbf{r}')\Psi_{\nu}(\mathbf{r})$ as a function of a sum of pair annihilation operators. This is another way of understanding the role of pairs in the determination of the binary correlation function expression (66).

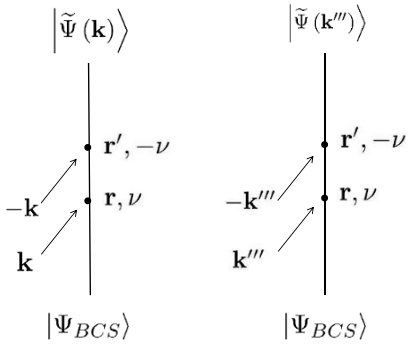


Figure 6: Diagram symbolizing two pair annihilation processes from the initial state $|\Psi_{BCS}\rangle$, leading respectively to the states $\left|\widetilde{\Psi}\,(\mathbf{k})\right\rangle$ and $\left|\widetilde{\Psi}\,(\mathbf{k}''')\right\rangle$. As these two states are not orthogonal, an interference effect occurs that is at the origin of the position dependent part of the binary correlation function.

2-c. Properties of the pair wave function, coherence length

The pair wave function plays an important role in the BCS theory, and not only for the binary correlation functions, as we already mentioned. Its range determines the coherence length of the system, and its norm is also related to the number of quanta present in the field of condensed pairs (Complement $\mathrm{A_{XVII}}$).

α. Form of the pair wave function

As the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ only depend on the modulus of $\mathbf{k}$, we can apply the Fourier transform formulas for this case – see Appendix I, relation (52). Replacing the discrete summation by an integral, the pair wave function (67) becomes:

$$\phi_{\mathrm{pair}}\,(\mathbf{r}) = \left(\frac{1}{2\pi}\right)^3 \int \mathrm{d}^3k\; e^{-i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}} v_{\mathbf{k}} = \frac{1}{(2\pi)^2}\;\frac{\Delta}{r}\int_0^\infty k\mathrm{d}k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}}\;\sin kr \tag{73}$$

Therefore,the pair wave function is real. Figure 1 gives a plot of $u_{\mathbf{k}}v_{\mathbf{k}}$ as a function of the energy e_k; it presents a maximum in the vicinity of the surface of the Fermi sphere, with a peak whose width is of the order of Δ. More details on the role this pair wave function plays in the correlation functions are given in Complement $\mathrm{A_{XVII}}$.

The Fourier transform of $\phi_{\mathrm{pair}}(\mathbf{r})$ is thus concentrated around values of the modulus of $\mathbf{k}$ of the order of the Fermi wave vector k_F. Its spread δk is such that the corresponding variation of energy e_k is of the order of the gap Δ, which leads to the condition:

$$\frac{\hbar^2}{2m} 2k_F\;\delta k \simeq \Delta \quad \text{that is} \quad \frac{\delta k}{k_F} \simeq \frac{\Delta}{E_F} \tag{74}$$

This wave function oscillates[9] as a function of the position $\mathbf{r}$, at a spatial frequency approximately equal to the wave vector at the Fermi surface. These oscillations are damped over a length of the order of ξ_{pair} defined by (the arbitrary factor $2/\pi$ is introduced to match the usual definition found in the literature):

$$\xi_{\text{pair}} = \frac{2}{\pi \delta k} = \frac{2\hbar^2 k_F}{\pi m \Delta} = \frac{1}{k_F}\frac{4E_F}{\pi \Delta} \tag{75}$$

which is of the order of the distance between fermions, multiplied by the ratio E_F/Δ, very large compared to 1. Each fermion pair extends over a relatively large volume, leading to a strong overlap between pairs. In a superconductor, the length ξ_{pair} is called the "coherence length"[10]; it characterizes the capacity of the physical system to adapt to spatial constraints, and plays a role analogous to the "healing length" in systems of condensed bosons (Complement C_{XV}, § 4-b).

We have shown that the pairing significantly modifies the correlation functions for opposite spins: the particles now become correlated, whereas this was absolutely not the case for an ideal gas. It is a positive correlation, leading to a bunching tendency (the opposite of a Pauli exclusion); this explains the decrease in the interaction energy of the particles (§ 2-c-γ). On the other hand, the pairing has no significant effect on the correlation function of particles having the same spin direction; it remains similar to the correlation function of an ideal gas, with an exchange hole whose width is of the order of $1/k_F$. Relation (75) indicates that the width of this exchange hole is much smaller than the distance ξ_{pair} over which the modifications of the correlation function for opposite spins occur.

As mentioned before, the pair wave function has little in common with the initial wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ used in Chapter XVII to build the N-particle variational ket, since (3) shows that the Fourier transform of χ is $g_{\mathbf{k}} = v_{\mathbf{k}}/u_{\mathbf{k}}$. This is not surprising: when building a trial ket by the repetitive action of the same pair creation operator, we do not simply juxtapose those pairs. The antisymmetrization effects are dominant, and in each term of the expansion to the power P of operator $\sum_k g_{\mathbf{k}}\, a^\dagger_{\mathbf{k},\uparrow} a^\dagger_{-\mathbf{k},\downarrow}$ in relation (1), the result is zero each time the same value of $\mathbf{k}$ is repeated (the square of a fermionic creation operator is zero; two fermions cannot occupy the same individual state). This is why the antisymmetrization effects completely remodel the pairs formed in the system of N identical particles.

β. Norm of the pair wave function

According to (67), the component on $|\mathbf{k}\rangle$ of the ket $|\phi_{\text{pair}}\rangle$ associated with the wave function $\phi_{\text{pair}}(\mathbf{r})$ is written as:

$$\langle \mathbf{k} \,|\phi_{\text{pair}}\rangle = \frac{1}{L^{3/2}} u_{\mathbf{k}} v_{\mathbf{k}} \tag{76}$$

[9] The existence of this oscillation is confirmed by the fact that its integral over the entire space is practically zero; this integral is proportional to $u_0 v_0$, i.e. $u_0\sqrt{1-u_0^2}$, which is indeed practically zero since $u_0 \simeq 1$.

[10] The coherence length ξ_{pair} should not be confused with the (London) "penetration depth" that characterizes the magnetic field exclusion from a superconductor, and that depends on the charges of the particles.

(the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are even). The square of this ket's norm is therefore:

$$\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{1}{L^3}\sum_{\mathbf{k}} |u_{\mathbf{k}} v_{\mathbf{k}}|^2 \tag{77}$$

Replacing the discrete summation by an integral, we get:

$$\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{1}{L^3}\left(\frac{L}{2\pi}\right)^3 \int \mathrm{d}^3 k \;\; |u_{\mathbf{k}} v_{\mathbf{k}}|^2 = \frac{\Delta^2}{4L^3}\int_0^\infty D\left(e_k\right) \;\; \frac{\mathrm{d}e_k}{\left(e_k - \overline{\mu}\right)^2 + \Delta^2} \tag{78}$$

where, in the second equality, we chose e_k as the integral variable, introducing the density of states $D\left(e_k\right)$ defined in (39). The function to be integrated converges since $D\left(e_k\right)$ varies as $\sqrt{e_k}$; this function is concentrated around $e_k = \overline{\mu}$, spreading over a width $\Delta \ll \overline{\mu}$. We can then, to a good approximation, replace $D\left(e_k\right)$ by its value D_F at the Fermi energy, and extend the lower bound of the integral to $-\infty$. As the integral of a Lorentz function is known (Appendix II, § 1-b):

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}e_k}{\left(e_k - \overline{\mu}\right)^2 + \Delta^2} = \frac{\pi}{\Delta} \tag{79}$$

we get:

$$\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{\pi}{4L^3} D_F \Delta \tag{80}$$

We showed in § 2-a-β of Complement $\mathrm{A_{XVII}}$ that the norm $\sqrt{\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle}$ yields the average value of the field $\langle \Phi_{\text{pair}}\left(\mathbf{R}\right)\rangle$, hence the value of the order parameter. In § 2-b-β of that same complement we showed that the square of the norm, equal to $\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle$, yields the large distance behavior of the average value $\left\langle \Phi^\dagger_{\text{pair}}\left(\mathbf{R}\right) \Phi_{\text{pair}}\left(\mathbf{R}'\right)\right\rangle$; the quantity $\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle$ is thus related[11] to the field pair intensity (or, in other words, to the total number of quanta in that field).

In addition, we just saw in the above § 2-b-α that a peak in the momentum distribution signaled the presence of a Bose-Einstein condensation. Inserting (76) into (58) shows that this peak height is:

$$\delta n^{(2)}_{\mathbf{k}} = |u_{\mathbf{k}} v_{\mathbf{k}}|^2 = L^3 \left|\langle \mathbf{k} \,|\phi_{\text{pair}}\rangle\right|^2 \tag{81}$$

The total particle number associated with this peak is:

$$\sum_{\mathbf{k}} \delta n^{(2)}_{\mathbf{k}} = L^3 \langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{\pi}{4} D_F \Delta \tag{82}$$

Consequently, the square of the norm, $\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle$, multiplied by the volume, is also the total particle number in the condensation peak found in § 2-b-α, which confirms the previous interpretation.

[11] The pair field operators Φ_{pair} and $\Phi^\dagger_{\text{pair}}$ do not exactly satisfy the boson commutation relations (Complement $\mathrm{A_{XVII}}$); the operator $\Phi^\dagger_{\text{pair}}\Phi_{\text{pair}}$ is thus not, stricto sensu, an operator giving the number of quanta in the pair field.

γ. Link to the interaction energy

The energy term on the third line of (7) can be written, in the zero range potential approximation (29):

$$\sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\ v_{\mathbf{k}}^{*}\ u_{\mathbf{k}}\ v_{\mathbf{k}'}\ u_{\mathbf{k}'} = -V \sum_{\mathbf{k}} v_{\mathbf{k}}^{*}\ u_{\mathbf{k}} \sum_{\mathbf{k}'} v_{\mathbf{k}'}\ u_{\mathbf{k}'}$$
$$= -VL^{6} \left|\phi_{\text{pair}}(0)\right|^{2} \tag{83}$$

This result yields an energy proportional to V and to the probability that the two components of a pair, described by the wave function $\phi_{\text{pair}}(\mathbf{r})$, are found at the same point; this makes sense since the pair size is very large compared to the interaction potential range.

δ. Non-diagonal order

When studying Bose-Einstein condensation for bosons, we showed in § 3 of Complement A_{XVI} that the one-particle non-diagonal correlation function did not go to zero at large distance, when a significant fraction of the particles occupy the same individual state. Nevertheless, in § 2-a of the present complement, we found that this was not the case for a system of paired fermions, where the non-diagonal order goes to zero over a microscopic distance. This can be understood from a physical point of view, since, in the present case, there is no accumulation of particles in the same individual quantum state. On the other hand, we saw in § 2-b-α that the center of mass of the pairs of particles is subject to a phenomenon of partial accumulation, reminiscent of a Bose-Einstein condensation. It is thus natural to examine the properties of the non-diagonal functions relative to pairs, and compute the "non-diagonal position" average value:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r})\Psi_{\downarrow}^{\dagger}(\mathbf{r})\Psi_{\downarrow}(\mathbf{r}')\Psi_{\uparrow}(\mathbf{r}') \right\rangle \tag{84}$$

With two positions $\mathbf{r}'$ on the right, and two positions $\mathbf{r}$ on the left, this expression is the exact transposition to two particles of the one-particle non-diagonal function: a couple of particles with opposite spins are annihilated at point $\mathbf{r}'$, and then recreated at a different point $\mathbf{r}$. In a more general way, we are going to evaluate the 4-point average value:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}_3)\Psi_{\uparrow}(\mathbf{r}_4) \right\rangle \tag{85}$$

which, following the same computation steps as for the one-particle functions, can be written as:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}'_2)\Psi_{\uparrow}(\mathbf{r}'_1) \right\rangle$$
$$= \frac{1}{L^{6}} \sum_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''} e^{i\left[\left(\mathbf{k}\cdot\mathbf{r}'_1+\mathbf{k}'\cdot\mathbf{r}'_2\right)-\left(\mathbf{k}''\cdot\mathbf{r}_2+\mathbf{k}'''\cdot\mathbf{r}_1\right)\right]} \left\langle a_{\mathbf{k}''',\uparrow}^{\dagger} a_{\mathbf{k}'',\downarrow}^{\dagger} a_{\mathbf{k}',\downarrow} a_{\mathbf{k},\uparrow} \right\rangle \tag{86}$$

In this equality, the matrix elements have already been evaluated in § 2-b-β. We are going to show that this expression can be written as:

$$\left\langle \Psi^{\dagger}_{\uparrow}(\mathbf{r}_1)\Psi^{\dagger}_{\downarrow}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}'_2)\Psi_{\uparrow}(\mathbf{r}'_1)\right\rangle \\ = G_1(\mathbf{r}_1,\uparrow;\mathbf{r}'_1,\uparrow)\; G_1(\mathbf{r}_2,\downarrow;\mathbf{r}'_2,\downarrow) + \phi^{*}_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\; \phi_{\text{pair}}\left(\mathbf{r}'_1-\mathbf{r}'_2\right) \tag{87}$$

where the one-particle non-diagonal distribution G_1 has been defined in relation (49) for the case $\nu = \nu'$, and the pair wave function ϕ_{pair} in relation (67). This equality is the same as relation (72) of Complement A_{XVII}, but is now obtained by another method.

Demonstration:

To compute expression (86), we distinguish several cases, as already explained on several occasions:

(**I**) Forward scattering terms; if the annihilation operators do not act on two states of the same pair ($(\mathbf{k} \neq -\mathbf{k}')$, this term will be non-zero only if $\mathbf{k}''' = \mathbf{k}$ and $\mathbf{k}'' = \mathbf{k}'$, in which case it is written:

$$\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}_1\right)+\mathbf{k}'\cdot\left(\mathbf{r}'_2-\mathbf{r}_2\right)\right]}\,|v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 = G_1(\mathbf{r}_1,\uparrow,\mathbf{r}'_1,\uparrow)\; G_1(\mathbf{r}_2,\downarrow,\mathbf{r}'_2,\downarrow) \tag{88}$$

As we already mentioned, for example in Chapter XVII (§ D-1-a), the constraints on the summation indices can be ignored if the size of the system, L, is macroscopic; the two summations then become independent.

(**II**) Terms corresponding to the annihilation-creation of different pairs; if $\mathbf{k} = -\mathbf{k}'$ and $\mathbf{k}'' = -\mathbf{k}'''$ but $\mathbf{k} \neq \mathbf{k}'''$ (annihilation-creation of different pairs), we get the contribution:

$$\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}''} e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}'_2\right)-\mathbf{k}''\cdot\left(\mathbf{r}_2-\mathbf{r}_1\right)\right]}\; u^{*}_{\mathbf{k}}v_{\mathbf{k}}u_{\mathbf{k}''}v^{*}_{\mathbf{k}''} = \phi_{\text{pair}}(\mathbf{r}'_1-\mathbf{r}'_2)\; \phi^{*}_{\text{pair}}(\mathbf{r}_2-\mathbf{r}_1) \tag{89}$$

As, in addition, the function $\phi_{\text{pair}}(\mathbf{r})$ is even, we indeed obtain the second term of the right-hand side of (87).

(**III**) If $\mathbf{k} = -\mathbf{k}'$ and $\mathbf{k}'' = -\mathbf{k}'''$, and furthermore $\mathbf{k} = \mathbf{k}'''$ (annihilation-creation of the same pair), we get:

$$\frac{1}{L^6}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'_1+\mathbf{r}_2-\mathbf{r}'_2-\mathbf{r}_1\right)}\,|v_{\mathbf{k}}|^2 \tag{90}$$

This term is negligible as it is proportional to $\langle N\rangle / L^6$ when all the positions are the same, whereas the term (**I**) is proportional to $\langle N\rangle^2 / L^6$.

Let us now assume the positions can be grouped two by two: $\mathbf{r}_1$ and $\mathbf{r}_2$ are close to each other, as are $\mathbf{r}'_1$ and $\mathbf{r}'_2$, but that the two groups' positions are further away from each other. Under these conditions, the first term in G_1 on the right-hand side of (87), which has a microscopic range in $(\mathbf{r}_1 - \mathbf{r}'_1)$ and $(\mathbf{r}_2 - \mathbf{r}'_2)$, becomes very small. We are then left with the product of the pair wave functions. It follows that the non-diagonal

correlation function is simply the product of pair wave functions calculated at the relative positions[12].

In the particular case where $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ and $\mathbf{r}'_2 = \mathbf{r}'_1 = \mathbf{r}'$, we get the pair correlation function (84), which obeys:

$$\left\langle \Psi^\dagger_\uparrow(\mathbf{r})\Psi^\dagger_\downarrow(\mathbf{r})\Psi_\downarrow(\mathbf{r}')\Psi_\uparrow(\mathbf{r}') \right\rangle \underset{\mathbf{r}-\mathbf{r}'\to\infty}{\longrightarrow} \left|\phi_{\text{pair}}(0)\right|^2 \tag{91}$$

This non-zero long distance limit signals the existence of a non-diagonal order for the two-particle density operator. It comes, as was already the case for the pair wave function, from contribution (**II**), meaning from terms corresponding to the annihilation-creation of different pairs. This situation is reminiscent of what we encountered in the case of a condensed boson gas; but in the present case, the non-diagonal order concerns the pairs and not the individual particles.

3. Physical discussion

In an ideal gas of fermions, and as we saw in Complement A_{XVI}, there already exist strong correlations between the particles, simply due to their indistinguishability (a purely statistical effect). In the presence of attractive interactions, the BCS mechanism introduces additional correlations (dynamic correlations) that lower the total system energy. We are going to show that this decrease in energy comes from a slight imbalance between an increase in kinetic energy and a decrease of the potential energy, the latter one slightly surpassing the first one.

For clarity, we shall discuss this using the short-range potential approximation (§ 1-c) where all the matrix elements of the interaction potential are replaced by a constant $-V$ (V being positive); all the $\Delta_{\mathbf{k}}$ are then equal to the same gap Δ.

3-a. Modification of the Fermi surface and phase locking

The energy written in (7) first includes a kinetic energy term, then a mean field term expressed in terms of the average particle number. If that average number is constant, this term is independent of the quantum state of the system, and hence not related to the BCS pairing mechanism. By contrast, the last term in (7), which is the one we optimized in the variational calculation, is far more interesting; we call it the "pairing term", and use the words "pairing potential energy" or else "condensation energy" for its optimized value E_{paired}. As the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$ are real, E_{paired} can be written as:

$$E_{\text{paired}} = -V\left[\sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}\right]^2 \tag{92}$$

where the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$ take the optimized values given in (27). We see that to get a large condensation energy, the sum $\sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}$ must take the largest possible value.

[12] As mentioned in note 8, the center of mass variables are not included in (87) since we assumed all the pairs to be at rest. If this were not the case, the long-range factorization of the non-diagonal order would be expressed as the product of a function of the two variables $(\mathbf{r}'_1 - \mathbf{r}'_2)$ and $(\mathbf{r}'_1 + \mathbf{r}'_2)/2$ by the complex conjugate of that same function of the two variables $(\mathbf{r}_1 - \mathbf{r}_2)$ and $(\mathbf{r}_1 + \mathbf{r}_2)/2$ (in other words by the product of a function of $\mathbf{r}'_1$ and $\mathbf{r}'_2$ by the complex conjugate of that same function of the two variables $\mathbf{r}_1$ and $\mathbf{r}_2$).

α. Compromise between the kinetic and potential energies

In an ideal gas, the ground state is the one for which all the individual states of energy lower than the Fermi level (chemical potential $\mu = E_F$) are occupied by one particle, and all the states above are totally empty. In the **k**-space, the particles each occupy a state inside one of the two Fermi spheres of radius k_F (with $E_F = \hbar^2 k_F^2/2m$), one associated with the + spin state, and one associated with the − spin state. Using the ket (1), such a state simply corresponds to the case where:

$$\begin{array}{llll} u_{\mathbf{k}} = 0 & \text{and} & v_{\mathbf{k}} = 1 & \text{for } k < k_F \\ u_{\mathbf{k}} = 1 & \text{and} & v_{\mathbf{k}} = 0 & \text{for } k > k_F \end{array} \qquad \text{(ideal gas)} \tag{93}$$

Whatever the value of **k**, one or the other of the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ will be zero, and so is the product $u_{\mathbf{k}} v_{\mathbf{k}}$; the condensation energy of an interacting system remains equal to zero as long as the state of the system does not differ from the ideal gas state. A condensation energy can only be obtained via a deformation of the Fermi distribution.

It is the attractive interactions that actually distort this distribution to create an overlap between regions where both functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are different from zero, as can be seen in Figure 1. This allows minimizing the pairing energy (92), but involves a transfer of particles from the inside of the Fermi sphere to the outside, hence toward states of higher kinetic energy; this has a cost in terms of kinetic energy. The optimization we performed amounts to looking for the most favorable balance between the gain in potential energy and the cost in kinetic energy. The condensation energy is proportional to the square of the integral of the dashed line curve in Figure 1, which presents a maximum in the vicinity of the Fermi energy E_F. The largest contributions come from energies close to E_F, over a width of the order of a few Δ - but the figure also shows that the contributions to the condensation energy spread relatively far from the Fermi surface (the curve only decreases as the inverse of the energy's distance from its maximum). The Fermi surface, which was perfectly defined for an ideal gas, becomes blurred over a certain energy domain.

The two-particle correlation function expresses in more detail this optimization of the attractive potential energy. Relation (64) shows that, for parallel spins, no significant change of the correlation function occurs, when compared to that of an ideal gas (for which an exchange hole is already present in the binary correlation function) – hence no significant change of the corresponding interaction energy. This lack of effect comes from the fact that the BCS wave function only pairs particles with opposite spins. On the other hand, when the spin directions are opposite, relation (66) shows that the probability of presence of two particles at a short distance from each other is increased; the larger the pair wave function modulus at the origin ($\mathbf{r} = \mathbf{r}'$), the higher this increase will be. It directly yields the gain in the attractive potential energy.

In other words, the BCS gain in energy comes from the fact that, because of the interactions, the system changes its wave function to optimize its pairing potential energy. It develops correlations that go beyond that of an ideal gas; they are referred to as "dynamic correlations", as opposed to the statistical correlations (coming solely from the indistinguishability of the particles, such as those studied in Complement A_{XVI}). This produces a deformation of the ideal gas Fermi distribution that, instead of presenting an abrupt transition between the occupied and empty states (perfectly well defined Fermi sphere), presents at its edge a more progressive transition region (blurred Fermi sphere). The system's state vector then becomes a superposition of states where the particle

number in each pair of states $(\mathbf{k}, -\mathbf{k})$ fluctuates. The potential energy term that drives the BCS mechanism is the "pair annihilation-creation" term computed in § 3-a-β of Complement B_{XVII}, and that is schematized on its Figure 4. It includes a sum of terms containing non-diagonal matrix elements of the form:

$$\left\langle n_{(\mathbf{k},-\mathbf{k})} = 2;\ n_{(\mathbf{k}'',-\mathbf{k}'')} = 0 \right| W_2 \left| n_{(\mathbf{k},-\mathbf{k})} = 0;\ n_{(\mathbf{k}'',-\mathbf{k}'')} = 2 \right\rangle \tag{94}$$

(the occupation numbers of all the other pairs remaining the same); between the ket and the bra, a pair $(\mathbf{k}'', -\mathbf{k}'')$ is replaced by another one $(\mathbf{k}, -\mathbf{k})$. The BCS energy gain is due to the summation over all these non-diagonal terms; they are sensitive to the coherence of the state vector between these two components (where the numbers of pairs fluctuate in a correlated way) and hence to their relative phase.

β. Phase locking and cooperative effects

In the computation of the § 1-b, the minimization of the energy led us to choose the phases of all the $v_{\mathbf{k}}$ to be equal, and they simply disappeared from the following calculations. To discuss the physical process at work, it is useful to reintroduce them with their non-specified values before the optimization, as they appeared in (14); when all the matrix elements of the interaction potential are equal, the average value of the pairing energy is written:

$$E_{\text{paired}} = -V \sum_{\mathbf{k},\mathbf{k}'} \sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}}\ \sin\theta_{\mathbf{k}'} \cos\theta_{\mathbf{k}'}\ e^{2i(\zeta_{\mathbf{k}'} - \zeta_{\mathbf{k}})} \tag{95}$$

We mentioned above that adding to the phase $\zeta_{\mathbf{k}}$ of each $v_{\mathbf{k}}$ any given common phase χ, left all the results unchanged. The energy is invariant with respect to a symmetry of the wave function, the one that concerns the global phase of the $v_{\mathbf{k}}$. It is often called the "$U(1)$ symmetry", referring to the $U(1)$ unitary symmetry group of rotations around a circle, isomorphic to the group of phase changes for a complex number.

On the other hand, changing one by one the phases of the $v_{\mathbf{k}}$, leads to an obvious reduction (in absolute value) of the pairing energy (95): in the complex plane, vectors that were perfectly aligned, now take different directions and the modulus of their sum is reduced. We saw that this term is at the origin of the gain in energy provided by the BCS mechanism; it is clearly linked to the acquisition of a common phase by all the pairs of individual states. This is an example of a phenomenon called in physics "spontaneous symmetry breaking": whatever the phase of each $v_{\mathbf{k}}$, which can take on any value, it is essential that it be the same for all, otherwise most of the gain in energy will be lost. In a similar way, in the ferromagnetic transition in a solid, the space direction along which the spins will align is not a priori fixed, but it is essential that it be the same[13] for all the spins.

Note finally the cooperative character of the energy gain obtained, which, mathematically, corresponds to the presence of a double summation over $\mathbf{k}$ and $\mathbf{k}'$. Starting

[13] For an ensemble of spins parallel to any given direction, each spin is in a state where the relative phase between the components on $|+\rangle$ and $|-\rangle$ is the same. For the BCS mechanism, it is the phase between the components where the occupation number of the couple of states $\mathbf{k}, -\mathbf{k}$ is 0 or 2 that takes on a value independent of $\mathbf{k}$. The corresponding energy lowering results from an interference effect between states where two pairs $\mathbf{k}$ and $\mathbf{k}'$ have respective occupation numbers $n_{\mathbf{k}} = 2, n_{\mathbf{k}'} = 0$ and $n_{\mathbf{k}} = 0, n_{\mathbf{k}'} = 2$; therefore it cannot be directly expressed in terms of pair populations.

from a perfectly phase locked situation, destroying the phase locking of a single pair leads to an energy loss proportional to the number of pairs that remained phase locked; the individual energy of a single pair is not what is at stake. On the other hand, starting from a situation where the phases of all the pairs are random, changing a single phase $\zeta_{\mathbf{k}}$ barely modifies the average energy. We are in the presence of a cooperative effect: the greater the number of pairs that are already phase locked, the higher the tendency for a new pair to become phase locked; this tendency can be seen, in a way, as resulting from a mean field created in a cumulative way by all the other pairs. Here again we see the analogy with a ferromagnetic material where, the greater the number of spins already aligned, the higher the gain in energy with the alignment of a new spin.

3-b. Gain in energy

We now compute the gain in energy resulting from the pairs' formation. We first insert in (7) relations (27), which yield the optimal values of the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$, and use the definition (25) for E_k; we also take all the potential matrix elements to be equal to the same constant $-V$. Since we then have:

$$|v_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 - \frac{\xi_k}{E_k}\right) = \frac{1}{2}\frac{E_k - \xi_k}{E_k} \tag{96}$$

and:

$$u_{\mathbf{k}} v_{\mathbf{k}} = \frac{1}{2}\sqrt{1 - \left(\frac{\xi_k}{E_k}\right)^2} = \frac{\Delta}{2E_k} \tag{97}$$

we get:

$$\begin{aligned}&\left\langle \widehat{H} - \mu \widehat{N} \right\rangle_{\mathrm{BCS}} \\ &= -\frac{V}{4}\langle N\rangle^2 + \sum_{\mathbf{k}} \left[\xi_k \, \frac{E_k - \xi_k}{E_k} - \frac{V\Delta^2}{4} \sum_{\mathbf{k},\mathbf{k}'} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \frac{1}{\sqrt{(\xi_{k'})^2 + \Delta^2}} \right]\end{aligned} \tag{98}$$

The first term on the right-hand side is the mean field term (as before, we have neglected 1 compared to the total number of particles). The second one corresponds to the kinetic energy, and the third one, to the interaction between pairs:

$$\frac{-V\Delta^2}{4} \sum_{\mathbf{k},\mathbf{k}'} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \frac{1}{\sqrt{(\xi_{k'})^2 + \Delta^2}} = -\frac{\Delta^2}{2} \sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{99}$$

where to get the second equality we have used relation (36) to eliminate the summation over $\mathbf{k}'$. Using again the definition (25) for E_k, we get:

$$\left\langle \widehat{H} - \mu \widehat{N} \right\rangle_{\mathrm{BCS}} = -\frac{V}{4}\langle N\rangle^2 + \sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \left[\xi_k \left(\sqrt{(\xi_k)^2 + \Delta^2} - \xi_k \right) - \frac{\Delta^2}{2} \right] \tag{100}$$

On the right-hand side of this expression, the first term, corresponding to the mean field, is of no particular interest. The second one accounts for the change in energy due to the dynamic correlations introduced by the interactions; it characterizes the BCS mechanism.

We first check the convergence of its summation over $\mathbf{k}$, for a fixed value of Δ. This is not the case for each term in the parenthesis, which tends toward a constant when $k \to \infty$, leading to a summation in $1/\xi_k \sim 1/k^2$ that diverges. We are now going to show that the divergent terms cancel each other. We can write:

$$\sqrt{(\xi_k)^2 + \Delta^2} - \xi_k \sim \frac{\Delta^2}{2\xi_k} - \frac{3}{8}\frac{\Delta^4}{(\xi_k)^3} + .. \tag{101}$$

and thus:

$$\frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}}\left[\xi_k\left(\sqrt{(\xi_k)^2 + \Delta^2} - \xi_k\right) - \frac{\Delta^2}{2}\right] \sim -\frac{3}{8}\frac{\Delta^4}{(\xi_k)^3} + .. \tag{102}$$

We have just shown that the divergent terms in the infinite summation of $\mathbf{k}$ in relation (100) cancel each other between the kinetic and interaction terms; the function in the summation goes to zero, for large k values, as $1/(\xi_k)^3 \sim 1/k^6$, which ensures the convergence and does not require the introduction of a cutoff frequency k_c (apart from the one we had to introduce before to ensure a finite value for Δ). This was also the case for the total number of particles. We thus see that once we have introduced an upper boundary (cutoff) k_c in the integral determining the gap Δ, all the other important physical quantities remain finite, without having to reimpose this cutoff frequency.

The precise determination of the energy requires, in general, the computation of somewhat complex integrals. It will not be detailed here, but yields the result:

$$\left\langle \widehat{H} - \mu\widehat{N} \right\rangle_{\text{BCS}} = \left\langle \widehat{H} - \mu\widehat{N} \right\rangle_0 - \frac{V}{4}\left[\langle N \rangle\right]^2 - \frac{1}{2}\Delta^2 D_F \tag{103}$$

(remember that D_F is the density of individual states at the surface of the Fermi sphere, and is proportional to the volume $\mathcal{V} = L^3$). Finally, the energy gain resulting from the BCS pairing is given by:

$$\delta E = -\frac{1}{2}\Delta^2 D_F \tag{104}$$

It can be shown that the values of ξ_k that contribute most to the energy are those that are lower than or comparable to the gap Δ; the energy change linked to the pairing phenomenon is mainly located in the vicinity of the Fermi surface. This result is often interpreted by saying that, in an ensemble of $\Delta \times D_F$ pairs, each pair gains an energy of the order of Δ, which explains the Δ^2 dependence of (104); while this image is simple, it has its shortcomings (see note 13).

3-c. Non-perturbative character of the BCS theory

Generally, the most basic way to take the interactions into account is to use a first order perturbation theory (Chapter XI), where the energy correction is the average value, in the initial non-perturbed state, of the perturbation Hamiltonian. Applied here, the first order correction to the energy is obtained by inserting the values (93) into (7).

The first term (kinetic energy) on the right-hand side of (7) is unchanged, and the third one remains zero since, according to (93), the product $u_{\mathbf{k}} v_{\mathbf{k}}$ is always zero for any value of $\mathbf{k}$. We are left with the second term, which produces a mean field correction. To the next perturbation order, the effect of the potential is to change the ground state by transferring pairs of particles, initially both inside the Fermi sphere, toward individual states whose momenta fall outside the sphere (all the while keeping the total momentum constant); this changes at the same time the average kinetic energy (which is increased) and the interaction potential energy. The computations become more and more complex as the perturbation order increases. And above all, it is clear that this approach to higher and higher perturbation orders cannot account for the existence of the gap obtained in (46): as the function $\Delta(V)$ has all its derivatives with respect to V equal to zero at $V = 0$, it can not be expanded as a series in V.

The BCS theory is a non-perturbative method that solves this difficulty. However, it is not an exact method since it is a variational method, but the chosen wave function is sufficiently well adapted to allow the inclusion of important physical effects, without using any perturbation theory.

4. Excited states

Up to now, we only studied the ground state of the system of attractive fermions. As soon as the temperature is no longer zero, excited states of the system begin to be populated. In this last section, we shall give a survey of the BCS theory predictions concerning the excited states. A study of the BCS theory at non-zero temperature can be found in more specialized books.

4-a. Bogolubov-Valatin transformation

Relations (E-3) and (E-4) of Chapter XVII define the Bogolubov-Valatin transformations of the creation and annihilation operators of spin 1/2 fermions. With the notation of the present complement where the spin directions are explicit, they become:

$$\begin{aligned} b_{\mathbf{k}} &= u_{\mathbf{k}}\, a_{\mathbf{k},\uparrow} - v_{\mathbf{k}}\, a^{\dagger}_{-\mathbf{k},\downarrow} \\ b_{-\mathbf{k}} &= u_{\mathbf{k}}\, a_{-\mathbf{k},\downarrow} + v_{\mathbf{k}}\, a^{\dagger}_{\mathbf{k},\uparrow} \end{aligned} \tag{105}$$

which, by conjugation, yield the definitions of the Hermitian conjugate operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$:

$$\begin{aligned} b^{\dagger}_{\mathbf{k}} &= u^{*}_{\mathbf{k}}\, a^{\dagger}_{\mathbf{k},\uparrow} - v^{*}_{\mathbf{k}}\, a_{-\mathbf{k},\downarrow} \\ b^{\dagger}_{-\mathbf{k}} &= u^{*}_{\mathbf{k}}\, a^{\dagger}_{-\mathbf{k},\downarrow} + v^{*}_{\mathbf{k}}\, a_{\mathbf{k},\uparrow} \end{aligned} \tag{106}$$

For each value of $\mathbf{k}$, we get a general transformation of the four initial creation and annihilation operators a or $a^{\dagger}$ into four new operators b and $b^{\dagger}$. We showed in Chapter XVII that these operators obey the usual anticommutation relations of fermions.

We also saw in that chapter that the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ defined in (2):

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\, a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{107}$$

is an eigenvector of the two operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ with a zero eigenvalue:

$$\begin{aligned} b_{\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle &= u_{\mathbf{k}} v_{\mathbf{k}} \left[a^{\dagger}_{-\mathbf{k},\downarrow} - a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle = 0 \\ b_{-\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle &= u_{\mathbf{k}} v_{\mathbf{k}} \left[-a^{\dagger}_{\mathbf{k},\uparrow} + a^{\dagger}_{\mathbf{k},\uparrow}\right] |0\rangle = 0 \end{aligned} \tag{108}$$

It follows that the variational ket $|\Psi_{\mathrm{BCS}}\rangle$ of the ground state, written in (1), is an eigenvector, for any value of $\mathbf{k}$, of all the operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$, with a zero eigenvalue. It is therefore also an eigenket of all the operators $b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}} b_{-\mathbf{k}}$ with a zero eigenvalue, which is a minimum eigenvalue for operators defined as positive or zero. Furthermore, we showed that the repeated action of the creation operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$ permits obtaining other states, which are also eigenvectors of the operators $b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}} b_{-\mathbf{k}}$. We are going to show that these operators $b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}} b_{-\mathbf{k}}$ can be interpreted as corresponding to the occupation numbers of the excitations present in the physical system.

4-b. Broken pairs and excited pairs

Letting $b^{\dagger}_{\mathbf{k}}$ act on (107), we get:

$$\begin{aligned} b^{\dagger}_{\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle &= \left[|u_{\mathbf{k}}|^2 a^{\dagger}_{\mathbf{k},\uparrow} - |v_{\mathbf{k}}|^2 a_{-\mathbf{k},\downarrow} a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle = \left[|u_{\mathbf{k}}|^2 a^{\dagger}_{\mathbf{k},\uparrow} + |v_{\mathbf{k}}|^2 a^{\dagger}_{\mathbf{k},\uparrow}\right] |0\rangle \\ &= a^{\dagger}_{\mathbf{k},\uparrow} |0\rangle \end{aligned} \tag{109}$$

which is a ket normalized to unity, and obviously non-zero (as opposed to the one resulting from the action of $b_{\mathbf{k}}$). Similarly, if we consider the action of $b^{\dagger}_{-\mathbf{k}}$, we get another non-zero ket:

$$b^{\dagger}_{-\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle = a^{\dagger}_{-\mathbf{k},\downarrow} |0\rangle \tag{110}$$

These two new normalized kets are orthogonal to the initial ket $|\overline{\varphi}_{\mathbf{k}}\rangle$, since they correspond to an occupation number equal to 1, whereas the occupation numbers of $|\overline{\varphi}_{\mathbf{k}}\rangle$ are 0 and 2. In these states, a pair has been replaced by a single particle, not belonging to any pair; they are called "broken pair" states.

As the squares of the operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$ are zero, the repeated application of any of the two does not allow constructing new orthogonal states; however this can be achieved with their cross product. Letting $b^{\dagger}_{\mathbf{k}}$ act on the ket (110), we get the ket $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$:

$$\left|\overline{\varphi}_{\mathbf{k}}\right\rangle^{e} = b^{\dagger}_{\mathbf{k}} b^{\dagger}_{-\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle = \left[u^{*}_{\mathbf{k}} a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow} - v^{*}_{\mathbf{k}}\right] |0\rangle \tag{111}$$

which is another normalized ket, and orthonormal, as can be easily checked, to $|\overline{\varphi}_{\mathbf{k}}\rangle$; letting now the two operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$, in the inverse order, act on $|\overline{\varphi}_{\mathbf{k}}\rangle$, we get the same ket, $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$, within a change of sign. The components of the two states $|\overline{\varphi}_{\mathbf{k}}\rangle$ and $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$ contain occupation numbers equal to 0 or 2; the ket $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$ is called "excited pair" state. To go from $|\overline{\varphi}_{\mathbf{k}}\rangle$ to $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$, we simply switch $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, change the sign of $v_{\mathbf{k}}$, and finally take their complex conjugate (this is true for the general case, but in the BCS pairing case, as $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ take on real values, this last step becomes unnecessary).

4-c. Stationarity of the energies

Let us now show that the energies of these new states are stationary with respect to the variational parameters.

α. Broken pair

According to (109), the action of $b^\dagger_{\mathbf{k}}$ on the ground state $|\Psi_{\rm BCS}\rangle$ leads to the ket:

$$b^\dagger_{\mathbf{k}}\,|\Psi_{\rm BCS}\rangle = \left[a^\dagger_{\mathbf{k},\uparrow}\,|0\rangle\right] \otimes |\Psi'_{\rm BCS}\rangle \tag{112}$$

where $|\Psi'_{\rm BCS}\rangle$ is just the ket $|\Psi_{\rm BCS}\rangle$ whose $\mathbf{k}$ pair component has been removed from the product:

$$|\Psi'_{\rm BCS}\rangle = \prod_{\otimes \mathbf{k}'\neq\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle \tag{113}$$

The energy average value in the state (112) is the sum of three terms:

(i) the kinetic energy $e_k = \hbar^2k^2/2m$ associated with the state $a^\dagger_{\mathbf{k},\uparrow}\,|0\rangle$

(ii) the energy associated with the state $|\Psi'_{\rm BCS}\rangle$; the computation of that energy is the same as for $|\Psi_{\rm BCS}\rangle$, including the pair interaction energy, except for the fact that one less pair is now involved in the calculation. This slightly modifies the value of Δ, and hence the optimal value of the parameters $\theta_{\mathbf{k}}$; however, since the relative variation of Δ is inversely proportional to the number of particles, we shall ignore this slight modification.

(iii) finally, the interaction energy between the particle in the individual state $a^\dagger_{\mathbf{k},\uparrow}\,|0\rangle$ and the particles described by $|\Psi'_{\rm BCS}\rangle$; the pair structure of that state means that the only contributions are a direct term in:

$$\begin{aligned}\frac{1}{2}\sum_{\mathbf{k}'\neq\mathbf{k},\nu} V_{\mathbf{0}}\left\langle a^\dagger_{\mathbf{k},\uparrow}a^\dagger_{\mathbf{k}',\nu}a_{\mathbf{k}',\nu}a_{\mathbf{k},\uparrow} + a^\dagger_{\mathbf{k}',\nu}a^\dagger_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}a_{\mathbf{k}',\nu}\right\rangle &= V_{\mathbf{0}}\sum_{\mathbf{k}',\nu}\langle\widehat{n}_{\mathbf{k}',\nu}\rangle \\ &= V_{\mathbf{0}}N'\end{aligned} \tag{114}$$

(where N' is the average particle number in the state $|\Psi'_{\rm BCS}\rangle$) and an exchange term in:

$$\frac{1}{2}\sum_{\mathbf{k}'\neq\mathbf{k}} V_{\mathbf{k}'-\mathbf{k}}\left\langle a^\dagger_{\mathbf{k},\uparrow}a^\dagger_{\mathbf{k}',\uparrow}a_{\mathbf{k},\uparrow}a_{\mathbf{k}',\uparrow} + a^\dagger_{\mathbf{k}',\uparrow}a^\dagger_{\mathbf{k},\uparrow}a_{\mathbf{k}',\uparrow}a_{\mathbf{k},\uparrow}\right\rangle = -\sum_{\mathbf{k}'\neq\mathbf{k}} V_{\mathbf{k}'-\mathbf{k}}\,\langle\widehat{n}_{\mathbf{k}',\uparrow}\rangle \tag{115}$$

Note again that, for an interaction that does not act on the spins, the exchange is only possible with particles having the same spin. With the short-range potential approximation (29):

$$V_{\mathbf{0}} = V_{\mathbf{k}} = -V \tag{116}$$

the term (114) becomes equal to $-VN'$, the term (115) to $VN'/2$, and their sum simply yields a constant $-VN'/2$.

The parameters defining the variational state (113) are the set of the $\theta_{\mathbf{k}'}$ and the $\zeta_{\mathbf{k}'}$ for $\mathbf{k}'\neq\mathbf{k}$ (the dependence on the θ_k parameter, which characterized the broken pair, is no longer present). These parameters are the ones that make the energy stationary, since we neglected the slight variation of the gap linked to the disappearance of a pair; they play no role either in (i) or in (iii).

We have just confirmed that $b^\dagger_{\mathbf{k}}\,|\Psi_{\rm BCS}\rangle$ renders the energy stationary (in the framework of the variational approximation we are using). A symmetry argument shows that this is obviously also the case for the state $b^\dagger_{-\mathbf{k}}\,|\Psi_{\rm BCS}\rangle$.

β. Excited pair

In the stationary relations (26), the change of $|\overline{\varphi}_{\mathbf{k}}\rangle$ into $|\overline{\varphi}_{\mathbf{k}}\rangle^e$ simply amounts to exchanging the signs ± (the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are real); the components of the ket $|\overline{\varphi}_{\mathbf{k}}\rangle^e$ are part of the stationary solutions we have discarded in writing (27). We thus confirm that the excited pair corresponds to a stationary energy; it is actually the highest possible energy for the pair of states $(\mathbf{k}, -\mathbf{k})$.

4-d. Excitation energies

In the 4-dimensional state space associated with each pair of states $(\mathbf{k}, -\mathbf{k})$, the creation operators $b^\dagger$ acting on the $|\overline{\varphi}_{\mathbf{k}}\rangle$ permit building a new basis of 4 orthonormal states, whose average energies are stationary. They can be considered to be the approximate eigenvectors of the system Hamiltonian. We now compute the corresponding eigenvalues.

In the case of the broken pair, the excited state $|\Psi'_{\mathrm{BCS}}\rangle$ does not contain the same number of particles as $|\Psi_{\mathrm{BCS}}\rangle$; it does not make sense to directly compare their energies. To make a valid comparison, we must take into account the presence of a particle reservoir whose energy increases by μ (chemical potential) each time it absorbs a particle. In other words, we must evaluate the variations of the average value $\left\langle \widehat{H} - \mu\widehat{N} \right\rangle$.

α. Broken pair

We now show that the variation of the average value $\left\langle \widehat{H} - \mu\widehat{N} \right\rangle$ associated with the breaking of the pair is simply the energy E_k defined in (25).

To do so, we compute the variation of this average value when the state $|\Psi_{\mathrm{BCS}}\rangle$ is replaced by expression (112). Several terms come into play:

(i) The variation of the average value of the kinetic energy is the difference between the energy e_k of a particle and that of the population $|v_{\mathbf{k}}|^2$ of the pair $(\mathbf{k}, -\mathbf{k})$, with a kinetic energy of $2e_k$; this difference is therefore $e_k\left(1 - 2|v_{\mathbf{k}}|^2\right)$.

(ii) As for the potential energy, the passage from $|\Psi_{\mathrm{BCS}}\rangle$ to the ket $|\Psi'_{\mathrm{BCS}}\rangle$ changes the average particle number $\langle N\rangle$ from $2|v_{\mathbf{k}}|^2$ (initial population of the pair) to 0, so that the variation of $\langle N\rangle$ is $\delta\langle N\rangle = -2|v_{\mathbf{k}}|^2$. The mean field term $-V\langle N\rangle^2/4$ in (7) varies by $-V\langle N\rangle\delta\langle N\rangle/2$, that is $V\langle N\rangle|v_{\mathbf{k}}|^2$. The following term in (7) is zero for a short-range potential. Finally, the breaking of a pair has an impact on the binding energy in the last term of (7); if we change the dummy summation index k into k', and k' into k'', the terms that will change correspond to the terms $k = k'$ and the terms $k = k''$, which double each other. The breaking of the pair leads to an increase of energy equal to:

$$2Vu_{\mathbf{k}}v_{\mathbf{k}}\sum_{\mathbf{k}'}u_{\mathbf{k}'}v_{\mathbf{k}'} = 2\Delta\, u_{\mathbf{k}}v_{\mathbf{k}} = \Delta\sqrt{1 - \frac{\xi_k^2}{(E_k)^2}} = \frac{\Delta^2}{E_k} \tag{117}$$

where we have used (31) and (27).

(iii) We saw that the unpaired particle has a potential energy $-VN'/2 \simeq -V\langle N\rangle/2$. This energy must be added to the variation of the mean field term calculated above, to give a contribution equal to $-V\langle N\rangle\left(1 - 2|v_{\mathbf{k}}|^2\right)/2$.

We now sum all the previous contributions and add the variation of $-\mu \left\langle \widehat{N} \right\rangle$ which yields a term $-\mu \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right)$. The total variation is then:

$$\delta \left\langle \widehat{H} - \mu \widehat{N} \right\rangle = \left(e_k - \frac{V \langle N \rangle}{2} - \mu \right) \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right) + \frac{\Delta^2}{E_k} \tag{118}$$

or else, taking (13) into account:

$$\begin{aligned} \delta \left\langle \widehat{H} - \mu \widehat{N} \right\rangle &= \xi_k \left[1 - \left(1 - \frac{\xi_k}{E_k} \right) \right] + \frac{\Delta^2}{E_k} \\ &= \frac{\xi_k^2 + \Delta^2}{E_k} = E_k \end{aligned} \tag{119}$$

We find the expected result[14].

β. Excited pair

We now assume that in the product that yields $|\Psi_{\text{BCS}}\rangle$, the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ is replaced by the orthogonal ket $|\overline{\varphi}_{\mathbf{k}}\rangle^e$ written in (111), and which describes an "excited pair". We are going to show that the variation of the average value $\left\langle \widehat{H} - \mu \widehat{N} \right\rangle$ associated with that excitation is $2E_k$, twice the excitation associated with the breaking of the pair.

To show this, we must, here again, add several variations. The first one comes from the kinetic energy and yields $2e_k \left(\left|u_{\mathbf{k}}\right|^2 - \left|v_{\mathbf{k}}\right|^2\right)$, that is $2e_k \left(1 - 2\left|v_{\mathbf{k}}\right|\right)^2\big)$ – see relation (4). The second one is the mean field term introduced by the fact that the average value of the total particle number varies by $2\left(\left|u_{\mathbf{k}}\right|^2 - \left|v_{\mathbf{k}}\right|^2\right)$, which leads to a potential energy variation $-V \langle N \rangle \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right)$. We must also account for a variation of the pair binding energy, which comes from the sign change of the product $u_{\mathbf{k}} v_{\mathbf{k}}$, which doubles the term (117). Because of the change in the average particle number, the term in $\mu \langle N \rangle$ gives a contribution of $2\mu \left(\left|u_{\mathbf{k}}\right|^2 - \left|v_{\mathbf{k}}\right|^2\right)$, that is $2\mu \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right)$. Finally, all the terms found for the broken pair are just doubled here and we indeed find $2E_k$.

γ. Spectrum of the elementary excitations

We now have the energy of the three excited states associated with each pair of states: the energy E_k (doubly degenerate since it corresponds to both kets $b_{\mathbf{k}}^\dagger |\overline{\varphi}_{\mathbf{k}}\rangle$ and $b_{-\mathbf{k}}^\dagger |\overline{\varphi}_{\mathbf{k}}\rangle$) and the energy $2E_k$ (non-degenerate). The value of these energies is given in (25). Figure 7 plots the dispersion relation of these elementary excitations (energy of these excitations as a function of their momentum). The solid line corresponds to the spectrum associated with the breaking of a pair $\mathbf{k}$, $-\mathbf{k}$, during which a particle disappears, as in relations (109) and (110); as the spectrum associated with the excitation of a pair can be simply obtained by the multiplication by a factor 2, it is not plotted in the figure. The dashed lines plot those same energies for an ideal gas (no interaction) for which $\Delta = 0$. The interaction effect creates a "gap" Δ, which yields a minimum value for the excitation energy that otherwise can go to zero for an ideal gas.

Upon the breaking of a pair, we saw that, on average, the system's total population changes by $1 - 2\left|v_{\mathbf{k}}\right|^2$. The curve in Figure 7 has a different interpretation, depending on

[14] The calculation would be the same for an ideal gas; in the particular case where $\Delta = 0$ and according to (25), we would get: $E_k = |\xi_k|$.

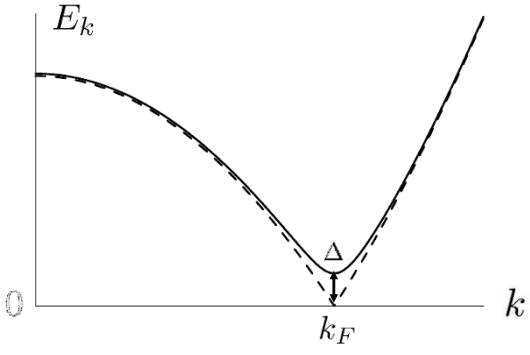


Figure 7: The solid line plots the variation as a function of k of the excitation energy $E_k = \sqrt{(\xi_k)^2 + \Delta^2}$, with $\xi_k = e_k - \mu$ and $e_k = \hbar^2 k^2/2m$ (for the sake of simplicity, we assumed $\mu \simeq \overline{\mu}$, so as to neglect the difference between $\overline{\mu}$ and μ in the expression of ξ_k). This energy presents a minimum equal to the gap Δ when k is equal the the Fermi wave vector $k = k_F$ (wave vector for which $e_k = \mu$). The dashed line curve is the same function, but for $\Delta = 0$ (zero gap), hence for an ideal gas.
The BCS theory predicts that the energy associated with the breaking of a pair is E_k, and the energy associated with the excitation of a pair is twice that amount, i.e. $2E_k$. The minimum of the energy E_k is therefore the minimum of energy that must be supplied to the system in its BCS ground state to produce one of the previously computed excitations. As explained in the text, the region on the left-hand side of the curve corresponds to "hole type" excitations (the excitation leads to a particle loss for the physical system) and the region on the right-hand side to "particle type" excitations (the physical system gains a particle).
For clarity, the figure does not take into account the mean field effects; these would lower all the energies by the same negative quantity, and change the chemical potential μ into $\overline{\mu}$ defined by (41), so that ξ_k would become $\xi_k = e_k - \overline{\mu}$. These effects slightly shift the curve plotting E_k to the left.

which part of the curve we analyze. On the left-hand side (decreasing function), the solid line curve nearly perfectly matches the dashed curve, meaning $|v_{\mathbf{k}}|^2$ is practically equal to 1 (see Figure 1). In that case, $1 - 2|v_{\mathbf{k}}|^2 \simeq -1$: a particle disappears in the course of the excitation, which is said to be of the "hole type". Its energy μ is the energy needed to push one particle towards the reservoir that fixes the chemical potential, diminished by its initial kinetic energy e_k, and to which we must add the mean field correction[15]; the excitation energy is therefore $\overline{\mu} - e_k$, a value that corresponds to the left-hand side of the curve. As for the right-hand side part (increasing function), the constant $|v_{\mathbf{k}}|^2$ is practically zero: the excitation adds a particle to the system, and is said to be of the "particle type". Its energy is equal to $e_k - \mu$, energy necessary to promote a particle from its energy μ to a state of kinetic energy e_k (with, as before, a mean field correction

[15]This correction changes the initial energy e_k into $e_k - V\langle N\rangle/2$, and hence $\mu - e_k$ into $\overline{\mu} - e_k$.

that changes μ into $\overline{\mu}$). Finally, for the central part of the curve, we have a "mixed" excitation, of both hole and particle type; it is the region of the spectrum where the solid line parts the most from the dashed line, and where the BCS mechanism, which creates the gap Δ plays an essential role.

From these four energy levels associated with each pair of states, quantum statistical mechanics allows obtaining a density operator describing the thermal equilibrium of the system at temperature T, as well as all the various thermodynamic functions. The corresponding development will not be exposed in this complement. We shall simply mention that it allows extending the validity of a certain number of results obtained at $T = 0$, by simply introducing a gap $\Delta(T)$ that depends on the temperature. At zero temperature, $\Delta(0)$ is still given by (46), but the gap decreases as T increases, and goes to zero for a certain critical temperature T_c. This cancelling of the gap corresponds to a phase transition: as a system of attractive fermions is cooled down, when it reaches a certain temperature the pair condensation phenomenon occurs, which leads to a number of physical consequences. For example, the system's specific heat first takes on values higher than in the absence of transition, then abruptly (exponentially) goes to zero as $T \to 0$.

Conclusion

In conclusion, the choice of a variational basis of paired states sheds new light on the behavior of an ensemble of attractive fermions. We focused on the case of weak attractions, corresponding to electrons in superconducting metals; in such a situation, $(\Delta/E_F) \ll 1$ and relation (75) shows that the pair range ξ_{pair} is very large compared to the distance between fermions. The one-particle distribution shown in Figure 1 is then very similar to the step function obtained for an ideal gas at zero temperature, the step being nevertheless rounded off over an energy band of width equal to Δ. In other words, the BCS pairing only slightly modifies the Fermi sphere of a perfect gas. Studying the properties of the optimal state, we were able to expose a number of important phenomena: existence of spatial dynamic correlations, which explain the increase of the average attraction (negative) energy between fermions of opposite spins; phase locking accounting for the cooperative aspect of the pairing (increase of the attractive energy overcoming the increase of kinetic energy, hence leading to a decrease of the system total energy); existence of a pair wave function describing fermions of opposite spins, and reminiscent of the Cooper pairs (Complement $\mathrm{D_{XVII}}$); appearance of a "gap" in the elementary excitation spectrum explaining the robustness of the system's ground state.

Another interesting limiting case concerns strong attractive interactions where the pair range becomes very small compared to the distance between particles. "Molecules" are then really formed with a binding energy $-e_L$ large (in absolute value) compared to the Fermi energy E_F. Saying that the size of the bound state is small compared to the distance between particles amounts to saying that its momentum distribution width is large compared to the Fermi wave vector k_F. This means that due to the attractive potential, the occupation of the individual states are spread over a large number of different momenta, which dilutes the effects of Pauli exclusion principle; these effects thus become negligible whereas they are essential in the BCS case. Instead of being positive, the chemical potential is now negative, close to $-e_L$. Relations (24) then show that the θ_k (and hence the populations of the individual states $\mathbf{k}$) always remain small, nevertheless

extending up to energies e_k of the order of e_L. In that special case, the pair wave function ϕ_{pair}, with its Fourier components $u_k v_k = \sin\theta_k \cos\theta_k e^{2i\zeta_k}$, practically coincides with the wave function having as Fourier components $(v_k/u_k) = \text{tg}\theta_k e^{2i\zeta_k}$, and which was initially used for building the paired state. As the molecules contain two strongly bound fermions, they behave as composite bosons (Complement A_{XVII}, § 3), which may undergo Bose-Einstein condensation. Paired states enable us to see the continuous passage from one limiting case (BCS situation with a weakly perturbed Fermi sphere) to the other (condensation of strongly bound "molecules"). A detailed discussion of this continuous passage and its physical consequences is given in §4-6 of reference [10] and in [11].

In this complement, we emphasized the physical interpretation of the results obtained in the detailed calculations we presented; this will give the reader the necessary base for studying the experimental aspects of superconductivity, which are not presented here. Among the many aspects of superconductivity that have not been studied in this complement, we can list: transport phenomena and the disappearance of the electrical resistance; behavior in the presence of a magnetic field (Meissner-Ochsenfeld effect); experimental study of the elementary excitation spectrum and gap measurements via different methods (tunnel effect, magnetic resonance); Josephson effect.

The interested reader can refer, for example, to the books of M. Tinkham [12], of R.D. Parks [13], or of A.J. Leggett already quoted [8]. The book of Combescot and Shiau [14] presents a good overview of the four main theoretical methods for studying superconductivity, the BCS variational method discussed in this complement being the first of this list.

Complement D_{XVII}

Cooper pairs

1 **Cooper model** **1927**
2 **State vector and Hamiltonian** **1927**
3 **Solution of the eigenvalue equation** **1929**
4 **Calculation of the binding energy for a simple case** **1929**

In this complement, we present the "Cooper model" which was a first step towards the complete BCS theory. It yields some results of that theory without having to deal with the difficulties inherent to an N-body problem. With this simplified model, we study the properties of two attractive fermions whose wave function, in the momentum representation, is excluded from the Fermi sphere. The model will be presented in § 1, where we show the existence of a bound state, occurring only because of the existence of that sphere. Furthermore, we shall see that the mathematical expression for the corresponding binding energy is reminiscent of the expression for the gap value Δ obtained in the BCS theory.

1. Cooper model

Among a large ensemble of identical fermions, we focus our attention on two of them, supposed to attract each other, in order to study their two-body wave function and energy levels. The presence of all the other fermions is simply accounted for by a Fermi sphere that, because of the Pauli exclusion principle, requires the components of that wave function to be zero inside that sphere. Such an approach is obviously not very rigorous: isolating two fermions among a large number of other indistinguishable fermions does not make much sense. Furthermore, it is hard to imagine why two of them would interact via an attractive potential, whereas all the others determining the Fermi sphere would be without interaction. However, the mathematical form of the results obtained with this model presents interesting similarities with the variational method where all the fermions are treated equally; it is thus useful to study this model.

2. State vector and Hamiltonian

Consider two attractive fermions in a singlet spin state $|S=0\rangle$:

$$|S=0\rangle = \frac{1}{\sqrt{2}}\left[|1:\uparrow\rangle\,|2:\downarrow\rangle - |1:\downarrow\rangle\,|2:\uparrow\rangle\right] \tag{1}$$

The relative motion of their position variables is described by the orbital ket $|\Psi_{\text{orb}}\rangle$, and their center of mass is described by a zero momentum ket $|\Phi_{\mathbf{K}=\mathbf{0}}\rangle$. Their state vector is:

$$|\Psi\rangle = |\Phi_{\mathbf{K}=\mathbf{0}}\rangle \otimes |\Psi_{\text{orb}}\rangle \otimes |S=0\rangle \tag{2}$$

The state $|\Psi_{\text{orb}}\rangle$ is characterized by a wave function $\Psi_{\text{orb}}(\mathbf{r})$:

$$\Psi_{\text{orb}}(\mathbf{r}) = \langle \mathbf{r} | \Psi_{\text{orb}} \rangle \tag{3}$$

where:

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{4}$$

is the difference between the positions of the two particles (relative position). As the singlet state is even with respect to the exchange of the two particles, their fermionic character requires the wave function $\Psi_{\text{orb}}(\mathbf{r})$ to be even with respect to particle exchange, i.e. with respect to a sign change of $\mathbf{r}$:

$$\Psi_{\text{orb}}(-\mathbf{r}) = \Psi_{\text{orb}}(\mathbf{r}) \tag{5}$$

We assume the operator describing the attractive interaction between the two particles to be independent of the spin. As in § B-2 of Chapter VII, we separate in the the two-particle Hamiltonian H the motion of the center of mass from the relative motion, and assume the center of mass is at rest. We are then left with a Hamiltonian H_{rel} that only acts on the space of the relative motion variables, and can be written:

$$H_{\text{rel}} = \frac{\hat{\mathbf{p}}^2}{m} + V(\hat{\mathbf{r}}) \tag{6}$$

where $\hat{\mathbf{r}} = \hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2$ is the operator associated with the relative position of the two particles; $\hat{\mathbf{p}}$ is the operator associated with the momentum of the relative motion, defined as a function of the momenta $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ of the two particles:

$$\hat{\mathbf{p}} = \frac{\hat{\mathbf{p}}_1 - \hat{\mathbf{p}}_2}{2} \tag{7}$$

As mentioned above, we assume the presence of an ensemble of non-interacting fermions, whose Fermi level is E_F. We must then solve the eigenvalue equation:

$$H_{\text{rel}} |\Psi_{\text{orb}}\rangle = (E + 2E_F) |\Psi_{\text{orb}}\rangle \tag{8}$$

when $|\Psi_{\text{orb}}\rangle$ does not have any component inside the Fermi sphere of radius k_F, related to the Fermi level E_F by:

$$E_F = \frac{\hbar^2 (k_F)^2}{2m} \tag{9}$$

In relation (8), E is the eigen-energy with respect to twice the Fermi level. It is indeed natural to take $2E_F$ as an energy reference; this is the minimal energy to be given to the two fermions under study, for their wave function $|\Psi_{\text{orb}}\rangle$ to have zero components inside the Fermi sphere, in the absence of interaction. With this convention for the energy origin, E simply reflects the effect of the attractive interaction.

3. Solution of the eigenvalue equation

We now expand $|\Psi_{\text{orb}}\rangle$ on the normalized eigenvectors $|\mathbf{k}\rangle$ (plane waves) of the momentum $\hat{\mathbf{p}}$:

$$|\Psi_{\text{orb}}\rangle = \sum_{\mathbf{k}} g_{\mathbf{k}} |\mathbf{k}\rangle \tag{10}$$

Projected onto $|\mathbf{k}\rangle$, the eigenvalue equation (8) reads:

$$\left(2e_k \; g_{\mathbf{k}} + \sum_{\mathbf{k}'} \langle \mathbf{k}| \widehat{V} |\mathbf{k}'\rangle \;\; g_{\mathbf{k}'} \right) = (E + 2E_F)\, g_{\mathbf{k}} \tag{11}$$

where we have set, as usual:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{12}$$

The absence of components of $|\Psi_{\text{orb}}\rangle$ inside the Fermi sphere leads to the relation:

$$g_{\mathbf{k}} = 0 \qquad \text{if} \quad |\mathbf{k}| \leq k_F \tag{13}$$

while (11) becomes:

$$[E + 2\,(E_F - e_k)]\, g_{\mathbf{k}} = \sum_{|\mathbf{k}'|>k_F} V_{\mathbf{k}\mathbf{k}'} \; g_{\mathbf{k}'} \tag{14}$$

The matrix elements of the interaction operator $\widehat{V}$ are noted $V_{\mathbf{k}\mathbf{k}'}$:

$$V_{\mathbf{k}\mathbf{k}'} = \langle \mathbf{k}| \widehat{V} |\mathbf{k}'\rangle \tag{15}$$

4. Calculation of the binding energy for a simple case

Let us further simplify the model and assume that the potential matrix elements $V_{\mathbf{k}\mathbf{k}'}$ are such that:

$$\begin{aligned} &V_{\mathbf{k}\mathbf{k}'} = -V \quad \text{if} \;\; |\mathbf{k}|\,, |\mathbf{k}'| \leq k_F + \Delta k \\ &V_{\mathbf{k}\mathbf{k}'} = 0 \qquad\qquad\qquad\quad \text{otherwise} \end{aligned} \tag{16}$$

where Δk defines a wave vector domain $\Delta k \ll k_F$; these matrix elements can therefore be factored. Note that the minus sign in front of the constant V was introduced to ensure that, for our present attractive potential, the constant V is positive. When $|\mathbf{k}| \leq k_F + \Delta k$, the summation on the right-hand side of (14) becomes a constant $-SV$ independent of $\mathbf{k}$, with:

$$S = \sum_{k_F<|\mathbf{k}'|\leq k_F+\Delta k} g_{\mathbf{k}'} \tag{17}$$

whereas, if $|\mathbf{k}| > k_F + \Delta k$, this summation is zero. The solution of this equation is simply:

$$\begin{aligned} &g_{\mathbf{k}} = \frac{-SV}{E + 2\,(E_F - e_k)} \qquad \text{if } k_F < |\mathbf{k}| \leq k_F + \Delta k \\ &g_{\mathbf{k}} = 0 \qquad\quad \text{otherwise} \end{aligned} \tag{18}$$

We must add the self-consistent condition we get from inserting this solution in the definition (17) of S:

$$S = \sum_{k_F<|\mathbf{k}'|\leq k_F+\Delta k} \frac{-SV}{E + 2\left(E_F - e_{k'}\right)} \tag{19}$$

that is, changing the sign of the denominator:

$$\frac{1}{V} = \sum_{k_F<|\mathbf{k}'|\leq k_F+\Delta k} \frac{1}{2\left(e_{k'} - E_F\right) - E} \tag{20}$$

This condition is also an implicit equation for obtaining the energy E. Assuming the system is enclosed in a cubic box with a very large edge length L, the discrete summation can be replaced by an integral, and we get:

$$\frac{1}{V} = \frac{L^3}{2\pi^2} \int_{k_F}^{k_F+\Delta k} \frac{k^2 dk}{2\left(e_k - E_F\right) - E} \tag{21}$$

We now choose, as the integral variable, the variable x:

$$x = e_k - E_F \tag{22}$$

As $\mathrm{d}e_k = \hbar^2 k \mathrm{d}k/m$, this integral now includes a density of states $D\left(e_k\right)$:

$$D\left(e_k\right) = \frac{L^3}{2\pi^2} k^2 \frac{\mathrm{d}k}{de_k} = \frac{L^3}{2\pi^2} \frac{m}{\hbar^2} k \tag{23}$$

or:

$$D\left(e_k\right) = \frac{L^3}{2\pi^2} \frac{m}{\hbar^3} \sqrt{2me_k} \tag{24}$$

The implicit equation for E then becomes:

$$\frac{1}{V} = \int_0^{\Delta E} \mathrm{d}x \frac{D\left(E_F + x\right)}{2x - E} \tag{25}$$

where the upper bound ΔE is defined by:

$$E_F + \Delta E = \frac{\hbar^2}{2m} \left(k_F + \Delta k\right)^2 \tag{26}$$

As we assumed $\Delta k \ll k_F$, we can replace[1] in (25) the density of states $D\left(E_F + x\right)$ by $D\left(E_F\right)$, no longer dependent on the variable x, and that we simply note D_F:

$$D_F = D\left(E_F\right) \tag{27}$$

We can then perform the integration, which yields:

$$\begin{aligned} \frac{1}{V} &= \frac{D_F}{2} \left\{\ln\left(2x - E\right)\right\}_0^{\Delta E} = \frac{D_F}{2} \ln\left[\frac{2\Delta E - E}{-E}\right] \\ &= -\frac{D_F}{2} \ln\left[\frac{-E}{2\Delta E - E}\right] \end{aligned} \tag{28}$$

[1] Replacing k by k_F in (23), one can easily compute an order of magnitude for the density of states at the Fermi level. We find $D\left(E_F\right) \sim \langle N \rangle / E_F$, hence a value proportional to the average particle number.

We then have:

$$\frac{-E}{2\Delta E - E} = e^{-2/VD_F} \tag{29}$$

The solution of this equation for E is:

$$E = -2\Delta E \ \frac{e^{-2/VD_F}}{1 - e^{-2/VD_F}} \tag{30}$$

which, when $VD(E_F) \ll 1$, can be simplified to:

$$E = -2\Delta E \ \exp\left[-2/VD_F\right] \tag{31}$$

We obtain a negative energy (with respect to $2E_F$), as expected for a bound state (the wave function can be normalized). As we cannot make a series expansion of the function $\exp(-1/x)$ in the vicinity of $x = 0$, this energy cannot be expressed as a power series of the interaction potential V, since all its derivative are zero at $x = 0$; consequently, this energy cannot be obtained by an ordinary perturbation calculation.

Note also that the energy E goes to zero (through negative values) if the density of states $D(E_F)$ goes to zero, i.e. if E_F goes to zero: the existence of the bound state is therefore linked to the presence of the Fermi sphere, whose role is to introduce a non-zero density of states. If the Fermi sphere disappears, so does the bound state.

We find results similar to those found in Complement C_{XVII} using the BCS theory, and in particular to the expression (46) of that complement, which yields the gap Δ. To obtain that expression, we had to introduce an upper bound $\hbar\omega_D$ for the variation (in absolute value) of the energies around the Fermi energy; this upper bound plays a role comparable to that played by the energy ΔE we just introduced in (25). We simply have to assume that $\hbar\omega_D = \Delta E$ for the two results to become quite similar as they only differ by a factor 2 in the exponential (the sign difference simply comes from the fact that the gap Δ was defined as a positive quantity, whereas a binding energy is negative). The interest of the Cooper model is to clearly highlight the essential role played by the density of states D_F (in the vicinity of the Fermi level) in the creation of the gap Δ in the BCS theory.

Complement E_{XVII}

Condensed repulsive bosons

1 Variational state, energy 1935
1-a Variational ket 1935
1-b Total energy 1936
1-c $N_0 \gg N_e$ approximation 1936
2 Optimization 1937
2-a Stationarity conditions 1938
2-b Solution of the equations 1939
3 Properties of the ground state 1940
3-a Particle number, quantum depletion 1940
3-b Energy 1942
3-c Phase locking; comparison with the BCS mechanism 1944
3-d Correlation functions 1946
4 Bogolubov operator method 1950
4-a Variational space, restriction on the Hamiltonian 1951
4-b Bogolubov Hamiltonian 1952
4-c Constructing a basis of excited states, quasi-particles 1954

In this complement we study the properties of an ensemble of repulsively interacting bosons[1], undergoing Bose-Einstein condensation. We know that, for an ideal gas, a system of bosons in its ground state is totally condensed: a single individual quantum state, corresponding to the lowest energy, is occupied by all the particles. In the presence of short-range interactions, and for a sufficiently diluted system, one expects its properties to remain close to that of an ideal gas, and in particular that a large fraction of the particles still occupy the same individual quantum state. We consider this to be the case for the system under study, and that the population of one individual quantum state is much larger than all the others. We shall assume that each of the states obeys the periodic boundary conditions in a box of side length L (Complement C_{XIV}), and that the state with the large population[2] is the state $\mathbf{k} = \mathbf{0}$ (whose momentum $\mathbf{p} = \hbar\mathbf{k}$ and kinetic energy are zero). Consequently, if N_0 is the average value of the population of this zero momentum level, we assume that:

$$N_0 \gg n_{\mathbf{k}} \qquad (\text{for any } \mathbf{k} \neq \mathbf{0}) \tag{1}$$

[1] We will not consider the case of attractive interactions, as they lead to an unstable physical state – see § 4-b of Complement H_{XV}.

[2] This hypothesis simplifies the writing of the equations, but is not essential; in the case where it is a state of non-zero momentum $\mathbf{k}_0$ that is highly populated, one can go to the reference frame where this momentum is zero. This amounts to adding, in the initial frame, $\mathbf{k}_0$ to all the wave vectors appearing in the equations.

where $n_{\mathbf{k}} = \langle \widehat{n}_{\mathbf{k}} \rangle$ is the average number of particles occupying an individual state $\mathbf{k} \neq \mathbf{0}$. The total particle number N is:

$$N = N_0 + \sum_{\mathbf{k} \neq \mathbf{0}} n_{\mathbf{k}} \tag{2}$$

We have already used, in Complement C_{XV}, a first approximation to study the ground state of a condensed boson system: we described the state of the N-particle system as the product of N identical individual state vectors. This led to the Gross-Pitaevskii equation. This approach implies that only one single individual state $\mathbf{k} = \mathbf{0}$ is occupied ($N_0 = N$ and $n_{\mathbf{k}} = 0$ for any $\mathbf{k}$), as for an ideal gas. This obviously cannot be exact: it is clear that the interactions introduce dynamic correlations between the particles, which cannot be accounted for by a state vector that is a simple product of individual kets (hence without correlations). Actually, the effect of the interaction potential on the ground state is to transfer at least a fraction of the particles[3] from the state $\mathbf{k} = \mathbf{0}$ to the states $\mathbf{k} \neq \mathbf{0}$; a model involving only one individual state is necessarily limited to the case where the potential effect is very weak, and hence $N_0 \simeq N$.

In Complement E_{XV}, we introduced another approximation, based on the Hartree-Fock method; it is more general than the previous one as it allows taking into account a non-zero temperature. However, it still implies that each particle moves in the mean field created by all the others, ignoring the dynamic correlations; its description of the ground state is no better than the one derived from the Gross-Pitaevskii equation. Furthermore, this latter method proved to be problematic for a boson system undergoing Bose-Einstein condensation: we noted in § 3-b-β of Complement G_{XV} that, for a system of condensed bosons, the Hartree-Fock approximation predicts, at the grand canonical equilibrium, very large fluctuations of the number of condensed particles. In the real world, these fluctuations are strongly limited by the repulsion between the particles, which clearly indicates that the predictions of the Hartree-Fock approximation concerning fluctuations are non-physical.

In the present complement, we shall try to address these two problems: on one hand, we shall take into account the dynamic correlations introduced by the interactions in the physical system; on the other, we shall not let the number of condensed particles fluctuate arbitrarily. We will use a variational method, choosing a variational state that takes into account the binary correlations between the particles, but does not introduce unrealistic fluctuations of the particle number. This variational state will be built with the help of a paired state, enabling us to directly use the results of Chapter XVII. We will add an extra component, to account for the Bose-Einstein condensation in the individual $\mathbf{k} = \mathbf{0}$ state. Obviously, this is still not an exact calculation, as it involves a variational approximation, but it allows describing a physical situation more complex than the simple Gross-Pitaevskii approximation. This approach also highlights the many analogies, but also the differences, between the pairing of condensed bosons and the pairing of fermions.

In a general way, this complement illustrates how variational methods allow changing the correlations between particle pairs. When dealing with binary interactions, as in a standard Hamiltonian, these correlations determine the average value of the potential energy (Chapter XV, § C-5-b-α). The higher order correlations (ternary, etc.) may be present and play a role in the system; but they are not directly involved in the energy.

[3] This phenomenon is traditionally called "quantum depletion" and will be discussed in more detail in § 3-a.

This is why using the paired states to optimize only the binary correlations can lead to fairly good results.

We introduce in § 1 the paired variational state depending on a certain number of parameters, and compute the corresponding average energy. In § 2, we shall search for the optimal values of these parameters that minimize this energy, using an approximation where $N_0 \simeq N$ so that we can neglect the interactions between the particles in the $\mathbf{k} \neq \mathbf{0}$ individual states. In § 3 we study the physical properties of the state thus obtained, such as the number of particles that are not in the $\mathbf{k} = \mathbf{0}$ state, the energy, and the correlation functions. We shall then develop, in § 4, a different point of view, the Bogolubov operator method. We shall choose a larger variational space, and use the results of § E in Chapter XVII to get the Bogolubov Hamiltonian, which can be directly diagonalized. This will confirm a certain number of previously obtained results. The reader only interested in the Bogolubov operator method can go directly to this paragraph, which is fairly self-contained. The conclusion of this complement will sum up the results obtained and the limits of this approximation method.

1. Variational state, energy

We are now going to directly apply the results of Complement $\mathrm{B}_{\mathrm{XVII}}$, for the choice of the variational ket as well as for the computation of its average energy.

1-a. Variational ket

The (normalized) variational ket is of the form:

$$|\Phi_B\rangle = |\varphi_{\mathbf{0}}\rangle \otimes |\Psi_{\text{paired}}\rangle = |\varphi_{\mathbf{0}}\rangle \otimes \prod_{\otimes \mathbf{k} \in D} |\overline{\varphi}_{\mathbf{k}}\rangle \tag{3}$$

where the subscript B refers to the name Bogolubov. In this expression, $|\Psi_{\text{paired}}\rangle$ is the paired state for spinless particles written in (B-8) of Chapter XVII, which is a tensor product of the normalized states (C-13):

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{1}{\cosh\theta_{\mathbf{k}}} \exp\left[-x_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right] |0\rangle \tag{4}$$

with:

$$x_{\mathbf{k}} = \tanh\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \qquad (\ \theta_{\mathbf{k}} \geq 0) \tag{5}$$

The domain D of the tensor product in (3) is half the $\mathbf{k}$-space, which prevents (as we saw in Chapter XVII, § B-2-a) the double appearance of each state $|\overline{\varphi}_{\mathbf{k}}\rangle = |\overline{\varphi}_{-\mathbf{k}}\rangle$; the origin $\mathbf{k} = \mathbf{0}$ is excluded from D.

As for $|\varphi_{\mathbf{0}}\rangle$, it is the coherent state already used in Complement $\mathrm{B}_{\mathrm{XVII}}$, relation (44):

$$|\varphi_{\mathbf{0}}\rangle = e^{-N_0/2}\ e^{\alpha_0 a^{\dagger}_0} |n_0 = 0\rangle \tag{6}$$

This state depends on a complex parameter α_0, characterized by its modulus $\sqrt{N_0}$ and its phase ζ_0:

$$\alpha_0 = \sqrt{N_0} e^{i\zeta_0} \tag{7}$$

It is the normalized eigenvector of the operator a_0 with eigenvalue α_0:

$$a_0 \left|\varphi_{\mathbf{0}}\right\rangle = \alpha_0 \left|\varphi_{\mathbf{0}}\right\rangle \tag{8}$$

The average particle number in the state $\mathbf{k} = \mathbf{0}$ is then:

$$\left\langle\varphi_0\right| a_0^\dagger a_0 \left|\varphi_0\right\rangle = \alpha_0^* \alpha_0 \left\langle\varphi_0 \right|\left.\varphi_0\right\rangle = N_0 \tag{9}$$

The width of the corresponding distribution is $\sqrt{N_0}$ (Complement G_V), hence negligible compared to N_0 (this number is supposed to be large).

The variational variables contained in the trial ket (3) are thus the set of $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$, as well as N_0 and ζ_0.

1-b. Total energy

Expression (61) of Complement B_{XVII} yields the total energy in the form:

$$\begin{aligned}\left\langle \widehat{H} \right\rangle_B &= \sum_{\mathbf{k}\neq\mathbf{0}} e_k \, \sinh^2\theta_{\mathbf{k}} + \frac{V_0}{2}\left(N_0 + N_e\right)^2 \\ &+ N_0 \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\left(\zeta_0 - \zeta_{\mathbf{k}}\right)\right] \\ &+ \frac{1}{2} \sum_{\mathbf{k},\mathbf{k}'\neq\mathbf{0}} V_{\mathbf{k}-\mathbf{k}'} \left[\sinh^2\theta_{\mathbf{k}} \, \sinh^2\theta_{\mathbf{k}'} + \sinh\theta_{\mathbf{k}} \, \sinh\theta_{\mathbf{k}'} \, \cosh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}'} \cos 2\left(\zeta_{\mathbf{k}} - \zeta_{\mathbf{k}'}\right)\right]\end{aligned} \tag{10}$$

where the matrix elements $V_{\mathbf{k}}$ of the particle interaction potential are defined, as in Chapter XVII, by:

$$V_{\mathbf{k}} = \frac{1}{L^3} \int \mathrm{d}^3 r \; e^{-i\mathbf{k}\cdot\mathbf{r}} \; W_2(\mathbf{r}) \tag{11}$$

The term on the second line of (10) corresponds to the momentum exchanges between the $\mathbf{k} \neq \mathbf{0}$ particles and the $\mathbf{k} = \mathbf{0}$ condensate, as well as the pair annihilation-creation processes originating from the condensate. The terms in the last line, with a double summation over $\mathbf{k}$ and $\mathbf{k}'$, correspond to interaction effects between $\mathbf{k} \neq \mathbf{0}$ particles.

1-c. $N_0 \gg N_e$ approximation

As already pointed out in the introduction, for an ideal gas in its ground state, only one individual state is occupied, corresponding to the lowest energy; in that case, the average total particle number is equal to N_0, and all the populations $n_{\mathbf{k}}$ of the other $\mathbf{k}$ states are zero. We are going to assume that the system we study is a dilute gas where the interaction effects are limited so that N_0 remains very large compared to the sum of all the populations $n_{\mathbf{k}}$:

$$N_0 \gg N_e = \sum_{\mathbf{k}\neq\mathbf{0}} n_k \tag{12}$$

This hypothesis is more constraining than the one initially proposed in (1), since now the population N_0 must largely exceed the sum of all the other populations. Nevertheless it allows a simplification of the following computations while highlighting a certain number of general physical ideas.

Under these conditions, the interactions between particles in the $\mathbf{k} \neq \mathbf{0}$ states and particles in the $\mathbf{k} = \mathbf{0}$ condensate are dominant compared to the interactions between particles both in the $\mathbf{k} \neq \mathbf{0}$ states. The interaction term on the second line of (10), proportional to N_0, is therefore much larger than the term on the last line, which does not contain N_0. This is why we use the approximate average value:

$$\left\langle \widehat{H} \right\rangle_B \simeq \sum_{\mathbf{k}\neq\mathbf{0}} e_k \, \sinh^2\theta_{\mathbf{k}} + \frac{V_0}{2} \left(N_0 + N_e\right)^2 + N_0 \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\left(\zeta_0 - \zeta_{\mathbf{k}}\right)\right] \tag{13}$$

We have yet to determine the optimal values of the variables appearing in (13) by minimizing this energy average value with respect to each of them.

2. Optimization

The variational state $|\Phi_B\rangle$ depends on the variables N_0 and ζ_0 associated with the individual state $\mathbf{k} = \mathbf{0}$ (condensate), as well as on the angles $\theta_{\mathbf{k}}$ and the phases $\zeta_{\mathbf{k}}$ associated with all the other $\mathbf{k} \neq \mathbf{0}$ states. On the other hand, N_e is not a variational variable, but a function of the previous variables determined by relation (53) of Complement $\mathrm{B_{XVII}}$:

$$N_e = \sum_{\mathbf{k}\neq\mathbf{0}} \left\langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \right\rangle = \sum_{\mathbf{k}\neq\mathbf{0}} \sinh^2\theta_{\mathbf{k}} \tag{14}$$

As in Complement $\mathrm{B_{XVII}}$, we introduce a Lagrange multiplier μ (chemical potential, see Appendix VI) to fix the average total particle number; we thus impose the stationarity of the difference of two average values:

$$A = \left\langle \widehat{H} \right\rangle_B - \mu \left\langle \widehat{N} \right\rangle_B \tag{15}$$

where $\left\langle \widehat{N} \right\rangle_B$ is the average total particle number in the variational state:

$$\left\langle \widehat{N} \right\rangle_B = \left\langle a_0^\dagger a_0 \right\rangle + \sum_{\mathbf{k}\neq\mathbf{0}} \left\langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \right\rangle = N_0 + N_e \tag{16}$$

The function to be minimized is therefore:

$$A = \frac{V_0}{2} \left(N_0 + N_e\right)^2 - \mu N_0 + \sum_{\mathbf{k}\neq\mathbf{0}} \left(e_k - \mu\right) \, \sinh^2\theta_{\mathbf{k}} + N_0 \sum_{\mathbf{k}\neq\mathbf{0}} F_{\mathbf{k}} \tag{17}$$

with:

$$F_{\mathbf{k}} = V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\left(\zeta_0 - \zeta_{\mathbf{k}}\right)\right] \tag{18}$$

2-a. Stationarity conditions

The function A must be made stationary with respect to all the variables. We shall start with the phases, then the parameters $\theta_{\mathbf{k}}$, and finally N_0.

α. Stationarity with respect to the phases: phase locking

The phases only intervene in the $F_{\mathbf{k}}$, as phase differences $\zeta_0 - \zeta_{\mathbf{k}}$. Since we have a repulsive potential, we assume $V_{\mathbf{k}}$ is positive; furthermore, as the variable $\theta_{\mathbf{k}}$ is always positive according to its definition in Chapter XVII, the product $\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}$ is also always positive. Expression (18) shows that, whatever the value of $\theta_{\mathbf{k}}$, the minimization of the function A with respect to the phases ζ_0 and $\zeta_{\mathbf{k}}$ requires the cosine to be equal to 1, that is:

$$\zeta_{\mathbf{k}} = \zeta_0 \qquad \text{for any } \mathbf{k} \tag{19}$$

Consequently, the phases used to build the paired states must all be equal to the phase defining the coherent state associated with $\mathbf{k} = \mathbf{0}$. We call this equality the "*phase locking condition*".

β. Stationarity with respect to the $\theta_{\mathbf{k}}$

The stationarity of A with respect to each parameter $\theta_{\mathbf{k}}$ implies that, for any $\mathbf{k}$:

$$0 = V_0\,(N_0 + N_e)\,\frac{\partial N_e}{\partial \theta_{\mathbf{k}}} + 2\,(e_k - \mu)\ \sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}} + N_0\frac{\partial F_{\mathbf{k}}}{\partial \theta_{\mathbf{k}}} \tag{20}$$

where the derivative of $F_{\mathbf{k}}$ is taken at the phase values that satisfy relation (19). Grouping on a first line the terms in $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$ (including those coming from the derivative of $\sinh^2\theta_{\mathbf{k}}$), and on a second, those coming from the derivative of $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$, we get:

$$\begin{aligned} 0 \;=\; & \left[V_0\,(N_0 + N_e) + e_k - \mu + N_0 V_{\mathbf{k}}\right] 2\,\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}} \\ & -N_0 V_{\mathbf{k}}\left(\cosh^2\theta_{\mathbf{k}} + \sinh^2\theta_{\mathbf{k}}\right) \end{aligned} \tag{21}$$

Relation (21) then becomes:

$$0 = \left[e_k - \mu + V_0\,(N_0 + N_e) + N_0 V_{\mathbf{k}}\right]\ \sinh 2\theta_{\mathbf{k}}\ - N_0 V_{\mathbf{k}}\,\cosh 2\theta_{\mathbf{k}} \tag{22}$$

that is:

$$\tanh 2\theta_{\mathbf{k}} = \frac{N_0 V_{\mathbf{k}}}{e_k - \mu + V_0\,(N_0 + N_e) + N_0 V_{\mathbf{k}}} \tag{23}$$

γ. Stationarity with respect to N_0

We now write the stationarity of A with respect to N_0. Taking into account relations (17) and (18), as well as the phase locking condition (19), we get:

$$0 = V_0\,(N_0 + N_e) - \mu + \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}}\left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\right] \tag{24}$$

This result shows that the chemical potential is equal to:

$$\mu = V_0 (N_0 + N_e) + \delta\mu \tag{25}$$

which is the sum of a mean field term $V_0 (N_0 + N_e)$ created by all the particles, and another term $\delta\mu$:

$$\delta\mu = \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}}\right] \tag{26}$$

This last term is also the sum of two terms of different signs: a positive contribution coming from momentum transfer processes, leading to an increase of the repulsion energy due to the boson bunching effect; a negative contribution due to the creation or annihilation of pairs from the condensate $\mathbf{k} = \mathbf{0}$ (Figure 7 of Complement B_{XVII}), and expressing the reduction of that energy, due to the dynamic correlations induced by the interactions.

Relation (23) then becomes:

$$\tanh 2\theta_{\mathbf{k}} = \frac{N_0 V_{\mathbf{k}}}{\widetilde{e}_k + N_0 V_{\mathbf{k}}} \tag{27}$$

with:

$$\widetilde{e}_k = e_k - \delta\mu \tag{28}$$

2-b. Solution of the equations

The ground state we are looking for depends on two parameters that are externally fixed, the volume L^3 of the physical system, and the chemical potential μ that controls the total number of particles. We must determine the variables $\theta_{\mathbf{k}}$ from (27), as well as N_0 from (24). This last relation includes N_e, which is not an independent variable since it is determined by (14). We have a set of non-linear equations whose solution is not obvious, a priori: relation (27) determines the $\theta_{\mathbf{k}} (N_0)$, and hence $N_e (N_0)$, as a function of N_0. But N_0 itself is determined as a function of μ and the variables $\theta_{\mathbf{k}}$ (directly, and indirectly through N_e) by the stationarity condition (24). Inserting the $\theta_{\mathbf{k}} (N_0)$ in this relation, we get an implicit equation for N_0, reminiscent of the implicit equation for the gap Δ in the BCS theory (Complement C_{VII}, § 1-c-β).

A first approach for solving this implicit equation is to proceed by successive iterations, as in the Hartree-Fock method. We start from an approximate, reasonable value of N_0, such as the value μ/V_0 obtained by assuming that N_e and $\delta\mu$ are both zero. Using (27), we then get a first approximation for the $\theta_{\mathbf{k}}$ and for N_e, that can be inserted in (24) to get a new value for N_0. Iterating the process, one can expect, as for the Hartree-Fock non-linear equations, a convergence after a certain number of cycles.

Another approach is to not arbitrarily fix the chemical potential, but rather deduce it from the computation. We then start from an arbitrary N_0 value, yielding the values of the angles $\theta_{\mathbf{k}}$, then the value of N_e using (14); this fixes the total particle number $N = N_0 + N_e$, and the relations (25)-(26) yield the chemical potential. We shall use this simpler approach in what follows.

3. Properties of the ground state

We start by computing the total particle number. To highlight the general ideas while dealing with equations as simple as possible, we shall use a model where the potential matrix elements $V_{\mathbf{k}}$ are all equal to the same constant V_0, or else equal to zero:

$$\begin{aligned} V_{\mathbf{k}} &= V_0 \quad \text{if } |\mathbf{k}| \leq k_c \\ V_{\mathbf{k}} &= 0 \quad \text{if } |\mathbf{k}| > k_c \end{aligned} \tag{29}$$

where the "cutoff value" k_c characterizes the potential range b ($k_c \sim 1/b$). To further simplify, we shall consider, in each calculation, the case where $\delta\mu = 0$, i.e. where $\widetilde{e}_k$ is the same as e_k.

3-a. Particle number, quantum depletion

We use relation (14) to compute N_e, the average number of particles in the individual states $\mathbf{k} \neq \mathbf{0}$. To get $\sinh^2\theta_{\mathbf{k}}$, let us first compute $\cosh 2\theta_{\mathbf{k}}$ using:

$$\begin{aligned} \cosh 2\theta_{\mathbf{k}} &= \sqrt{\frac{\cosh^2 2\theta_{\mathbf{k}}}{\cosh^2 2\theta_{\mathbf{k}} - \sinh^2 2\theta_{\mathbf{k}}}} = \sqrt{\frac{1}{1 - \tanh^2 2\theta_{\mathbf{k}}}} \\ &= \frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} \end{aligned} \tag{30}$$

Since we have:

$$2\sinh^2\theta_{\mathbf{k}} = \cosh^2\theta_{\mathbf{k}} + \sinh^2\theta_{\mathbf{k}} - 1 = \cosh 2\theta_{\mathbf{k}} - 1 \tag{31}$$

we can write:

$$\sinh^2\theta_{\mathbf{k}} = \frac{1}{2}\left[\frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{32}$$

Inserting this relation in (14) we get:

$$N_e = \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{33}$$

Let us see what becomes of this expression in the simple model where the $\widetilde{e}_k$ are equal to the e_k ($\delta\mu$ is supposed to be negligible). Replacing the summation in (33) by an integral, we get:

$$N_e = \frac{L^3}{16\pi^3}\int \mathrm{d}^3k \ \left[\frac{\frac{\hbar^2k^2}{2m} + N_0 V_{\mathbf{k}}}{\sqrt{\frac{\hbar^2k^2}{2m}\left(\frac{\hbar^2k^2}{2m} + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{34}$$

Using the matrix elements of the simplified potential, relation (29), the function to be integrated only depends on the modulus k of $\mathbf{k}$, and goes to zero if $k > k_c$; this means

that, using spherical coordinates, the integral over dk only goes from 0 to k_c. We define the integral variable $\mathbf{s}$:

$$\mathbf{s} = \sqrt{\frac{\hbar^2}{2mN_0V_0}}\ \mathbf{k} = \xi\ \mathbf{k} \tag{35}$$

where ξ is the "healing length" introduced[4] in § 4-b of Complement $\mathrm{C_{XV}}$:

$$\xi = \sqrt{\frac{\hbar^2}{2mN_0V_0}} \tag{36}$$

Noting S_c the upper bound of $\mathbf{s}$ coming from the upper bound k_c of k:

$$S_c = \xi\ k_c \tag{37}$$

we can write:

$$N_e = \frac{L^3}{4\pi^2}\left(\frac{2mN_0V_0}{\hbar^2}\right)^{3/2}\int_0^{S_c} s^2\ \mathrm{d}s\ \left[\frac{s^2+1}{\sqrt{s^2\,(s^2+2)}} - 1\right] \tag{38}$$

The integral in (38) is still convergent if S_c goes to infinity since, when $s \to \infty$, one can make a limited expansion in powers of the infinitely small $x = 1/s^2$ and write:

$$\frac{s^2+1}{\sqrt{s^2\,(s^2+2)}} = \frac{1+1/s^2}{\sqrt{1+2/s^2}} = 1 + \frac{1}{2s^4} + .. \tag{39}$$

The integral also converges at the origin (the function to be integrated diverges as s^{-1}, but the differential element is $s^2\mathrm{d}s$, which eliminates the divergence). This integral can be readily calculated, and for an infinite value of S_c (very short-range potential) is equal to $\sqrt{2}/3$. We then get:

$$N_e = \frac{L^3}{3\pi^2}\left(\frac{mN_0V_0}{\hbar^2}\right)^{3/2} \tag{40}$$

We find that N_e is proportional to the volume and to the product N_0V_0 to the power 3/2. When the interaction potential V_0 is zero, we confirm that all the particles are in the individual state $\mathbf{k} = \mathbf{0}$. As V_0 starts increasing, the "non-condensed fraction" $N_e/\,(N_0 + N_e)$ varies, at the beginning, proportionally to the power 3/2 of V_0.

We have found that the effect of the interaction potential is to transfer a certain number of particles from the individual state $\mathbf{k} = \mathbf{0}$ towards the $\mathbf{k} \neq \mathbf{0}$ states. This effect is often called "quantum depletion". It has nothing to do with a thermal excitation effect that would bring some particles from their ground state towards excited states, as a result of the coupling with a thermal reservoir at a non-zero temperature. The calculations we are performing in this complement concern the ground state, and we assume the temperature is rigorously zero.

[4] In Complement $\mathrm{C_{XV}}$, we defined in (61) a constant ξ as a function of the parameter g associated with an interaction potential in $g\delta(\mathbf{r})$. Such a potential corresponds to $V_0 = g/\mathcal{V}$ (where $\mathcal{V}$ is the volume), and relation (36) is indeed equivalent to $\xi = \sqrt{\hbar^2/2mgn_0}$.

Comment:

Note however that the result (40) was established with the hypothesis $N_e \ll N_0$. It is therefore only valid if:

$$V_0 \ll \frac{\hbar^2}{mL^2} (N_0)^{-1/3} \tag{41}$$

If b is the range of the potential, and U its order of magnitude, relation (11) shows that the order of magnitude of the matrix elements V_0 is Ub^3/L^3; the previous condition is then written:

$$U \ll \frac{\hbar^2}{mb^2} \left(\frac{L^3}{b^3 N_0} \right)^{1/3} \tag{42}$$

The result is thus valid if U remains small compared to the kinetic energy of a particle localized within the potential range b, multiplied by the ratio between the average particle distance in the state $\mathbf{k} = \mathbf{0}$ and b. This requires the potential range b to be sufficiently small.

3-b. Energy

We now compute the energy, taking successively all the different terms of (13) into account.

α. Kinetic energy

The first contribution comes from the kinetic energy, which according to (32) is written:

$$E_c = \sum_{\mathbf{k}\neq\mathbf{0}} e_k \sinh^2\theta_{\mathbf{k}} = \frac{1}{2} \sum_{\mathbf{k}\neq\mathbf{0}} e_k \left[\frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1 \right] \tag{43}$$

This term reminds us of the one we encountered in the computation of N_e in (33), but the presence of the factor e_k in the summation changes its properties. In the simplified potential model where the $V_{\mathbf{k}}$ are given by (29), and if we furthermore assume that $\delta\mu = 0$, the change of variable (35) leads to:

$$E_c = \frac{L^3}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} [N_0 V_0]^{5/2} \int_0^{S_c} \mathrm{d}s \; s^4 \left[\frac{s^2+1}{\sqrt{s^2\left(s^2+2\right)}} - 1 \right] \tag{44}$$

According to (39), when $s \to \infty$, the function to be integrated behaves as:

$$s^4 \left[\frac{s^2+1}{\sqrt{s^2\left(s^2+2\right)}} - 1 \right] = s^4 \left[1 + \frac{1}{2s^4} + .. - 1 \right] = \frac{1}{2} + 0\left(\frac{1}{s^2}\right) \tag{45}$$

and tends towards a constant. Consequently its integral over $\mathrm{d}s$ is divergent if S_c is infinite; if S_c is large but finite, the integral value depends linearly on the choice of S_c.

β. Interaction with the condensate

In relation (13), the mean field term in the energy $V_0 \left(N_0 + N_e\right)^2 /2$ is known since N_e has been obtained previously – see relation (38). We do not need to compute specifically its contribution to the average energy.

The second term in the potential energy is proportional to N_0, and corresponds to the interactions between atoms outside the condensate ($\mathbf{k} \neq \mathbf{0}$ individual states) and inside the condensate (population of the $\mathbf{k} = \mathbf{0}$ state); this term contains the sum of the $F_{\mathbf{k}}$ defined in (18). In order to evaluate that term, we need to compute the product $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$:

$$\begin{aligned} \sinh\theta_{\mathbf{k}}\, \cosh\theta_{\mathbf{k}} &= \frac{1}{2}\sinh 2\theta_{\mathbf{k}} = \frac{1}{2}\sqrt{\frac{\sinh^2 2\theta_{\mathbf{k}}}{\cosh^2 2\theta_{\mathbf{k}} - \sinh^2 2\theta_{\mathbf{k}}}} \\ &= \frac{1}{2}\sqrt{\frac{\tanh^2 2\theta_{\mathbf{k}}}{1 - \tanh^2 2\theta_{\mathbf{k}}}} \end{aligned} \tag{46}$$

or else, according to (27):

$$\sinh\theta_{\mathbf{k}}\, \cosh\theta_{\mathbf{k}} = \frac{N_0 V_{\mathbf{k}}}{2\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} \tag{47}$$

Taking into account the phase locking condition (19) for the optimal variational state, the $F_{\mathbf{k}}$ defined in (18) are equal to:

$$\begin{aligned} F_{\mathbf{k}} &= V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\, \cosh\theta_{\mathbf{k}}\right] \\ &= \frac{V_{\mathbf{k}}}{2}\left[\frac{\widetilde{e}_k}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \end{aligned} \tag{48}$$

We obtain the contribution E_{N_0} to the energy:

$$E_{N_0} = N_0 \sum_{\mathbf{k}\neq\mathbf{0}} F_{\mathbf{k}} = \frac{N_0}{2} \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\frac{\widetilde{e}_k}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{49}$$

For the simple model (29) already used before, and if $\delta\mu = 0$, this results becomes:

$$E_{N_0} = \frac{L^3}{4\pi^2} N_0 V_0 \left(\frac{2mN_0 V_0}{\hbar^2}\right)^{3/2} \int_0^{S_c} s^2 \, \mathrm{d}s \, \left[\sqrt{\frac{s^2}{s^2+2}} - 1\right] \tag{50}$$

The function to be integrated behaves, at infinity, as:

$$s^2 \left[\frac{s}{\sqrt{s^2+2}} - 1\right] \sim s^2 \left[1 - \frac{1}{s^2} + .. - 1\right] = -1 + 0\left(\frac{1}{s^2}\right) \tag{51}$$

which means the integral is not convergent when $S_c \to \infty$, but depends linearly on S_c.

γ. Ground state total energy

With the same approximation $\delta\mu = 0$, the sum of (43) and (49) yields:

$$E_c + E_{N_0} = \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\sqrt{e_k\left(e_k + 2N_0V_{\mathbf{k}}\right)} - e_k - N_0V_{\mathbf{k}}\right] \tag{52}$$

or else, using again the simplified model (29) for the interaction potential and adding (44) and (50):

$$E_c + E_{N_0} = \frac{1}{4\pi^2}\left(\frac{2mL^2}{\hbar^2}\right)^{3/2}\left(N_0V_0\right)^{5/2}\int_0^{S_c} ds\ \left[s^3\sqrt{(s^2+2)} - s^4 - s^2\right] \tag{53}$$

Relations (45) and (51) show that the function to be integrated tends towards $-1/2$ when $s \to \infty$; we have again a divergent integral when $k_c \to \infty$ (or $S_c \to \infty$). Consequently, its value is a linear function of the chosen cutoff frequency k_c. However, the fact that the limit of the function to be integrated is negative indicates that, for large values of S_c, the decrease in potential energy overcomes the increase in kinetic energy.

The ground state total energy E_{ground} is the sum of all the energies we just computed, including the mean field term:

$$E_{\text{ground}} = \frac{V_0}{2}\left(N_0 + N_e\right)^2 + E_c + E_{N_0} \tag{54}$$

where $E_c + E_{N_0}$ is given by (53).

Remarque:

The divergences appearing when $S_c \to \infty$ in our calculation of the energy are not fundamental. They occur when one assumes that the matrix elements $V_{\mathbf{k}}$ of the potential written in (29) remain constant when $k \to \infty$, while they tend to zero with a realistic potential. It is indeed possible to perform a more careful treatment of the potential, such as that mentioned in § 4.2 of Référence [15], and to obtain a finite result.

3-c. Phase locking; comparison with the BCS mechanism

We just saw in § 2-a-α how all the phases $\zeta_{\mathbf{k}}$ had to become equal to the phase ζ_0 associated with the state $\mathbf{k} = \mathbf{0}$ in order to minimize the repulsive energy between any particle in the $\mathbf{k} \neq \mathbf{0}$ states and any particle in the $\mathbf{k} = \mathbf{0}$ state. Fixing the phase differences to be zero is what we called the "phase locking condition". This situation reminds us of the "symmetry breaking" of the BCS mechanism, where a common locking of all the relative phases of all the pairs $(\mathbf{k}, -\mathbf{k})$ enabled the building of a gap Δ, through a collective effect. For bosons, the equivalent of the collective mean field created by the pairs is the one created by the condensate of particles in the $\mathbf{k} = \mathbf{0}$ state. This is why it is now the relative phase of the pair states with respect to that of the condensate that plays a role; in other words, we have now an "external" instead of an "internal" mechanism. The gain in energy will be exactly the same, whatever value is chosen for the phase ζ_0; only the relative phase $\zeta_{\mathbf{k}} - \zeta_0$ is relevant. Accordingly, and just as for fermions, the arbitrary choice of the phase leads to a symmetry breaking phenomenon. The analogy is reinforced by the fact that, for a fixed value of μ, we found in § 2-b that the value of the

particle number N_0 in the $\mathbf{k} = \mathbf{0}$ state is given by an implicit equation in N_0. Similarly, in the BCS theory, it is also an implicit equation that fixes the value of the gap Δ.

Relations (40) and (53) include non-integer powers of the interaction potential V_0, and are thus non-analytic functions of that potential. They cannot be obtained by a perturbation theory as a power series expansion of V_0, and this is another analogy with the results of Complement B_{XVII}.

We also saw in that complement that, in the BCS mechanism, it is the energies in the vicinity of the Fermi level that play the most important role. This is not the case for a system of repulsive bosons. For example, in (44), the function whose integration over s yields the kinetic energy is:

$$f_c(s) = s^4 \left[\frac{s^2+1}{\sqrt{s^2(s^2+2)}} - 1 \right] \tag{55}$$

whereas the one yielding the interaction potential energy between particles in the $\mathbf{k} \neq \mathbf{0}$ states and particles in the $\mathbf{k} = \mathbf{0}$ state is:

$$f_{N_0}(s) = s^2 \left[\frac{s^2}{\sqrt{s^2(s^2+2)}} - 1 \right] \tag{56}$$

Figure 1 plots these two functions with dashed lines, as well as their sum with a solid line. It illustrates how, in the case of repulsive bosons, the effects of minimization of the potential energy overcome those of the kinetic energy. This minimization of the repulsion necessarily comes with a modification of the particles' position correlation function, which must decrease at short distances; this interpretation will be substantiated in § 3-d-β where we study the binary correlation functions. We also note in the figure that the accumulated gain of energy is not due to a particular energy band: all the s values contribute up to the limit imposed by the upper bound of the integral.

We can further refine the analysis of the energy balance by looking at the gain of energy per individual quantum state. We must then remove from the previous relations the factor k^2 coming from the density of states, and hence remove a factor s^2 from (55) and (56). Figure 2 plots the resulting functions, which show that as long as s is small or of the order of 1, the decrease in potential energy largely overcomes the increase in kinetic energy; on the other hand, the two contributions balance each other when $s \gg 1$. According to (35), the condition $s \lesssim 1$ corresponds to:

$$k \lesssim \xi^{-1} \qquad \text{that is:} \qquad e_k \lesssim N_0 V_0 \tag{57}$$

This means that individual states of low energy provide most of the decrease of the repulsive potential energy; the corresponding energy domain is proportional both to N_0 and to the interaction matrix element V_0. Relations (27) also show that it is those energies in the system ground state that the energy minimization affects the most. From the physical point of view, it is understandable that particles having a low kinetic energy compared to the interaction energy $N_0 V_0$ are the most affected by the interactions, whereas those with a kinetic energy large compared to $N_0 V_0$ have their correlations only slightly modified by the interaction potential. However, as we noted before, even though the individual contribution of the highest energy state to the energy reduction is reduced, their large number (corresponding to a density of states proportional to k^2) means that their contribution to the total energy remains significant.

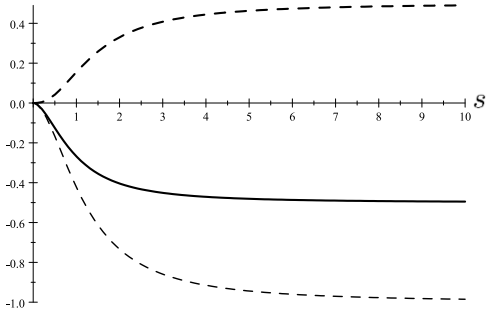


Figure 1: Plots as a function of s of the functions whose integral over s yield the kinetic energy E_c (upper dashed curve), the potential energy for the interaction with the condensate E_{N_0} (lower dashed curve), as well as their sum (solid line). The increase in the kinetic energy is overcome by the decrease in the potential energy, which ends up lowering the total energy.

3-d. Correlation functions

As the system is contained in a box and obeys periodic boundary conditions, we expect the properties of the one-body correlation functions to be translation invariant. This does not rule out a possible spatial dependence of the correlation functions, as far as the differences in positions are concerned. This is what we want to elucidate now.

α. One particle

Expanding the field operator $\Psi(\mathbf{r})$ on the annihilation operators, according to relation (A-3) of Chapter XVI, we get for the one-particle correlation function G_1

$$G_1(\mathbf{r},\mathbf{r}') = \langle\Phi_B|\,\Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}')\,|\Phi_B\rangle = \frac{1}{L^3}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'-\mathbf{k}\cdot\mathbf{r}\right)}\,\langle\Phi_B|\,a^\dagger_{\mathbf{k}}a_{\mathbf{k}'}\,|\Phi_B\rangle \tag{58}$$

where expression (3) determines $|\Phi_B\rangle$. Since in this state the particle number in an individual state $\mathbf{k}$ is always the same as that number for the individual state $-\mathbf{k}$, the average value of $a^\dagger_{\mathbf{k}}a_{\mathbf{k}'}$ in $|\Phi_B\rangle$ will be different from zero only if $\mathbf{k}=\mathbf{k}'$. In the summation over $\mathbf{k}$, the $\mathbf{k}=\mathbf{0}$ contribution introduces a term in N_0; adding to it all the other contributions, we get:

$$G_1(\mathbf{r},\mathbf{r}') = \frac{1}{L^3}\left[N_0 + \sum_{\mathbf{k}\neq\mathbf{0}}\langle\widehat{n}_{\mathbf{k}}\rangle\, e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)}\right] \tag{59}$$

When $\mathbf{r}=\mathbf{r}'$, the function G_1 is simply equal to the total particle density n_{tot}:

$$n_{\text{tot}} = \frac{N_0+N_e}{L^3} \tag{60}$$

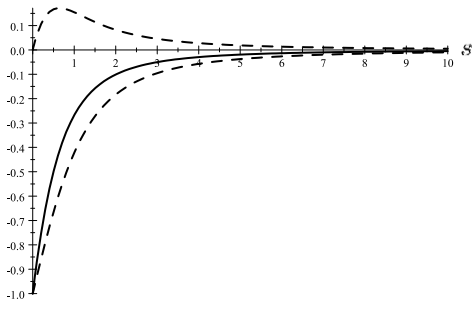


Figure 2: Plots as a function of s of the kinetic energy (upper dashed curve), the potential energy (lower dashed curve), and the total energy (solid line) per individual state. It shows that it is the lowest kinetic energy states that make the largest contribution to the lowering of the energy.

When $\mathbf{r}$ and $\mathbf{r}'$ are different, the function G_1 is the sum of two terms:

– one term corresponding to particles in the $\mathbf{k} = \mathbf{0}$ state (condensed particles), independent of the positions; this term does not decrease at large distance, but has an infinite range.

– a second term corresponding to particles in the $\mathbf{k} \neq \mathbf{0}$ state, which is the transform of the particle distribution $\langle \widehat{n}_{\mathbf{k}} \rangle$, and therefore goes to zero when $\mathbf{r}$ and $\mathbf{r}'$ move away from each other (it has a microscopic range).

We find again the Penrose-Onsager criterion according to which it is the condensed fraction of a boson system that leads to an infinite range of the non-diagonal one-body correlation function (in the case of paired fermions, we found in Complement $\mathrm{C_{XVII}}$, §§ 2-a-β and 2-b-β, that this long range does not occur for the one-body correlation function, but only for the two-body correlation function).

β. Two particles

The diagonal two-particle correlation function is written:

$$G_2(\mathbf{r},\mathbf{r};\mathbf{r}',\mathbf{r}') = \langle \Phi_B | \Psi^\dagger(\mathbf{r})\Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}')\Psi(\mathbf{r}) | \Phi_B \rangle$$
$$= \frac{1}{L^6} \sum_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''} e^{i(\mathbf{k}'''-\mathbf{k})\cdot\mathbf{r}} e^{i(\mathbf{k}''-\mathbf{k}')\cdot\mathbf{r}'} \langle \Phi_B | a_{\mathbf{k}}^\dagger a_{\mathbf{k}'}^\dagger a_{\mathbf{k}''} a_{\mathbf{k}'''} | \Phi_B \rangle \qquad (61)$$

We get the same simplifications as in § 4-b-β of Complement $\mathrm{B_{XVII}}$ in the computation of the average values of products of creation and annihilation operators: operators a_0 placed on the right each yield a factor α_0 and operators $a_0^\dagger$ placed on the left, each a factor α_0^*; the average values of the other operator products are given by the results of § C in Chapter XVII. We must distinguish between several cases, depending on the number of values, among the 4 summation indices, which are equal to $\mathbf{k} = \mathbf{0}$; we shall proceed by decreasing values of that number.

(i) If the four operators concern the $\mathbf{k} = \mathbf{0}$ state (case represented in Figure 6 of Complement B_{XVII}), we get the contribution:

$$\frac{N_0 (N_0 - 1)}{L^6} \simeq \left(\frac{N_0}{L^3} \right)^2 \tag{62}$$

which is position independent.

The case where only three of the summation indices are zero is not possible, as the corresponding term would contain the average value of an operator $a_{\mathbf{k}}$ (or of its Hermitian conjugate) in the state $|\Phi_B\rangle$, which is zero.

(ii) If one creation and one annihilation operator concern the individual $\mathbf{k} = \mathbf{0}$ state, two cases may occur and yield different types of terms:

– direct terms in $a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}'}^{\dagger} a_{\mathbf{k}'} a_{\mathbf{0}}$ or $a_{\mathbf{k}'}^{\dagger} a_{\mathbf{0}}^{\dagger} a_{\mathbf{0}} a_{\mathbf{k}'}$; both contributions are equal and their sum leads to:

$$\frac{2 N_0 N_e}{L^6} \tag{63}$$

which is also position independent.

– exchange terms in $a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}'}^{\dagger} a_{\mathbf{0}} a_{\mathbf{k}'}$ or $a_{\mathbf{k}'}^{\dagger} a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}'} a_{\mathbf{0}}$; the corresponding terms are also equal and their sum is written:

$$\begin{aligned} &\frac{N_0}{L^6} \left\{ \sum_{\mathbf{k}' \neq \mathbf{0}} e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}')} \langle \widehat{n}_{\mathbf{k}'} \rangle + \sum_{\mathbf{k} \neq \mathbf{0}} e^{i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r})} \langle \widehat{n}_{\mathbf{k}} \rangle \right\} \\ &\quad = \frac{2 N_0}{L^6} \sum_{\mathbf{k} \neq \mathbf{0}} \langle \widehat{n}_{\mathbf{k}} \rangle \cos \left[\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r}) \right] \end{aligned} \tag{64}$$

which now depends on the difference in the positions $\mathbf{r}$ and $\mathbf{r}'$. These terms reflect the existence of a bunching effect between bosons; relation (C-28) of Chapter XVII yields the value of $\langle \widehat{n}_{\mathbf{k}} \rangle$:

$$\langle \widehat{n}_{\mathbf{k}} \rangle = \sinh^2 \theta_{\mathbf{k}} \tag{65}$$

(iii) If the operators corresponding to $\mathbf{k} = \mathbf{0}$ are of the same nature (both creation or both annihilation operators), we get terms such as the ones represented on Fig 7 of Complement B_{XVII}, corresponding to the creation or annihilation of a pair from the condensate $\mathbf{k} = \mathbf{0}$:

– for a product of the type $a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger} a_{\mathbf{0}} a_{\mathbf{0}}$ where $\mathbf{k}$ and $\mathbf{k}'$ are not zero but opposite (for the same reason as explained before), we get an anomalous average value in the state $|\overline{\varphi}_{\mathbf{k}}\rangle$, multiplied by the average value of the product of two operators a_0. The first average value is given by relation (C-51) of Chapter XVII, and the second yields $(\alpha_0)^2$, that is $N_0 e^{2i\zeta_0}$ according to (7).

– for a product of the type $a_{\mathbf{0}}^{\dagger} a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}''} a_{-\mathbf{k}''}$ where $\mathbf{k}''$ and $\mathbf{k}'''$ are not zero but opposite, we get the complex conjugate of the previous result.

The sum of the two previous results is then written:

$$-\frac{N_0}{L^6}\left\{\sum_{\mathbf{k}\neq\mathbf{0}} e^{i\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})}\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}e^{2i(\zeta_0-\zeta_{\mathbf{k}})}\right.$$
$$\left.+\sum_{\mathbf{k}''\neq\mathbf{0}} e^{i\mathbf{k}''\cdot(\mathbf{r}'-\mathbf{r})}\sinh\theta_{\mathbf{k}''}\,\cosh\theta_{\mathbf{k}''}e^{-2i(\zeta_0-\zeta_{\mathbf{k}''})}\right\}$$
$$=-\frac{2N_0}{L^6}\sum_{\mathbf{k}\neq\mathbf{0}}\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\,\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})+2\left(\zeta_0-\zeta_{\mathbf{k}}\right)\right] \tag{66}$$

(iv) Finally we have terms where none of the wave vectors is zero, corresponding to cases where the particles are in $\mathbf{k}\neq\mathbf{0}$ states before and after the interaction. They include a direct term:

$$\frac{(N_e)^2}{L^6} \tag{67}$$

which is constant, an exchange term, and finally a pair annihilation-creation term. Compared to the previous terms (which are proportional N_0), their relative value is of the order of N_e/N_0. Taking into account the exchange term and the pair creation-annihilation term leads to simple calculations of a type already performed; however, to be consistent with (12) and the corresponding energy approximations, we shall ignore those terms.

The sum of (62), (63) and (67) yield the constant N^2/L^6, to which we add (64) and (66) and get:

$$G_2\left(\mathbf{r},\mathbf{r};\mathbf{r}',\mathbf{r}'\right)\simeq\frac{N^2}{L^6}+\frac{2N_0}{L^6}\sum_{\mathbf{k}\neq\mathbf{0}}\left\{\sinh^2\theta_{\mathbf{k}}\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})\right]\right.$$
$$\left.-\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\,\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})+2\left(\zeta_0-\zeta_{\mathbf{k}}\right)\right]\right\} \tag{68}$$

The position dependent term on the second line shows how the relative phases introduced in $|\Phi_B\rangle$ control the relative particle position in the physical system; it confirms that the choice $\zeta_{\mathbf{k}}=\zeta_0$ does indeed decrease the probability of finding two particles close to each other.

When the phase locking condition (19) is satisfied, the position dependent contribution becomes:

$$\frac{2N_0}{L^6}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\sinh^2\theta_{\mathbf{k}}-\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}\right]\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})\right] \tag{69}$$

Since $\sinh\theta_{\mathbf{k}}<\cosh\theta_{\mathbf{k}}$, the cosine in each term has a negative coefficient; it does decrease the probability $G_2\left(\mathbf{r},\mathbf{r};\mathbf{r},\mathbf{r}\right)$ of finding two particles at the same point $\mathbf{r}$: the dynamic correlations appearing in the system tend to "antibunch" the particles, and hence reduce their repulsive interactions. The final result is a compromise between a $\sinh^2\theta_{\mathbf{k}}$ term that leads to bunching (as for a non-interacting boson gas) and a antibunching term in $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$ that is larger, and involves anomalous average values (creation or annihilation of particle pairs in the condensate).

Comments:

(i) The correlation function (68) is invariant with respect to the exchange of $\mathbf{r}$ and $\mathbf{r}'$, as seen from the way the terms $\mathbf{k}$ and $-\mathbf{k}$ are accounted for in the summation. Its Fourier transform only contains terms in $\cos[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})]$, which can take on any value by an appropriate choice of the $\theta_{\mathbf{k}}$ and the $\zeta_{\mathbf{k}}$. As mentioned in the introduction, the variational state can lead to any correlation function; the results discussed above concern the optimal value of this correlation function.

(ii) On several occasions, we assumed the chemical potential correction $\delta\mu$, defined in (26), to be zero, which enabled us to replace the e_k by the $\widetilde{e}_k$. Let us now check that a non-zero value of this correction does not radically change the results we obtained.

Using the model (29) where the non-zero matrix elements $V_{\mathbf{k}}$ are all equal to the same constant V_0, expression (26) is simply written:

$$\delta\mu = V_0 \sum_{\mathbf{k}\neq\mathbf{0}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\right] = -V_0 \sum_{\mathbf{k}\neq\mathbf{0}} \sinh\theta_{\mathbf{k}} e^{-\theta_{\mathbf{k}}} \leq 0 \tag{70}$$

where the summation is limited to vectors $\mathbf{k}$ having a modulus less than k_c. Setting:

$$\delta\mu = -xV_0N_0 \tag{71}$$

with $x > 0$, relation (27) that fixes the $\theta_{\mathbf{k}}$ becomes:

$$\tanh 2\theta_{\mathbf{k}} = \frac{N_0V_0}{e_k + N_0V_0\,(1+x)} \tag{72}$$

The corrective effect of x and hence $\delta\mu$, is to lower the $\theta_{\mathbf{k}}$; this correction is however negligible if $e_k \gg xN_0V_0$. Consequently, the populations of the individual states $\mathbf{k}$ (which are equal to $\sinh^2\theta_{\mathbf{k}}$) decrease when $e_k \lesssim xN_0V_0$, but remain practically unchanged in the opposite case. The quantities resulting from a summation over $\mathbf{k}$, such as N_e, are then barely affected: the change in the function to be integrated only occurs for small values of s, whose contributions, in any case, are weak because of the factor s^2 in the integral (38). As for the energy, this is accentuated since the integral in (53) diverges if S_c is infinite, which means it mainly depends on the large values of s (if $S_c \gg 1$). Turning now to the correlation functions computed in § 3-d, they contain summations over $\mathbf{k}$ that lead to the same integrals; they are thus fairly insensitive to the value of $\delta\mu$. This explains why, aside from the predictions concerning the populations of small wave vectors, the approximation $\delta\mu = 0$ used in § 3 is reasonable in many cases.

To push the analysis a step further we need to compute the value of the x coefficient. This requires improving the precision of the calculations, and in particular taking into account the interactions of the particles in the $\mathbf{k} \neq \mathbf{0}$ individual states. This is beyond the scope of this complement, and we shall simply accept that x is small and note that only the small k populations are changed when $\delta\mu$ is not equal to zero.

4. Bogolubov operator method

We now present a different point of view and introduce the Bogolubov method; it is based on the search for a readily diagonalizable operator form of the Hamiltonian (or of an approximate expression of this Hamiltonian). This method not only applies to the ground state, but it also enables the study of the excited states. We shall use the results of § E in Chapter XVII to introduce new operators that simplify the diagonalization of the Hamiltonian.

4-a. Variational space, restriction on the Hamiltonian

The variational set we consider has been defined in (3); we assume:

$$|\Phi_B\rangle = |\varphi_{\mathbf{0}}\rangle \otimes |\Psi_{\text{paired}}\rangle \tag{73}$$

where $|\varphi_{\mathbf{0}}\rangle$ is the coherent state (6), and $|\Psi_{\text{paired}}\rangle$, any paired state in the Fock space spanned by all the individual states others than $\mathbf{k} = \mathbf{0}$. We call $\mathcal{E}_B(N_0)$ the ensemble of kets expressed as (73).

We now take the general Hamiltonian operator $\widehat{H}$ written in (8) of Complement $\mathrm{B_{XVII}}$, and consider its action restricted to such states; the corresponding matrix elements are of the type:

$$\langle \Phi_B | \, \widehat{H} \, | \Phi'_B \rangle \tag{74}$$

where $|\Phi_B\rangle$ and $|\Phi'_B\rangle$ are any two kets of $\mathcal{E}_B(N_0)$. In the computation of this matrix element, the same simplifications as in § 1 will occur: any annihilation operator a_0 on the right can be replaced by the number α_0, any creation operator $a_0^\dagger$ on the left by the complex conjugate α_0^*. We will further simplify the problem by assuming, as in (12), that the total population N_e of the $\mathbf{k} \neq \mathbf{0}$ individual levels is much smaller than $N_0 = |\alpha_0|^2$, and keeping only certain terms among the Hamiltonian interaction terms.

First, we study the forward scattering terms, which are the terms ($\mathbf{k}'' = \mathbf{k}$ and $\mathbf{k}''' = \mathbf{k}'$) in relation (20) of Complement $\mathrm{B_{XIII}}$. Their expression is:

$$\frac{V_0}{2} \sum_{\mathbf{k},\mathbf{k}'} a_{\mathbf{k}}^\dagger a_{\mathbf{k}'}^\dagger a_{\mathbf{k}'} a_{\mathbf{k}} = \frac{V_0}{2} \sum_{\mathbf{k},\mathbf{k}'} \left[a_{\mathbf{k}}^\dagger a_{\mathbf{k}} a_{\mathbf{k}'}^\dagger a_{\mathbf{k}'} - \delta_{\mathbf{k},\mathbf{k}'} \; a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \right] = \frac{V_0}{2} \widehat{N} \left(\widehat{N} - 1 \right) \tag{75}$$

In this equality, $\widehat{N}$ is the total number of particles operator:

$$\widehat{N} = \widehat{N}_0 + \widehat{N}_e \tag{76}$$

where $N_0 = a_0^\dagger a_0$ is the operator associated with the population of the individual state $\mathbf{k} = \mathbf{0}$, and $\widehat{N}_e$ that associated with the total number of particles in the states $\mathbf{k} \neq \mathbf{0}$:

$$\widehat{N}_e = \sum_{\mathbf{k} \neq \mathbf{0}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \tag{77}$$

As for all the other interactions terms, we shall proceed as in § 1-c and will only keep the terms that contain N_0 – the others correspond to interactions between particles in the $\mathbf{k} \neq \mathbf{0}$ individual states, assumed to be negligible when inequality (12) is satisfied. In all the terms we keep, there are either four or two creation or annihilation operators concerning the $\mathbf{k} = \mathbf{0}$ state.

Those containing the product $a_0^\dagger a_0^\dagger a_0 a_0$, or one of the two products $a_{\mathbf{k}}^\dagger a_0^\dagger a_0 a_{\mathbf{k}}$ and $a_0^\dagger a_{\mathbf{k}}^\dagger a_{\mathbf{k}} a_0$, are already taken into account in the mean field term (75). We simply have to add:

– the terms containing the products $a_{\mathbf{k}}^\dagger a_0^\dagger a_{\mathbf{k}} a_0$ or $a_0^\dagger a_{\mathbf{k}}^\dagger a_0 a_{\mathbf{k}}$, i.e. the exchange terms of § 4-b-β in Complement $\mathrm{B_{XVII}}$; they yield a contribution:

$$N_0 V_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \tag{78}$$

– the terms in $a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a_0 a_0$, corresponding to the pair creation from the condensate, or the terms in $a^\dagger_0 a^\dagger_0 a_{\mathbf{k}} a_{-\mathbf{k}}$ corresponding to the annihilation of pairs into the condensate; their contribution is:

$$\frac{N_0 V_{\mathbf{k}}}{2}\left[\mathrm{e}^{2i\zeta_0}\ a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + \mathrm{e}^{-2i\zeta_0}\ a_{\mathbf{k}} a_{-\mathbf{k}}\right] \tag{79}$$

With the above conditions, we get a simplified version of the Hamiltonian $\widehat{H}$, which becomes a reduced Hamiltonian $\widehat{H}_R$:

$$\begin{aligned}\widehat{H}_R =& \frac{V_0 \widehat{N}(\widehat{N}-1)}{2} \\ &+ \sum_{\mathbf{k}\neq\mathbf{0}} \left\{ e_k a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + N_0 V_{\mathbf{k}} \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{1}{2}\left(\mathrm{e}^{2i\zeta_0}\ a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + \mathrm{e}^{-2i\zeta_0}\ a_{\mathbf{k}} a_{-\mathbf{k}} \right) \right] \right\}\end{aligned} \tag{80}$$

If the number of particles is fixed, the first term in the right hand side (mean field) introduces only the same displacement of all energies, without physical consequence. In the Bogolobov approximation, where the condition $N_0 \gg N_e$ is assumed, one often merely replaces $\widehat{N}$ by $\widehat{N}_0$ in this term, which amounts to restricting the sum of (75) to the terms $\mathbf{k} = \mathbf{k}' = \mathbf{0}$. Since the kets (73) are eigenvectors of a_0 with eigenvalue $\sqrt{N_0}$, we may then replace the mean field operator (75) by the number $V_0 N_0^2/2$. If, moreover, one assumes that $\zeta_0 = 0$, one obtains the simpler expression:

$$\widehat{H}_R = \frac{V_0 N_0^2}{2} + \sum_{\mathbf{k}\neq\mathbf{0}} e_k a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + N_0 V_{\mathbf{k}} \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + a_{\mathbf{k}} a_{-\mathbf{k}}}{2} \right] \tag{81}$$

Either (80) or (81) can be used as the Hamiltonian within the Bogolubov approximation.

Neither of these operators conserves the total particle number, because of its terms proportional to the product of two creation or two annihilation operators. Complement B_{XVII} explained how such "anomalous" terms can, nevertheless, account for the interaction effects within the framework of certain approximations. We are now going to show that this expression can be put in the form of a Hamiltonian of independent particles, provided the operators undergo the transformation introduced in § E of Chapter XVII.

4-b. Bogolubov Hamiltonian

We obtained in Chapter XVII the expression (E-29) of the Hamiltonian operator $\widehat{H}_B$:

$$\widehat{H}_B = \sum_{\mathbf{k}\in D} \hbar\omega_k \left[b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} \right] \tag{82}$$

which includes the Bogolubov operators for bosons:

$$\begin{aligned} b_{\mathbf{k}} &= u_{\mathbf{k}} a_{\mathbf{k}} + v_{\mathbf{k}} a^\dagger_{-\mathbf{k}} \\ b_{-\mathbf{k}} &= u_{\mathbf{k}} a_{-\mathbf{k}} + v_{\mathbf{k}} a^\dagger_{\mathbf{k}} \end{aligned} \tag{83}$$

Remember that D is half the momentum space, avoiding double counting of the same pairs of states in (82). Relation (E-15) of Chapter XVII expresses the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ in terms of the two parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$:

$$\begin{aligned} u_{\mathbf{k}} &= \cosh\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sinh\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}} \end{aligned} \tag{84}$$

As for the value of the parameter ω_k, it will be fixed later.

We then have:

$$b^\dagger_{\mathbf{k}} b_{\mathbf{k}} = |u_{\mathbf{k}}|^2\ \ a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + |v_{\mathbf{k}}|^2\ \ a_{-\mathbf{k}} a^\dagger_{-\mathbf{k}} + \ \ \left(u^*_{\mathbf{k}} v_{\mathbf{k}} a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} + u_{\mathbf{k}} v^*_{\mathbf{k}} a_{-\mathbf{k}} a_{\mathbf{k}}\right) \tag{85}$$

and:

$$b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} = |u_{\mathbf{k}}|^2\ \ a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}} + |v_{\mathbf{k}}|^2\ \ a_{\mathbf{k}} a^\dagger_{\mathbf{k}} + \ \ \left(u^*_{\mathbf{k}} v_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + u_{\mathbf{k}} v^*_{\mathbf{k}} a_{\mathbf{k}} a_{-\mathbf{k}}\right) \tag{86}$$

The operators in these equalities can be rearranged in normal order, using the proper commutation; adding them both, we get:

$$\begin{aligned} b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} = &\left(|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2\right)\left(a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right) \\ &+ 2\,|v_{\mathbf{k}}|^2 + 2u^*_{\mathbf{k}} v_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + 2u_{\mathbf{k}} v^*_{\mathbf{k}} a_{\mathbf{k}} a_{-\mathbf{k}} \end{aligned} \tag{87}$$

that is, taking (84) into account:

$$\begin{aligned} b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} = &\cosh 2\theta_{\mathbf{k}}\ \ \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right] \\ &+ 2\sinh^2\theta_{\mathbf{k}} + \sinh 2\theta_{\mathbf{k}} \left[a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} + a_{\mathbf{k}} a_{-\mathbf{k}}\ e^{-2i\zeta_{\mathbf{k}}}\right] \end{aligned} \tag{88}$$

Operator $\widehat{H}_B$ can therefore be written:

$$\begin{aligned} \widehat{H}_B = \sum_{\mathbf{k}\in D} \hbar\omega_k \Big\{ &\cosh 2\theta_{\mathbf{k}}\ \ \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right] \\ &+ 2\sinh^2\theta_{\mathbf{k}} + \sinh 2\theta_{\mathbf{k}} \left[a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} + a_{\mathbf{k}} a_{-\mathbf{k}}\ e^{-2i\zeta_{\mathbf{k}}}\right] \Big\} \end{aligned} \tag{89}$$

Now this Hamiltonian may be identified with the approximate Hamiltonian (80). To see this, we replace $\cosh 2\theta_{\mathbf{k}}$ by expression (30), $\sinh 2\theta_{\mathbf{k}}$ by the double of (47), and $\sinh^2\theta_{\mathbf{k}}$ by (32); we are still in the simplified model where $\delta\mu = 0$, and hence the $\widetilde{e}_k$ are replaced by the e_k. Finally we choose for ω_k the value:

$$\hbar\omega_k = \sqrt{e_k\,(e_k + 2N_0 V_{\mathbf{k}})} \tag{90}$$

and we assume that all the $\zeta_{\mathbf{k}}$ are zero. We then get:

$$\begin{aligned} \widehat{H}_B = \sum_{\mathbf{k}\in D} \Big\{ &(e_k + N_0 V_{\mathbf{k}}) \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right] \\ &+ \frac{N_0 V_{\mathbf{k}}}{2} \left[a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}}\ + a_{\mathbf{k}} a_{-\mathbf{k}}\right] \Big\} - \Delta E_{\text{fond}} \end{aligned} \tag{91}$$

with, again using the value (32) for $\sinh^2\theta_{\mathbf{k}}$:

$$\Delta E_{\text{ground}} = -2\sum_{\mathbf{k}\in D}\hbar\omega_k\,\sinh^2\theta_{\mathbf{k}} = -\sum_{\mathbf{k}\neq\mathbf{0}}\hbar\omega_k\,\sinh^2\theta_{\mathbf{k}}$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\sqrt{e_k\,(e_k+2N_0V_{\mathbf{k}})}-e_k-N_0V_{\mathbf{k}}\right] \tag{92}$$

Comparison with relations (52) and (54) shows that ΔE_{ground} is none other than the energy E_{ground} already obtained, shifted by the mean field value:

$$\Delta E_{\text{ground}} = E_{\text{ground}} - \frac{V_0N^2}{2} \tag{93}$$

Finally, taking (80) into account, if the $\zeta_{\mathbf{k}}$ are all chosen equal to zero (phase locking condition), we simply have:

$$\widehat{H}_R = \widehat{H}_B + E_{\text{ground}} \tag{94}$$

4-c. Constructing a basis of excited states, quasi-particles

As E_{ground} is a number, it introduces a simple energy shift in the eigenvalues of $\widehat{H}_R$ compared to those of $\widehat{H}_B$, with no effect on the eigenvectors. Now we saw in § E-3 of Chapter XVII that the eigenvalues of $\widehat{H}_B$ are known, and can be written as:

$$E_B = \sum_{\mathbf{k}\in D}\left[n(b_{\mathbf{k}})+n(b_{-\mathbf{k}})\right]\hbar\omega_k \tag{95}$$

where $n(b_{\mathbf{k}})$ and $n(b_{-\mathbf{k}})$ are any positive or zero integers. As for the eigenstates associated with these energies, they can be simply obtained by the action on the ground state of the following product of creation operators:

$$\prod_{\mathbf{k}\in D}\left[b^\dagger_{\mathbf{k}}\right]^{n(b_{\mathbf{k}})}\left[b^\dagger_{-\mathbf{k}}\right]^{n(b_{-\mathbf{k}})} \tag{96}$$

All things considered, the operator $\widehat{H}_B$ shares a lot of properties with the Hamiltonian of an ensemble of non-interacting particles. Just as the usual creation operators permit adding particles in a system of free identical particles, the $b^\dagger_{\mathbf{k}}$ and $b^\dagger_{-\mathbf{k}}$ creation operators can be considered as adding an extra "quasi-particle" to the physical system. When acting on the ground state, the operator $b^\dagger_{\mathbf{k}}$ yields a ket where both the energy and the momentum are well-defined: the energy is increased by the amount $\hbar\omega_k$ specified in (90), the momentum is increased by $\hbar\mathbf{k}$ with respect to the zero momentum of the ground state. This exact change of momentum occurs because the action of $b^\dagger_{\mathbf{k}}$ on any ket creates two components: one component where one particle with momentum $\hbar\mathbf{k}$ is added, the other component where one particle with momentum $-\hbar\mathbf{k}$ is suppressed. In both cases, the total momentum has increased[5] by the same amount $\hbar\mathbf{k}$. The operator $b^\dagger_{\mathbf{k}}$

[5]This result can also be verified by calculating the commutator $\left[\mathbf{P},b^\dagger_{\mathbf{k}}\right]$ of the total momentum $\mathbf{P}=\sum_{\mathbf{k}}\hbar\mathbf{k}\;a^\dagger_{\mathbf{k}}a_{\mathbf{k}}$ with $b^\dagger_{\mathbf{k}}$. One obtains $\left[\mathbf{P},b^\dagger_{\mathbf{k}}\right]=u^*_{\mathbf{k}}\left[\hbar\mathbf{k}a^\dagger_{\mathbf{k}}a_{\mathbf{k}},a^\dagger_{\mathbf{k}}\right]+v^*_{\mathbf{k}}\left[-\hbar\mathbf{k}a^\dagger_{-\mathbf{k}}a_{-\mathbf{k}},a_{-\mathbf{k}}\right]$, or: $\left[\mathbf{P},b^\dagger_{\mathbf{k}}\right]=\hbar\mathbf{k}\left(u^*_{\mathbf{k}}a^\dagger_{\mathbf{k}}+v^*_{\mathbf{k}}a_{-\mathbf{k}}\right)=\hbar\mathbf{k}\;b^\dagger_{\mathbf{k}}$. As a consequence, the effect of $b^\dagger_{\mathbf{k}}$ on any eigenstate of $\mathbf{P}$ is to increase its eigenvalue by $\hbar\mathbf{k}$.

therefore creates a quasi-particle of well-defined energy and momentum, and $b_{\mathbf{k}}$ annihilates it; of course, $b^{\dagger}_{-\mathbf{k}}$ and $b_{-\mathbf{k}}$ have the same properties for the quasi-particle of opposite momentum. These quasi-particles do not coincide with particles of a system without interactions, as can be seen from the expression of those creation operators. They yield, however, a basis of states that permits reasoning as if there were no interactions; this provides a very powerful point of view in many fields of physics.

We can assume, as in Complement $\mathrm{D_{XV}}$, that the interaction potential has a zero range:

$$W_2(\mathbf{r}, \mathbf{r}') = g\ \delta(\mathbf{r} - \mathbf{r}') \tag{97}$$

Relation (11) then becomes:

$$V_{\mathbf{k}} = \frac{1}{L^3} \int \mathrm{d}^3 r\ e^{-i\mathbf{q}\cdot\mathbf{r}}\ W_2(\mathbf{r}) = \frac{g}{L^3} \tag{98}$$

and equality (90) is written as:

$$\hbar\omega(\mathbf{k}) = \sqrt{e_k\,(e_k + 2gn_0)} = \frac{\hbar^2}{2m}\sqrt{k^2\,(k^2 + k_0^2)} \tag{99}$$

with:

$$k_0 = \frac{2}{\hbar}\sqrt{mgn_0} = \frac{\sqrt{2}}{\xi} \tag{100}$$

In this last equation, ξ is the healing length defined in (36). This equation is the same as relation (34) of Complement $\mathrm{D_{XV}}$, whose Figure 1 represents the quasi-particle spectrum. When the modulus k of the wave vector $\mathbf{k}$ is smaller than the wave vector k_0, we get a linear spectrum whose slope corresponds to the sound velocity in the boson system; for values larger than k_0, the spectrum becomes quadratic, as for free particles.

Conclusion

The calculations presented in this complement illustrate the analogy between the pairing phenomena for attractive fermions and for repulsive bosons. In both cases, binary position correlations are introduced by the dynamic interactions, resulting in a decrease of the interaction potential energy of the physical system; the paired states are a valuable tool for understanding this effect. In both cases, a relative phase locking phenomenon occurs, but the precise nature of that locking is, however, different.

For fermions, the energy gain is due to a collective effect, involving the pair-pair interactions and the relative phase of every pair of states $(\mathbf{k}, -\mathbf{k})$; each contributes to the value of the gap Δ which, in turn, has an effect on all the others – this is translated mathematically by the presence of a double summation over $\mathbf{k}$ and $\mathbf{k}'$ in the energy. This is reminiscent of a ferromagnetic system, where each spin contributes to the collective exchange field that act on all its neighbors. As the interactions are supposed to be attractive, the phase locking to zero maximizes the pair-pair interactions, and hence minimizes the energy.

For bosons, the major role is played by the relative phase of the pairs with respect to that of the reservoir composed of all the particles in the $\mathbf{k} = \mathbf{0}$ state (condensate).

The physical process involved is illustrated in Figure 7 of Complement B_{XVII}, where two particles emerge from the condensate to form a pair, or vice-versa – mathematically, the energy term contains only one summation over **k**. The relative phase locking it introduces will minimize the repulsion between these pairs and the condensate, and hence the total energy. Compared to the fermion case, the presence of a condensate independent of the pairs radically changes the nature of the phase locking.

Chapter XVIII

Review of classical electrodynamics

A Classical electrodynamics 1959
A-1 Basic equations and relations 1959
A-2 Description in the reciprocal space 1960
A-3 Elimination of the longitudinal fields from the expression of the physical quantities 1965
B Describing the transverse field as an ensemble of harmonic oscillators 1968
B-1 Brief review of the one-dimensional harmonic oscillator . . . 1968
B-2 Normal variables for the transverse field 1969
B-3 Discrete modes in a box 1974
B-4 Generalization of the mode concept 1975

Introduction

In the three previous chapters, we studied ensembles of identical particles, which allowed us to introduce the concept of quantum field operators. We now begin a new series of three chapters where this quantum field concept is applied to an important particular case: the electromagnetic field, made of identical bosons called "photons". We start by noting that, in classical electromagnetism, the dynamics of the different field modes is exactly similar to that of a series of harmonic oscillators. Each of these modes may be quantized by the same method as that used for an elementary harmonic oscillator, for a single particle; this method has the great advantage of simplicity. It requires, however, establishing beforehand the equivalence between modes of the classical electrodynamic field and harmonic oscillators; this is the main purpose of the present chapter.

For the presentation to be self-contained, we first review a certain number of properties of classical electromagnetism. One complement is also devoted to a synthetic

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

presentation of the Lagrangian formalism applied to this case. The reader already familiar with those aspects of classical electrodynamics may wish to go directly to the quantum treatment presented in Chapter XIX.

We start in § A with the equations of Maxwell-Lorentz describing the coupled evolution of the electric field $\boldsymbol{E}(\boldsymbol{r},t)$, the magnetic field $\boldsymbol{B}(\boldsymbol{r},t)$ and the coordinates and speeds of the particles acting as source for this electromagnetic field[1]. We shall give the expressions for a certain number of constants of motion, such as the energy, or the linear and angular momenta of the global system "field + particles". The vector potential $\boldsymbol{A}(\boldsymbol{r},t)$ and scalar potential $U(\boldsymbol{r},t)$ will also be introduced, as well as the gauge transformations that can be performed on these potentials.

We shall then show that it is useful to take the spatial Fourier transforms of these fields, since in the reciprocal space, Maxwell's equations have a simpler form. For a free electromagnetic field (in the absence of charged particles), they are no longer partial differential equations, as in ordinary space, but ordinary time-dependent differential equations. Furthermore, the concept of longitudinal or transverse field vectors has a clear geometrical significance in the reciprocal space[2]. A field vector $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ is longitudinal if $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ is parallel to $\boldsymbol{k}$ at every point $\boldsymbol{k}$ of the reciprocal space, transverse if $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ is perpendicular to $\boldsymbol{k}$ at every point $\boldsymbol{k}$. We will show that two of the four Maxwell's equations yield the value of the longitudinal electrical and magnetic fields, whereas the other two describe the evolution of the transverse fields. It will become clear that the longitudinal electric field is simply the Coulomb electrostatic field created by the charged particles. Consequently, it is not an independent field variable since it only depends on the coordinates of the particles[3]. Furthermore, choosing the Coulomb gauge amounts to choosing the longitudinal potential vector equal to zero; this permits eliminating the longitudinal fields from the expressions for all the physical quantities.

In § B, we establish the equivalence between the radiation field and an ensemble of one-dimensional harmonic oscillators. Maxwell's equations for transverse fields enable introducing linear combinations of the vector potentials and transverse electric fields, whose time evolution, in the absence of particles, is of the form $e^{-i\omega t}$ where $\omega = ck$. These variables, called *normal variables*, thus describe the eigenmodes of the free field vibrations. The dynamics of each of these eigenmodes is similar to that of a one-dimensional harmonic oscillator. The normal mode variable is the equivalent of the linear combination of the position and velocity of the associated operator, and becomes, in the quantization process, the annihilation operator, fundamental in the quantum theory of the harmonic oscillator. Replacing the normal variables and their complex conjugates by annihilation and creation operators will yield, in Chapter XIX, the expressions for the various operators of the quantum theory.

[1] We assume that the speeds of the particles are small compared to the speed of light, so as to use a non-relativistic description.

[2] We shall note $\tilde{G}(\boldsymbol{k})$ the spatial Fourier transform of $G(\boldsymbol{r})$, the symbol "tilde" allowing a clear distinction between the functions in ordinary and reciprocal space.

[3] As for the longitudinal magnetic field, it is simply zero.

A. Classical electrodynamics

A-1. Basic equations and relations

A-1-a. Maxwell's equations

There are four Maxwell's equations in vacuum, and in the presence of sources:

$$\boldsymbol{\nabla} \cdot \boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{\varepsilon_0}\rho(\boldsymbol{r},t) \tag{A-1a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{B}(\boldsymbol{r},t) = 0 \tag{A-1b}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r},t) = -\frac{\partial}{\partial t}\boldsymbol{B}(\boldsymbol{r},t) \tag{A-1c}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r},t) = \frac{1}{c^2}\frac{\partial}{\partial t}\boldsymbol{E}(\boldsymbol{r},t) + \frac{1}{\varepsilon_0 c^2}\boldsymbol{j}(\boldsymbol{r},t) \tag{A-1d}$$

where c is the velocity of light in vacuum and ε_0 the vacuum permittivity. These equations yield the divergence and the curl of the electric field $\boldsymbol{E}(\boldsymbol{r},t)$ and the magnetic field $\boldsymbol{B}(\boldsymbol{r},t)$. The charge density $\rho(\boldsymbol{r},t)$ and current density $\boldsymbol{j}(\boldsymbol{r},t)$ appearing in those equations can be expressed, in the non-relativistic limit, in terms of the positions $\boldsymbol{r}_a(t)$ and the speeds $\boldsymbol{v}_a(t) = \mathrm{d}\boldsymbol{r}_a(t)/\mathrm{d}t$ of the various particles a of the system, each having a mass m_a and a charge q_a:

$$\rho(\boldsymbol{r},t) = \sum_a q_a \delta\left[\boldsymbol{r} - \boldsymbol{r}_a(t)\right] \tag{A-2a}$$

$$\boldsymbol{j}(\boldsymbol{r},t) = \sum_a q_a \boldsymbol{v}_a(t) \delta\left[\boldsymbol{r} - \boldsymbol{r}_a(t)\right] \tag{A-2b}$$

A-1-b. Lorentz Equations

Lorentz equations describe the dynamics of each particle a submitted to the electric and magnetic forces exerted by the fields $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$:

$$m_a \frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{r}_a(t) = q_a\left[\boldsymbol{E}\left(\boldsymbol{r}_a(t),t\right) + \boldsymbol{v}_a(t) \times \boldsymbol{B}\left(\boldsymbol{r}_a(t),t\right)\right] \tag{A-3}$$

The particle and field evolutions are coupled: the particles move under the effect of the forces the fields exert on them, but they also act as sources for the evolution of those fields.

A-1-c. Constants of motion

Definitions (A-2a) of $\rho(\boldsymbol{r},t)$ and (A-2b) of $\boldsymbol{j}(\boldsymbol{r},t)$ lead to the continuity equation:

$$\frac{\partial}{\partial t}\rho(\boldsymbol{r},t) + \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r},t) = 0 \tag{A-4}$$

which implies the time invariance of the total charge of the particle system:

$$Q = \int \mathrm{d}^3 r \rho(\boldsymbol{r},t) = \sum_a q_a \tag{A-5}$$

Other constants of motion exist: the total energy H, the total momentum $\boldsymbol{P}$ and the total angular momentum $\boldsymbol{J}$ of the system field + particles. They are respectively given by:

$$H = \sum_a \frac{1}{2} m_a \boldsymbol{v}_a^2(t) + \frac{\varepsilon_0}{2} \int \mathrm{d}^3 r \left[\boldsymbol{E}^2(\boldsymbol{r},t) + c^2 \boldsymbol{B}^2(\boldsymbol{r},t) \right] \tag{A-6a}$$

$$\boldsymbol{P} = \sum_a m_a \boldsymbol{v}_a(t) + \varepsilon_0 \int \mathrm{d}^3 r \; \boldsymbol{E}(\boldsymbol{r},t) \times \boldsymbol{B}(\boldsymbol{r},t) \tag{A-6b}$$

$$\boldsymbol{J} = \sum_a \boldsymbol{r}_a(t) \times m_a \boldsymbol{v}_a(t) + \varepsilon_0 \int \mathrm{d}^3 r \; \boldsymbol{r} \times [\boldsymbol{E}(\boldsymbol{r},t) \times \boldsymbol{B}(\boldsymbol{r},t)] \tag{A-6c}$$

Using (A-1) and (A-3), we can verify that the derivatives with respect to time of H, $\boldsymbol{P}$ and $\boldsymbol{J}$ are indeed zero (for H and $\boldsymbol{P}$, see for example exercise 1 in Complement $\mathrm{C_I}$ of [16] and its correction).

A-1-d. Scalar and vector potentials: gauge transformations

As we already saw in Complement $\mathrm{H_{III}}$, the fields $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$ can always be written in the form:

$$\boldsymbol{E}(\boldsymbol{r},t) = -\boldsymbol{\nabla} U(\boldsymbol{r},t) - \frac{\partial}{\partial t} \boldsymbol{A}(\boldsymbol{r},t) \tag{A-7a}$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r},t) \tag{A-7b}$$

where $\boldsymbol{A}(\boldsymbol{r},t)$ and $U(\boldsymbol{r},t)$ are the vector and scalar potentials defining a *gauge*. For any function $\chi(\boldsymbol{r},t)$ of $\boldsymbol{r}$ and of t, the transformation of these potentials obeying the relations:

$$\boldsymbol{A}(\boldsymbol{r},t) \Rightarrow \boldsymbol{A}'(\boldsymbol{r},t) = \boldsymbol{A}(\boldsymbol{r},t) + \boldsymbol{\nabla} \chi(\boldsymbol{r},t) \tag{A-8a}$$

$$U(\boldsymbol{r},t) \Rightarrow U'(\boldsymbol{r},t) = U(\boldsymbol{r},t) - \frac{\partial}{\partial t} \chi(\boldsymbol{r},t) \tag{A-8b}$$

leads to the same expression for $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$; the same physical fields can therefore be represented by several different potentials $\boldsymbol{A}(\boldsymbol{r},t)$ and $U(\boldsymbol{r},t)$. The transformation (A-8) associated with the function $\chi(\boldsymbol{r},t)$ is called a *gauge transformation*.

Relations (A-8) allow a flexibility on the choice of the gauge $\{\boldsymbol{A}, U\}$, which allows introducing an additional condition. The Coulomb gauge, which we will use in this chapter and the following, is defined by the condition:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r},t) = 0 \tag{A-9}$$

A geometrical interpretation of condition (A-9) in the reciprocal space will be given later.

A-2. Description in the reciprocal space

Using Fourier transforms, the equations of electrodynamics can be put in a form that simplifies calculations.

A-2-a. Spatial Fourier transforms

Let us introduce the Fourier transform of the electric field $\boldsymbol{E}(\boldsymbol{r},t)$:

$$\tilde{\boldsymbol{E}}(\boldsymbol{k},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 r\, \boldsymbol{E}(\boldsymbol{r},t) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{A-10}$$

which enables us to write $\boldsymbol{E}(\boldsymbol{r},t)$ as:

$$\boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 k\, \tilde{\boldsymbol{E}}(\boldsymbol{k},t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{A-11}$$

Analogous expressions can be written for all the physical quantities we just introduced: magnetic field, charge and current densities, scalar and vector potentials.

It will be useful in what follows to recall the Parseval-Plancherel relation (Appendix I, § 2-c) showing the identity of the scalar products of two functions expressed in position space or in reciprocal space[4]:

$$\int \mathrm{d}^3 r\, F^*(\boldsymbol{r}) G(\boldsymbol{r}) = \int \mathrm{d}^3 k\, \tilde{F}^*(\boldsymbol{k}) \tilde{G}(\boldsymbol{k}) \tag{A-12}$$

and the fact that the product of two functions in reciprocal space, is the Fourier transform of their convolution in position space:

$$\tilde{F}(\boldsymbol{k})\tilde{G}(\boldsymbol{k}) \underset{\mathrm{FT}}{\Leftrightarrow} \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 r'\, F(\boldsymbol{r}') G(\boldsymbol{r}-\boldsymbol{r}') \tag{A-13}$$

A-2-b. Maxwell's equations in reciprocal space

Maxwell's equations take on a simpler form in the reciprocal space, clearly showing the differences between the longitudinal and transverse components of the various fields. Any vector field $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ can be decomposed into a longitudinal field $\tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t)$, parallel at any point $\boldsymbol{k}$ to the vector $\boldsymbol{k}$, and a transverse field $\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{k},t)$ perpendicular to $\boldsymbol{k}$:

$$\tilde{\boldsymbol{V}}(\boldsymbol{k},t) = \tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t) + \tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{k},t) \tag{A-14}$$

with:

$$\tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t) = \boldsymbol{\kappa}\, \left(\boldsymbol{\kappa}\cdot\tilde{\boldsymbol{V}}(\boldsymbol{k},t)\right) = \boldsymbol{k}\, \left(\boldsymbol{k}\cdot\tilde{\boldsymbol{V}}(\boldsymbol{k},t)\right)/k^2 \tag{A-15a}$$

$$\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{k},t) = \tilde{\boldsymbol{V}}(\boldsymbol{k},t) - \tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t) \tag{A-15b}$$

where

$$\boldsymbol{\kappa} = \boldsymbol{k}/k \tag{A-16}$$

is the unit vector along $\boldsymbol{k}$.

[4]The space of the vectors $\boldsymbol{r}$ (ordinary space) is called "position space" whereas "reciprocal space" is the space of the wave vectors $\boldsymbol{k}$.

As the operator $\boldsymbol{\nabla}$ in position space corresponds to the operator $i\boldsymbol{k}$ in reciprocal space, Maxwell's equations (A-1) become in reciprocal space:

$$i\boldsymbol{k}\cdot\tilde{\boldsymbol{E}}(\boldsymbol{k},t)=\frac{1}{\varepsilon_0}\tilde{\rho}(\boldsymbol{k},t) \tag{A-17a}$$

$$i\boldsymbol{k}\cdot\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=0 \tag{A-17b}$$

$$i\boldsymbol{k}\times\tilde{\boldsymbol{E}}(\boldsymbol{k},t)=-\frac{\partial}{\partial t}\tilde{\boldsymbol{B}}(\boldsymbol{k},t) \tag{A-17c}$$

$$i\boldsymbol{k}\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=\frac{1}{c^2}\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}(\boldsymbol{k},t)+\frac{1}{\varepsilon_0 c^2}\tilde{\boldsymbol{j}}(\boldsymbol{k},t) \tag{A-17d}$$

Taking into account definitions (A-15) for the longitudinal and transverse components of a vector field, the first two equations (A-17a) and (A-17b) determine the longitudinal parts, projections of the fields $\tilde{\boldsymbol{E}}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$ onto $\boldsymbol{k}$:

$$\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)=-\frac{i}{\varepsilon_0}\tilde{\rho}(\boldsymbol{k},t)\frac{\boldsymbol{k}}{k^2} \tag{A-18a}$$

$$\tilde{\boldsymbol{B}}_{\parallel}(\boldsymbol{k},t)=\boldsymbol{0} \tag{A-18b}$$

The last two equations (A-17c) and (A-17d) yield the rate of change $\partial\tilde{\boldsymbol{E}}(\boldsymbol{k},t)/\partial t$ and $\partial\tilde{\boldsymbol{B}}(\boldsymbol{k},t)/\partial t$ of the fields $\tilde{\boldsymbol{E}}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$, and are the equations of motion of these fields. In the absence of sources ($\tilde{\boldsymbol{j}}(\boldsymbol{k},t)=\boldsymbol{0}$), i.e. for what we will call a "free" field, they are time-dependent differential equations, and no longer partial derivative equations as is the case in position space.

A-2-c. Longitudinal electric and magnetic fields

Equation (A-18b) shows the longitudinal magnetic field $\tilde{\boldsymbol{B}}_{\parallel}(\boldsymbol{k},t)$ is zero. Equation (A-18a) expresses $\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)$ as a product of two functions of $\boldsymbol{k}$, $\tilde{\rho}(\boldsymbol{k},t)$ and $-i\boldsymbol{k}/\varepsilon_0 k^2$ whose Fourier transforms are written (relation (63) of Appendix I):

$$\tilde{\rho}(\boldsymbol{k},t)\ \underset{\text{FT}}{\leftrightarrow}\ \rho(\boldsymbol{r},t) \tag{A-19a}$$

$$-\frac{i}{\varepsilon_0}\frac{\boldsymbol{k}}{k^2}\ \underset{\text{FT}}{\leftrightarrow}\ \frac{(2\pi)^{3/2}}{4\pi\varepsilon_0}\frac{\boldsymbol{r}}{r^3} \tag{A-19b}$$

Using relation (A-13) then leads to:

$$\begin{aligned}\boldsymbol{E}_{\parallel}(\boldsymbol{r},t)&=\frac{1}{4\pi\varepsilon_0}\int \mathrm{d}^3r'\rho(\boldsymbol{r}',t)\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\\&=\frac{1}{4\pi\varepsilon_0}\sum_a q_a\frac{\boldsymbol{r}-\boldsymbol{r}_a(t)}{|\boldsymbol{r}-\boldsymbol{r}_a(t)|^3}\end{aligned} \tag{A-20}$$

This means that at time t, the longitudinal electric field coincides with the Coulomb field produced by the charge distribution $\rho(\boldsymbol{r},t)$, computed as if this distribution were static and fixed at that instant t.

Comment

The fact that the longitudinal electric field instantaneously follows the evolution of the charge distribution $\rho(\boldsymbol{r},t)$ should not lead us to believe in an *action at a distance* propagating at an infinite speed. The contribution of the transverse field must also be taken into account, as only the total electric field $\boldsymbol{E}=\boldsymbol{E}_{\parallel}+\boldsymbol{E}_{\perp}$ has a real physical meaning. It can be shown that the transverse electric field also has an instantaneous component, which balances exactly the longitudinal component so that the total field is always retarded (to $t-|r-r'|/c$), as the electromagnetic interactions propagate at the speed of light c (see exercise 3 and its correction in Complement C_I of reference [16]).

The previous results show that the longitudinal fields are not independent quantities: they are either zero (in the case of the longitudinal magnetic field), or simply related to the particle coordinates $\boldsymbol{r}_a(t)$ (in the case of the longitudinal electric field, whose expression is given by (A-20)).

A-2-d. Time evolution of the transverse fields

Now that we showed that the first two Maxwell's equations determine the longitudinal part of the fields, let us consider the last two equations (A-17c) and (A-17d) and focus on their transverse components. Since $\boldsymbol{k}\times\boldsymbol{E}=\boldsymbol{k}\times\boldsymbol{E}_{\perp}$, they can be rewritten as:

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=-i\boldsymbol{k}\times\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t) \tag{A-21a}$$

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=ic^2\boldsymbol{k}\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)-\frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_{\perp}(\boldsymbol{k},t) \tag{A-21b}$$

which yield the time evolution of the transverse fields $\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$.

Comment

One can also study the longitudinal projections of the two Maxwell's equations (A-17c) and (A-17d). The result is trivial for the first one: as both sides of the equation are transverse, their longitudinal projections are zero. As for the second equation, (A-17d), it leads to:

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)+\frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_{\parallel}(\boldsymbol{k},t)=0. \tag{A-22}$$

Taking the scalar product of $\boldsymbol{k}$ with each side of this equation, using (A-18a) and the fact that $\boldsymbol{k}\cdot\tilde{\boldsymbol{j}}=\boldsymbol{k}\cdot\tilde{\boldsymbol{j}}_{\parallel}$, we find:

$$\frac{\partial}{\partial t}\tilde{\rho}(\boldsymbol{k},t)+i\boldsymbol{k}\cdot\tilde{\boldsymbol{j}}(\boldsymbol{k},t)=0 \tag{A-23}$$

which is simply the continuity equation (A-4) in the reciprocal space, and does not provide any new information.

A-2-e. Potentials

In the reciprocal space, relations (A-7a) and (A-7b) between fields and potentials become:

$$\tilde{\boldsymbol{E}}(\boldsymbol{k},t)=-i\boldsymbol{k}\tilde{U}(\boldsymbol{k},t)-\frac{\partial}{\partial t}\tilde{\boldsymbol{A}}(\boldsymbol{k},t) \tag{A-24a}$$

$$\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=i\boldsymbol{k}\times\tilde{\boldsymbol{A}}(\boldsymbol{k},t) \tag{A-24b}$$

and the gauge transformations relations (A-8a) and (A-8b) are written:

$$\tilde{\boldsymbol{A}}(\boldsymbol{k},t) \to \tilde{\boldsymbol{A}}'(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}(\boldsymbol{k},t) + i\boldsymbol{k}\tilde{\chi}(\boldsymbol{k},t) \tag{A-25a}$$

$$\tilde{U}(\boldsymbol{k},t) \to \tilde{U}'(\boldsymbol{k},t) = \tilde{U}(\boldsymbol{k},t) - \frac{\partial}{\partial t}\tilde{\chi}(\boldsymbol{k},t) \tag{A-25b}$$

where $\tilde{\chi}(\boldsymbol{k},t)$ is the Fourier transform of $\chi(\boldsymbol{r},t)$.

Since the last term in (A-25a) is a longitudinal vector, it is clear that a gauge transformation does not change the transverse part $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$, which thus defines a gauge invariant physical field:

$$\tilde{\boldsymbol{A}}'_\perp(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{A-26}$$

Since $\boldsymbol{k} \times \tilde{\boldsymbol{A}}_\parallel = \boldsymbol{0}$, the transverse projections of relations (A-24a) and (A-24b) yield the equations:

$$\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t) = -\frac{\partial}{\partial t}\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{A-27a}$$

$$\tilde{\boldsymbol{B}}(\boldsymbol{k},t) = i\boldsymbol{k} \times \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{A-27b}$$

Note that equation (A-27b) allows expressing $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ as a function of $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$, as we now show. Taking the vector product of $\boldsymbol{k}$ with each side of this equation, and using the identity:

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}, \tag{A-28}$$

and the fact that $\boldsymbol{k} \cdot \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = 0$, we get:

$$\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \frac{i}{k^2}\left(\boldsymbol{k} \times \tilde{\boldsymbol{B}}(\boldsymbol{k},t)\right) \tag{A-29}$$

This equation, together with equation (A-27a), allow rewriting the two time evolution equations (A-21a) and (A-21b) for the transverse fields in a form only involving $\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$:

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = -\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t) \tag{A-30a}$$

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t) = c^2k^2\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) - \frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) \tag{A-30b}$$

In the absence of sources ($\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) = \boldsymbol{0}$), we get two coupled time evolution equations for the transverse fields $\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$. They will be useful later on for introducing the field normal variables, and for the demonstration of the equivalence between the transverse field and an ensemble of harmonic oscillators.

Time evolution equation for the transverse potential vector

The time evolution equation for $\tilde{\boldsymbol{A}}_\perp$ can be obtained by replacing $\tilde{\boldsymbol{E}}_\perp$ in (A-30b) by $-\partial\tilde{\boldsymbol{A}}_\perp/\partial t$. We obtain:

$$\left[\frac{\partial^2}{\partial t^2} + c^2k^2\right]\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) \tag{A-31}$$

which is written, in the position space:

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta\right]\boldsymbol{A}_\perp(\boldsymbol{r},t) = \frac{1}{\varepsilon_0 c^2}\boldsymbol{j}_\perp(\boldsymbol{r},t) \tag{A-32}$$

A-2-f. Coulomb gauge

Condition $\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r},t) = 0$, which defined in (A-9) the Coulomb gauge, becomes in the reciprocal space:

$$i\boldsymbol{k} \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k},t) = 0 \quad \longleftrightarrow \quad \tilde{\boldsymbol{A}}_{\parallel}(\boldsymbol{k},t) = 0 \tag{A-33}$$

In the Coulomb gauge, the longitudinal vector potential is therefore equal to zero; there only remains the transverse vector potential, which, as mentioned above, is a physical field.

What can be said about the scalar potential U in the Coulomb gauge? Let us consider the longitudinal part of each side of equation (A-24a). As the last term on the right-hand side is transverse in the Coulomb gauge, we get $\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t) = -i\boldsymbol{k}\tilde{U}(\boldsymbol{k},t)$, which reads, in position space, $\boldsymbol{E}_{\parallel}(\boldsymbol{r},t) = -\boldsymbol{\nabla}U(\boldsymbol{r},t)$. The scalar potential is the potential whose gradient yields the longitudinal electric field. Equation (A-20) then shows that, to within a constant, $U(\boldsymbol{r},t)$ is equal to:

$$U(\boldsymbol{r},t) = \frac{1}{4\pi\varepsilon_0} \sum_a q_a \frac{1}{|\boldsymbol{r} - \boldsymbol{r}_a(t)|} \tag{A-34}$$

which is the Coulomb potential created by the charge distribution.

Lorenz gauge

In the present chapter and the next one, we shall mainly use the Coulomb gauge. Another gauge often used, in particular in the clearly covariant formulations of electrodynamics, is the *Lorentz gauge*[5] defined by the condition:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r},t) + \frac{1}{c^2}\frac{\partial}{\partial t} U(\boldsymbol{r},t) = 0 \tag{A-35}$$

which can be written, using covariant notation:

$$\sum_{\mu} \partial_\mu A^\mu = 0 \tag{A-36}$$

The condition defining the Lorenz gauge thus keeps the same form in every Lorentz reference frame, which is not the case for the Coulomb gauge (since in relativity, a transverse field of zero divergence in one reference frame is no longer necessarily transverse in another frame). Nevertheless, an advantage of the Coulomb gauge is that it allows the immediate identification, in a given reference frame, of the field variables that are really independent.

A-3. Elimination of the longitudinal fields from the expression of the physical quantities

It will be useful for the following discussion to eliminate the longitudinal fields from the expressions of the total energy H and the total momentum given by equations (A-6a) and (A-6b). We shall express these physical quantities only in terms of the truly independent variables, such as particle coordinates and speeds, and transverse fields.

[5]The danish physicist Ludwig Lorenz is often confused with the dutch physicist Hendrik Lorentz.

A-3-a. Total energy

We start by eliminating the longitudinal electric field from the last term in expression (A-6a). Using the Parseval-Plancherel equality (A-12) and the fact that $\tilde{\boldsymbol{E}}^*_{\parallel}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=0$, we can rewrite this term as:

$$\frac{\varepsilon_0}{2}\int \mathrm{d}^3r\left[\boldsymbol{E}^2(\boldsymbol{r},t)+c^2\boldsymbol{B}^2(\boldsymbol{r},t)\right]=H_{\mathrm{long}}+H_{\mathrm{trans}} \tag{A-37}$$

where:

$$H_{\mathrm{long}}=\frac{\varepsilon_0}{2}\int \mathrm{d}^3k\;\tilde{\boldsymbol{E}}^*_{\parallel}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t) \tag{A-38a}$$

$$H_{\mathrm{trans}}=\frac{\varepsilon_0}{2}\int \mathrm{d}^3k\left[\tilde{\boldsymbol{E}}^*_{\perp}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)+c^2\tilde{\boldsymbol{B}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{B}}(\boldsymbol{k},t)\right] \tag{A-38b}$$

In (A-38a), we replace $\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)$ by expression (A-18a). We get, taking (A-12) and (A-13) into account:

$$\begin{aligned}H_{\mathrm{long}}&=\frac{1}{2\varepsilon_0}\int \mathrm{d}^3k\;\frac{\tilde{\rho}^*(\boldsymbol{k},t)\tilde{\rho}(\boldsymbol{k},t)}{k^2}\\&=\frac{1}{8\pi\varepsilon_0}\int\int \mathrm{d}^3r\,\mathrm{d}^3r'\;\frac{\tilde{\rho}(\boldsymbol{r},t)\tilde{\rho}(\boldsymbol{r}',t)}{|\boldsymbol{r}-\boldsymbol{r}'|}\\&=\sum_a h^a_{\mathrm{Coul}}+\frac{1}{8\pi\epsilon_0}\sum_{a\neq b}\frac{q_aq_b}{|\boldsymbol{r}_a-\boldsymbol{r}_b|}=V_{\mathrm{Coul}}\end{aligned} \tag{A-39}$$

The longitudinal field energy is thus equal to the Coulomb electrostatic energy V_{Coul} of the charge distribution $\rho(\boldsymbol{r},t)$. In addition to the Coulomb interaction energy between different particles a and b, V_{Coul} also contains the energy h^a_{Coul} of the Coulomb field of each particle a, which diverges for point particles.

Expression (A-38b) for H_{trans} can be rewritten as a function of the variables $\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=-\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)$ introduced above for the transverse field:

$$H_{\mathrm{trans}}=\frac{\varepsilon_0}{2}\int \mathrm{d}^3k\;\left[\dot{\tilde{\boldsymbol{A}}}^*_{\perp}(\boldsymbol{k},t)\cdot\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)+\omega^2\tilde{\boldsymbol{A}}^*_{\perp}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\right] \tag{A-40}$$

Finally, the energy of the global system field + particles can be expressed in the form:

$$H=\frac{1}{2}\sum_a m_a\dot{\boldsymbol{r}}^2_a(t)+V_{\mathrm{Coul}}+H_{\mathrm{trans}} \tag{A-41}$$

where we used the simplified notation $\dot{\boldsymbol{r}}_a(t)=\mathrm{d}\boldsymbol{r}_a(t)/\mathrm{d}t=\boldsymbol{v}_a(t)$. It is the sum of the kinetic energy of the particles, of their Coulomb energy, and of the energy of the transverse field.

A-3-b. Total momentum

Similar computations can be carried out for the total momentum $\boldsymbol{P}$. The field contribution contained in the last term of (A-6b) can be written as:

$$\varepsilon_0 \int \mathrm{d}^3k\ \tilde{\boldsymbol{E}}^*(\boldsymbol{k},t)\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t) = \underbrace{\varepsilon_0 \int \mathrm{d}^3k\ \tilde{\boldsymbol{E}}^*_{\parallel}(\boldsymbol{k},t)\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)}_{\boldsymbol{P}_{\mathrm{long}}} + \underbrace{\varepsilon_0 \int \mathrm{d}^3k\ \tilde{\boldsymbol{E}}^*_{\perp}(\boldsymbol{k},t)\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)}_{\boldsymbol{P}_{\mathrm{trans}}} \tag{A-42}$$

where we have separated the contributions to $\boldsymbol{P}$ coming from the longitudinal and transverse components of the electric field[6]. Using (A-18a) and (A-27b), taking into account identity (A-28) and the fact that $\boldsymbol{k}\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)=0$, we get:

$$\begin{aligned}\boldsymbol{P}_{\mathrm{long}} &= \varepsilon_0 \int \mathrm{d}^3k\ \frac{i\tilde{\rho}^*(\boldsymbol{k},t)}{\varepsilon_0}\frac{\boldsymbol{k}}{k^2}\times\big(i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\big)\\ &= \int \mathrm{d}^3k\ \tilde{\rho}^*(\boldsymbol{k},t)\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\end{aligned} \tag{A-43}$$

We then have:

$$\begin{aligned}\boldsymbol{P}_{\mathrm{long}} &= \int \mathrm{d}^3r\ \rho(\boldsymbol{r},t)\boldsymbol{A}_{\perp}(\boldsymbol{r},t)\\ &= \sum_a q_a\boldsymbol{A}_{\perp}(\boldsymbol{r}_a,t).\end{aligned} \tag{A-44}$$

As we did above for (A-40), we can rewrite the expression of $\boldsymbol{P}_{\mathrm{trans}}$ as a function of the variables $\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=-\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)$ of the transverse field:

$$\begin{aligned}\boldsymbol{P}_{\mathrm{trans}} &= -\varepsilon_0\int \mathrm{d}^3k\ \dot{\tilde{\boldsymbol{A}}}^*_{\perp}(\boldsymbol{k},t)\times\big[i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\big]\\ &= -i\varepsilon_0\int \mathrm{d}^3k\ \ \boldsymbol{k}\ \ \big[\dot{\tilde{\boldsymbol{A}}}^*_{\perp}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\big]\end{aligned} \tag{A-45}$$

The momentum of the global system field + particles can be written in the form:

$$\boldsymbol{P} = \sum_a \left[m_a\dot{\boldsymbol{r}}_a(t) + q_a\boldsymbol{A}_{\perp}(\boldsymbol{r}_a,t)\right] + \boldsymbol{P}_{\mathrm{trans}} \tag{A-46}$$

Let us finally introduce the quantity:

$$\boldsymbol{p}_a(t) = m_a\dot{\boldsymbol{r}}_a(t) + q_a\boldsymbol{A}_{\perp}(\boldsymbol{r}_a,t) \tag{A-47}$$

We shall see later that, in the Coulomb gauge electrodynamics, $\boldsymbol{p}_a(t)$ is the conjugate momentum of $\boldsymbol{r}_a(t)$, hence different from the mechanical momentum $m_a\dot{\boldsymbol{r}}_a(t)$. Expressed

[6]The notation $\boldsymbol{P}_{\mathrm{long}}$ should not lead us to believe that $\boldsymbol{P}_{\mathrm{long}}$ is a longitudinal field vector itself: it is actually the vector yielding the longitudinal electric field contribution to the momentum vector; the same comment applies to $\boldsymbol{P}_{\mathrm{trans}}$.

as a function of $\boldsymbol{p}_a(t)$, the total energy (A-41) and total momentum (A-46) are written as:

$$H = \frac{1}{2m_a}\sum_a \left[\boldsymbol{p}_a - q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a, t)\right]^2 + V_{\text{Coul}} + H_{\text{trans}} \tag{A-48}$$

$$\boldsymbol{P} = \sum_a \boldsymbol{p}_a(t) + \boldsymbol{P}_{\text{trans}} \tag{A-49}$$

where H_{trans} and $\boldsymbol{P}_{\text{trans}}$ were introduced in equations (A-38b) and (A-42). We shall see that H actually coincides with the Hamiltonian in the Coulomb gauge of the global system field + particles.

A-3-c. Total angular momentum

Calculations similar to ones just presented, but that will not be detailed here[7], show that the contribution of the longitudinal electric field to the total angular momentum is equal to:

$$\boldsymbol{J}_{\text{long}} = \varepsilon_0 \int \mathrm{d}^3 r\, \boldsymbol{r} \times (\boldsymbol{E}_\parallel \times \boldsymbol{B}) = \sum_a q_a \boldsymbol{r}_a \times \boldsymbol{A}_\perp(\boldsymbol{r}_a). \tag{A-50}$$

Adding $\boldsymbol{J}_{\text{long}}$ to the particles' angular momenta, we get, taking (A-47) into account:

$$\sum_a \boldsymbol{r}_a \times m_a \dot{\boldsymbol{r}}_a + \boldsymbol{J}_{\text{long}} = \sum_a \boldsymbol{r}_a \times \boldsymbol{p}_a \tag{A-51}$$

so that we can finally write:

$$\boldsymbol{J} = \sum_a \boldsymbol{r}_a \times \boldsymbol{p}_a + \boldsymbol{J}_{\text{trans}} \tag{A-52}$$

where:

$$\boldsymbol{J}_{\text{trans}} = \varepsilon_0 \int \mathrm{d}^3 r\, [\boldsymbol{r} \times (\boldsymbol{E}_\perp \times \boldsymbol{B})] \tag{A-53}$$

B. Describing the transverse field as an ensemble of harmonic oscillators

B-1. Brief review of the one-dimensional harmonic oscillator

The energy of a harmonic oscillator is given by:

$$E = \frac{1}{2} m \dot{x}^2 + \frac{1}{2} m \omega^2 x^2 \tag{B-1}$$

where $\omega/2\pi$ is the oscillation frequency, and $\dot{x}$ the oscillator velocity:

$$\frac{\mathrm{d}}{\mathrm{d}t}\, x = \dot{x} \tag{B-2}$$

[7] These calculations can be found in § 1 of Complement $\mathrm{B_I}$ in [16].

This velocity obeys:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\dot{x} = -\omega^2 x \tag{B-3}$$

so that the equation of motion of x is:

$$\ddot{x} + \omega^2 x = 0 \tag{B-4}$$

Consequently, the time evolution of $x(t)$ is given by a (real) linear combination of $\cos(\omega t)$ and $\sin(\omega t)$.

The dynamic state of the classical harmonic oscillator is defined at each instant by two real variables $x(t)$ and $\dot{x}(t)$. It is often useful to combine them into a single complex variable $\alpha(t)$ by setting:

$$\alpha(t) = C\left[x(t) + i\,\frac{\dot{x}(t)}{\omega}\right] \tag{B-5}$$

where C is an arbitrary (time-independent) constant. Relations (B-2) and (B-3) show that $\alpha(t)$ obeys the first order differential equation:

$$\dot{\alpha} = C(\dot{x} - i\omega x) = -i\omega C\left(x + i\frac{\dot{x}}{\omega}\right) = -i\omega\alpha \tag{B-6}$$

The time dependence of the new variable $\alpha(t)$ is therefore simply $e^{-i\omega t}$.

One can invert the system formed by equation (B-5) and its complex conjugate yielding α^*, and compute x and $\dot{x}$ as a function of α and α^*. Inserting the expressions thus obtained in equation (B-1) for the energy E, we obtain by a simple calculation[8]:

$$E = \frac{m\omega^2}{4C^2}(\alpha^*\alpha + \alpha\alpha^*) \tag{B-7}$$

The constant C can be chosen so that:

$$\frac{m\omega^2}{4C^2} = \frac{\hbar\omega}{2} \tag{B-8}$$

This leads, after quantization, to the Hamiltonian operator:

$$\hat{H} = \frac{\hbar\omega}{2}(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger) \tag{B-9}$$

which is the Hamiltonian of a harmonic oscillator[9].

B-2. Normal variables for the transverse field

B-2-a. Vibration eigenmodes of the free transverse field

In the reciprocal space, expression (A-40) for the free transverse field energy H_{trans} is a sum of quadratic functions of $\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$. For each value of $\boldsymbol{k}$, we get a harmonic oscillator Hamiltonian. The evolution introduces no coupling between the various spatial Fourier components of the transverse field. We see the advantage of

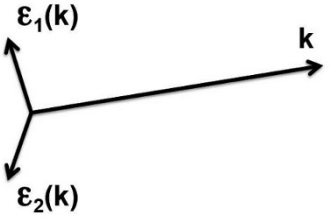


Figure 1: For each vector $\boldsymbol{k}$, the transverse fields can have two polarizations characterized by unit vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ perpendicular both to each other and to $\boldsymbol{k}$.

working in the reciprocal space: it enables us to identify the eigenmodes of the field vibrations, in the absence of sources.

Actually, for each $\boldsymbol{k}$, the transverse field can have two different polarizations [10] characterized by unit vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$, both perpendicular to $\boldsymbol{k}$ and to each other, so that we can write for $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$, as an example:

$$\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \tilde{A}_{\perp\boldsymbol{\varepsilon}_1(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_1(\boldsymbol{k}) + \tilde{A}_{\perp\boldsymbol{\varepsilon}_2(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_2(\boldsymbol{k}) = \sum_{\boldsymbol{\varepsilon}_i(\boldsymbol{k})} \tilde{A}_{\perp\boldsymbol{\varepsilon}_i(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_i(\boldsymbol{k}) \quad \text{(B-10)}$$

with:

$$\tilde{A}_{\perp\boldsymbol{\varepsilon}_i(\boldsymbol{k})}(\boldsymbol{k},t) = \boldsymbol{\varepsilon}_i(\boldsymbol{k}) \cdot \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \quad \text{(B-11)}$$

The set $\{\boldsymbol{k}, \boldsymbol{\varepsilon}_i(\boldsymbol{k})\}$ defines what we shall call in this chapter a free field *mode*; they are the eigenmodes of the free field vibration, with a frequency:

$$\omega = ck \quad \text{(B-12)}$$

To simplify the notation, we shall write the last summation in (B-10) in a more compact form:

$$\sum_{\boldsymbol{\varepsilon}_i(\boldsymbol{k})} \tilde{A}_{\perp\boldsymbol{\varepsilon}_i(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_i(\boldsymbol{k}) \equiv \sum_{\boldsymbol{\varepsilon}} \tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\boldsymbol{\varepsilon} \quad \text{(B-13)}$$

Let us rewrite expression (A-40) for $\tilde{H}_{\text{trans}}$ expliciting the components of the fields $\boldsymbol{A}_\perp(\boldsymbol{k},t)$ and $\dot{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ on the polarization vectors. We get:

$$H_{\text{trans}} = \frac{\varepsilon_0}{2} \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + \omega^2 \tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \right] \quad \text{(B-14)}$$

[8] In view of the quantization where α and α^* will be replaced by non-commuting operators $\hat{a}$ and $\hat{a}^\dagger$, we keep the sequence of α and α^* as they appear in the computations.

[9] If $\hat{x}$ and $\hat{p}$ obey the canonical commutation relation $[\hat{x},\hat{p}] = i\hbar$, relation (B-8) for the choice of C also leads to the commutation relation $[\hat{a},\hat{a}^\dagger] = 1$.

[10] We choose real vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ corresponding to linear polarizations, but the choice of these two polarizations is arbitrary, since they can always be rotated by any angle around $\boldsymbol{k}$. It is also possible to perform a more general change of basis with complex vectors defining elliptical polarizations, for instance the right and left circular polarizations $\boldsymbol{\varepsilon}_\pm = (\boldsymbol{\varepsilon}_1 \pm i\boldsymbol{\varepsilon}_2)/\sqrt{2}$. Circular polarizations are useful when discussing electromagnetic spin — see § 3 of Complement $\mathrm{B}_{\mathrm{XIX}}$. If complex (orthonormal) polarizations are used, $\boldsymbol{\varepsilon}_i(\boldsymbol{k})$ should be replaced by $\boldsymbol{\varepsilon}^*_i(\boldsymbol{k})$ in the right side hand of relation (B-11).

Note that the components on the two polarizations $\boldsymbol{\varepsilon}$ are truly independent dynamic variables (generalized coordinates and velocities). This is not the case for the Cartesian components $\tilde{A}_{\perp i}(\boldsymbol{k},t)$ and $\dot{\tilde{A}}_{\perp i}(\boldsymbol{k},t)$ (with $i=x,y,z$), because of the transversality condition. For example, the components $\tilde{A}_{\perp i}(\boldsymbol{k},t)$ must obey $\sum_i k_i \tilde{A}_{\perp i}=0$.

Constraints on the dynamic variables in the reciprocal space

Since the fields are real in real space, we have the condition $\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t)$. In half the reciprocal space, the variables $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ can be considered as independent .

B-2-b. Definition of the normal variables, free field case

Let us first assume that we are in the free field case ($\tilde{\boldsymbol{j}}_\perp = \boldsymbol{0}$), and we can replace the field $\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)$ by $-\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$ in equations (A-30a) and (A-30b). As $\omega = kc$, we get two equations exactly similar to those of a harmonic oscillator (B-2) and (B-3), with $\boldsymbol{A}_\perp(\boldsymbol{k},t)$ instead of $x(t)$. This analogy suggests introducing, as in (B-5), a new transverse variable:

$$\begin{aligned}\boldsymbol{\alpha}(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) + \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)\right] \\ &= i\mathcal{N}(k)\left[\frac{\boldsymbol{k}}{k^2}\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t) - \frac{1}{\omega}\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)\right]\end{aligned} \tag{B-15}$$

where $\mathcal{N}(k)$ is a real constant, not yet specified, which can depend on k (its value will be chosen at the beginning of the next chapter). This definition, together with (A-30b), yields the equation of motion for $\boldsymbol{\alpha}(\boldsymbol{k},t)$:

$$\dot{\boldsymbol{\alpha}}(\boldsymbol{k},t) + i\omega\boldsymbol{\alpha}(\boldsymbol{k},t) = 0 \tag{B-16}$$

As opposed to $\boldsymbol{A}_\perp(\boldsymbol{k},t)$ that, according to (A-31), obeys a second order equation, this new variable $\boldsymbol{\alpha}(\boldsymbol{k},t)$ obeys a first order equation. It is a complex variable whose time evolution is proportional to $e^{-i\omega t}$, and not, as is the case for the variable $\boldsymbol{A}_\perp(\boldsymbol{k},t)$, to a linear superposition of $e^{-i\omega t}$ and $e^{+i\omega t}$. It will be useful in what follows to consider the complex conjugate of equation (B-15):

$$\begin{aligned}\boldsymbol{\alpha}^*(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}^*_\perp(\boldsymbol{k},t)\right] \\ &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(-\boldsymbol{k},t)\right]\end{aligned} \tag{B-17}$$

To go from the first to the second line of (B-17), we used the fact that $\boldsymbol{A}_\perp$ is real in the real space, which leads to:

$$\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t) \tag{B-18}$$

A similar relation exists for $\dot{\tilde{\boldsymbol{A}}}_\perp$. The transverse variables $\boldsymbol{\alpha}(\boldsymbol{k},t)$ and $\boldsymbol{\alpha}^*(\boldsymbol{k},t)$ are called the transverse field *normal variables*. We will see in the next chapter that the quantization process will transform these variables into annihilation and creation operators of photons.

B-2-c. Equation of motion for the normal variables in the presence of sources

In the presence of sources, $\tilde{\boldsymbol{j}}_\perp$ is no longer zero. We can still define the normal variables $\boldsymbol{\alpha}(\boldsymbol{k},t)$ by relations (B-15), but we must now keep the term in $\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t)$ on the right-hand side of equation (A-30b). The same transformation that led us from equations (A-30a) and (A-30b) to (B-16) now yields a new equation of motion in the presence of sources:

$$\dot{\boldsymbol{\alpha}}(\boldsymbol{k},t) + i\omega\boldsymbol{\alpha}(\boldsymbol{k},t) = \frac{i\mathcal{N}(k)}{\varepsilon_0\ \omega}\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) \tag{B-19}$$

This equation is strictly equivalent to Maxwell's equations for the transverse fields. One can see this by taking the time derivative of equations (B-22a) and (B-22b) given below, and using (B-19) to get the time-dependent evolution equations (A-30a) and (A-30b) of these fields.

Independence of the normal variables

Another interest of the normal variables is that they are independent: there is no relation between $\boldsymbol{\alpha}(\boldsymbol{k},t)$ and $\boldsymbol{\alpha}^*(-\boldsymbol{k},t)$ such as the one that exists between $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}^*_\perp(-\boldsymbol{k},t)$. This is because the real and imaginary parts of $\boldsymbol{\alpha}(\boldsymbol{k},t)$ depend on two independent degrees of freedom, $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and its time derivative. It is easy to check, by changing the sign of $\boldsymbol{k}$ in (B-15) and by using (B-18) that:

$$\boldsymbol{\alpha}(-\boldsymbol{k},t) = \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) + \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}^*_\perp(\boldsymbol{k},t)\right] \neq \boldsymbol{\alpha}^*(\boldsymbol{k},t) \tag{B-20}$$

The knowledge of the $\boldsymbol{\alpha}(\boldsymbol{k},t)$ in the entire reciprocal space does not entail the knowledge of the $\boldsymbol{\alpha}^*(\boldsymbol{k},t)$. Consequently, the integrals over $\boldsymbol{k}$ of the normal variables must be taken over the entire space, and not be limited to half the reciprocal space.

B-2-d. Expression of the physical quantities in terms of the normal variables

We are going to show that all the physical quantities can be expressed in terms of the normal variables.

α. Transverse fields in the reciprocal space

Replacing $\boldsymbol{k}$ by $-\boldsymbol{k}$, we can rewrite equation (B-17) as:

$$\boldsymbol{\alpha}^*(-\boldsymbol{k},t) = \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)\right] \tag{B-21}$$

Using (B-15) and (B-21), we can now express $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and $\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$ as a function of $\boldsymbol{\alpha}(\boldsymbol{k},t)$ and $\boldsymbol{\alpha}^*(-\boldsymbol{k},t)$. We get:

$$\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \frac{1}{2\mathcal{N}(k)}\left[\boldsymbol{\alpha}(\boldsymbol{k},t) + \boldsymbol{\alpha}^*(-\boldsymbol{k},t)\right] \tag{B-22a}$$

$$\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t) = -i\frac{\omega}{2\mathcal{N}(k)}\left[\boldsymbol{\alpha}(\boldsymbol{k},t) - \boldsymbol{\alpha}^*(-\boldsymbol{k},t)\right] \tag{B-22b}$$

β. Energy and momentum of the transverse field

We insert relations (B-22a) and (B-22b) for $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and $\dot{\tilde{A}}_\perp(\boldsymbol{k},t)$ in the expression (A-40) for H_{trans}, using the more compact notation:

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}(\boldsymbol{k},t) \quad ; \quad \boldsymbol{\alpha}_-=\boldsymbol{\alpha}(-\boldsymbol{k},t) \tag{B-23}$$

We get:

$$\begin{aligned} H_{\text{trans}} &= \frac{\varepsilon_0}{2}\int \mathrm{d}^3k\ \frac{\omega^2}{4\mathcal{N}^2(k)}\left[(\boldsymbol{\alpha}^*-\boldsymbol{\alpha}_-)\cdot(\boldsymbol{\alpha}-\boldsymbol{\alpha}_-^*)+(\boldsymbol{\alpha}^*+\boldsymbol{\alpha}_-)\cdot(\boldsymbol{\alpha}+\boldsymbol{\alpha}_-^*)\right] \\ &= \frac{\varepsilon_0}{2}\int \mathrm{d}^3k\ \frac{\omega^2}{4\mathcal{N}^2(k)}\left[2\boldsymbol{\alpha}^*\cdot\boldsymbol{\alpha}+2\boldsymbol{\alpha}_-\cdot\boldsymbol{\alpha}_-^*\right] \end{aligned} \tag{B-24}$$

(in these equations, we keep the ordered sequence of $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ as they appear in the computations, even though $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ are commuting numbers; the reason is that similar computations can be carried out in the quantum theory where $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ will be replaced by non-commuting operators). A change of variable $\boldsymbol{k}\to-\boldsymbol{k}$ in the integral of the terms in $\boldsymbol{\alpha}_-\cdot\boldsymbol{\alpha}_-^*$ yields an integral of $\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^*$. We then get:

$$H_{\text{trans}}=\varepsilon_0\int \mathrm{d}^3k\ \frac{\omega^2}{4\mathcal{N}^2(k)}\left[\boldsymbol{\alpha}^*\cdot\boldsymbol{\alpha}+\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^*\right] \tag{B-25}$$

Expliciting the components of $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ on the two polarization vectors $\boldsymbol{\varepsilon}$ perpendicular to $\boldsymbol{k}$, and using the simplified notation (B-13), we finally get:

$$H_{\text{trans}}=\varepsilon_0\int \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\frac{\omega^2}{4\mathcal{N}^2(k)}\left[\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\right] \tag{B-26}$$

This expression looks a lot like a sum of harmonic oscillator Hamiltonians; a suitable choice for the constant $\mathcal{N}$ will be made in the next chapter.

Similar calculations can be carried out for the transverse field momentum $\boldsymbol{P}_{\text{trans}}$[11]. Using equations (A-45), (B-22a) and (B-22b), we get:

$$\boldsymbol{P}_{\text{trans}}=\varepsilon_0\int \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\frac{\omega}{4\mathcal{N}^2(k)}\ \boldsymbol{k}\ \left[\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\right] \tag{B-27}$$

γ. Transverse fields in real space

Let us consider first the transverse potential vector $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$, whose expression in terms of the normal variables is given by (B-22a). To get its expression in real space, one must, taking (A-11) into account, multiply (B-22a) by $(2\pi)^{-3/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ and integrate over $\boldsymbol{k}$. Making the change of variable $\boldsymbol{k}\to-\boldsymbol{k}$ in the integral containing $\boldsymbol{\alpha}^*(-\boldsymbol{k},t)$, we finally get:

$$\boldsymbol{A}_\perp(\boldsymbol{r},t)=\frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\frac{1}{2\mathcal{N}(k)}\left[\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}\ e^{i\boldsymbol{k}\cdot\boldsymbol{r}}+\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}^*\ e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{B-28}$$

[11] The expression of the angular momentum $\boldsymbol{J}_{\text{trans}}$ of the transverse field, in terms of the normal variables, will be computed in Complement $\mathrm{B}_{\mathrm{XIX}}$.

This relation (as well as the next two equations) is written in the general case where the polarizations may be complex, elliptical or circular (*cf.* note 10). This is why the term in $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ contains a complex conjugate polarization $\boldsymbol{\varepsilon}^*$.

Similar calculations can be carried out for the transverse electric field as well as for the magnetic field. They yield:

$$\boldsymbol{E}_\perp(\boldsymbol{r},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{i\omega}{2\mathcal{N}(k)} \left[\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{B-29}$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{ik}{2\mathcal{N}(k)} \left[\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{B-30}$$

where $\boldsymbol{\kappa}$ has been defined in (A-16) as the unit vector parallel to $\boldsymbol{k}$.

B-3. Discrete modes in a box

So far, we have considered radiation propagating in an infinite space and used continuous Fourier transforms; in relation (A-11), the electric field is expanded on a continuous basis of normalized plane waves $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}/(2\pi)^{3/2}$. It is often useful, however, to use a discrete basis, assuming the radiation to be contained in a box of finite volume, generally defined as a cube of edge length L; this will frequently occur in the next two chapters when dealing with the quantized radiation. The components of each wave vector must obey the boundary conditions in the box[12], and hence take on discrete values:

$$k_{x,y,z} = 2\pi\ n_{x,y,z}/L \tag{B-31}$$

At the end of the computation, nothing prevents us from choosing a very large value of L in order to check that the final result does not depend on L.

Instead of continuous spatial Fourier transforms, one must now introduce discrete Fourier series where each physical quantity is expanded in terms of normalized plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}/L^{3/2}$. The expansion (A-11) of the electric field then becomes:

$$\boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{L^{3/2}} \sum_{\mathbf{k}} \tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{B-32}$$

with[13]:

$$\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t) = \frac{1}{L^{3/2}} \int_{\mathcal{V}} \mathrm{d}^3r\, \boldsymbol{E}(\boldsymbol{r},t) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{B-33}$$

The summation in (B-32) is discrete, and the integral in (B-33) is now limited to the volume $\mathcal{V}$ of the box.

[12] One can choose to impose the field being zero on the walls, but it is generally easier to enforce periodic boundary conditions (B-31), which lead to the same k density of states.

[13] In Appendix I, we used a slightly different definition for the Fourier series, with which the factor $1/L^{3/2}$ would be missing from (B-32), but where (B-33) would contain a factor $1/L^3$. The definition we use here is chosen to directly yield an expansion of $\boldsymbol{E}(\boldsymbol{r},t)$ on plane waves normalized in the cube.

Note that if the field is zero outside the box, it is obviously possible to use the continuous Fourier transform (A-10) to get the field component $\tilde{\boldsymbol{E}}(\boldsymbol{k},t)$; however, this latter component is different from the discrete component $\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t)$, because of the coefficients introduced in the definitions. The two components are related by:

$$\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t) = \left(\frac{2\pi}{L}\right)^{3/2} \tilde{\boldsymbol{E}}(\boldsymbol{k},t) \tag{B-34}$$

The same changes can be made on the Fourier transforms of all the other physical quantities such as the magnetic field, the vector potential, as well as the charge and current densities. The equations in the reciprocal space such as equations (A-17), (A-21), (A-25) and the following, remain valid if we replace the continuous variables $\mathbf{k}$ by discrete variables, since each side of those equations are multiplied by the same factor $(2\pi)^{3/2}\, L^{-3/2}$. In the case of a zero field outside the box, the $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ are also replaced by :

$$\alpha_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t) = \left(\frac{2\pi}{L}\right)^{3/2} \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{B-35}$$

Coming back to ordinary space $(\boldsymbol{r},t)$ via the inverse Fourier transform, we must use relations of the type (B-32) instead of (A-11). Consequently, once we replace in the integral over d^3k the $\tilde{\boldsymbol{E}}_(\boldsymbol{k},t)$ by the $\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t)$, we must also introduce a multiplicative factor[14]:

$$\int \mathrm{d}^3k \implies \left(\frac{2\pi}{L}\right)^{3/2} \sum_{\mathbf{k}} \tag{B-36}$$

B-4. Generalization of the mode concept

In the absence of sources, the solution of the equation of motion (B-16) for the normal variable $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ is very simple, since it is an exponential with an angular frequency $\omega = ck$:

$$\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},0)e^{-i\omega t} \tag{B-37}$$

Inserting (B-37) in the expressions we just obtained for the transverse fields and the other physical quantities, we see that the fields are linear superpositions of progressive plane waves, propagating independently of each other. The free field energy and momentum are the sum of the squared moduli of the various normal variables, each being time-independent and proportional to $|\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},0)|^2$.

The modes $\{\boldsymbol{k},\boldsymbol{\varepsilon}\}$ introduced in this chapter permit expanding the free transverse fields on progressive plane waves. Nevertheless, other expansions on monochromatic waves that are not necessarily plane waves are also possible; they involve other families of modes, as we are now show, coming back to equation (A-31). In the absence of sources, any monochromatic solution of this equation, of the form $\boldsymbol{A}_{\perp}^{(+)}(\boldsymbol{r})e^{-i\omega t}$, necessarily obeys equation:

$$(\Delta + k^2)\boldsymbol{A}_{\perp}^{(+)}(\boldsymbol{r}) = 0 \tag{B-38}$$

[14]The product of the multiplicative factor of (B-34) and that of (B-35) yields the usual factor $(2\pi/L)^3$, obtained directly from (B-31).

(which is simply the Helmholtz equation) with $k = \omega/c$. The plane waves $e^{\pm i\boldsymbol{k}\cdot\boldsymbol{r}}$ are a possible basis of eigenfunctions for this eigenvalue equation, but not the only one. There exists other bases, such as the basis of stationary waves $\cos \boldsymbol{k}\cdot\boldsymbol{r}$ and $\sin \boldsymbol{k}\cdot\boldsymbol{r}$, the basis of multipolar waves (radiation modes with a specific angular momentum, whereas plane waves have a specific linear momentum), or the basis corresponding to Gaussian modes. More generally, any linear combination of plane waves with the same modulus $\boldsymbol{k}$ can become a mode. Whatever basis is chosen, the transverse field energy will be a sum of the squared moduli of normal variables introduced in the expansions of the transverse fields on the eigenfunctions of that basis. The expression of the other physical quantities, however, will only have a simple form in a particular basis. As an example, the momentum of the transverse field is a sum of squared moduli only in the basis of progressive plane waves, whereas the field angular momentum has a simple form only in the basis of multipolar waves.

Note finally that the field can be contained in a cavity with well defined boundary conditions. Finding the eigenfunctions of equation (B-38) obeying these boundary conditions is a way to determine the eigenmodes of this cavity.

To conclude this chapter, we can say that the free radiation field is equivalent to an ensemble of one-dimensional harmonic oscillators associated with the modes $\{\boldsymbol{k}, \boldsymbol{\varepsilon}\}$ labeled by their wave vector and their transverse polarization. Each mode is associated with a field normal variable, similar to the classical variable of the corresponding classical oscillator, and which will become, in the quantization process, the oscillator annihilation operator. The results established in this chapter will be the simple starting point for the radiation quantization explained in the next chapter.

COMPLEMENT OF CHAPTER XVIII, READER'S GUIDE

A_{XVIII} : LAGRANGIAN FORMULATION OF ELECTRODYNAMICS

The dynamic equations for the electrodynamic field (Maxwell's equations) can be obtained from the Lagrangian formalism based on a principle of least action. This enables introducing expressions for the conjugate momenta of the various field variables, as well as for the field Hamiltonian when coupled to charged particles. The results of this complement are not indispensable for reading the other chapters and complements. They offer, however, an overview of a more general approach to quantum electrodynamics, which is essential for a relativistic treatment of these problems and for the use of path integrals (Appendix IV).

Complement A_{XVIII}

Lagrangian formulation of electrodynamics

1 Lagrangian with several types of variables 1980
1-a Lagrangian formalism with discrete and real variables 1980
1-b Extension to complex variables 1982
1-c Lagrangian with continuous variables 1984
2 Application to the free radiation field 1986
2-a Lagrangian densities in real and reciprocal spaces 1986
2-b Lagrange's equations 1987
2-c Conjugate momentum of the transverse potential vector . . . 1987
2-d Hamiltonian; Hamilton-Jacobi equations 1988
2-e Field commutation relations 1989
2-f Creation and annihilation operators 1990
2-g Discrete momentum variables 1991
3 Lagrangian of the global system field + interacting particles 1992
3-a Choice for the Lagrangian 1992
3-b Lagrange's equations 1993
3-c Conjugate momenta 1995
3-d Hamiltonian 1995
3-e Commutation relations 1996

Introduction

As shown in Appendix III, the dynamics of a system of point particles in an external potential can be described either by Newton's equations, or by a Lagrangian with the principle of least action leading to Lagrange's equations, equivalent to Newton's equations. An advantage of the Lagrangian formalism is that it facilitates the quantization of the theory: it directly leads to the definition of the conjugate momenta of the particles' coordinates, and of the system's Hamiltonian, which is a function of the coordinates and the conjugate momenta. It then naturally introduces the canonical commutation relations, fundamental for the quantum description of the system. This complement will show, in a succinct way, how the Maxwell-Lorentz equations, studied in this chapter and the next, can be deduced from a Lagrangian and a principle of least action. This will give a more general justification for the expression of the Hamiltonian of the system "field + particles" postulated in Chapter XIX and for the commutation relations also postulated in that chapter[1]. Another advantage of the Lagrangian formalism, that we shall not exploit here, is that it is well suited to a relativistic description of the system "field + particles" which is why it is used in the quantum theory of relativistic fields.

[1]The relations postulated in Chapter XIX are justified a posteriori by the fact that they lead to the correct Heisenberg equations for the quantum operators associated with the particles and the fields.

We start in § 1 by extending the computations of Appendix III to the case where the system coordinates are complex and not real, even though the Lagrangian remains a real quantity. We will also show that the principle of least action and Lagrange's equations can be generalized to the field case, that is to a case where the system coordinates no longer depend on a discrete but on a continuous index, such as the point $\boldsymbol{r}$ in real space.

The discussion will be illustrated in § 2, which studies the Lagrangian of the free radiation field in the absence of sources. The field will be described by its components in reciprocal space, which are complex quantities. The Lagrangian then depends only on the field components and their time derivatives, hence making the computations easier than if the fields were described by the real components in real space (this is because the Lagrangian in real space depends not only on the field components and their time derivatives, but also on their spatial derivatives). In this study, we shall establish the expression for the field Hamiltonian, and the canonical commutation relations of the components of that field.

Finally, we give in § 3 the expression for the electrodynamic Lagrangian in the Coulomb gauge in the presence of sources; we show how Lagrange's equations deduced from this Lagrangian coincide with the Maxwell-Lorentz equations studied in Chapter XVIII. Several important relations for the quantization of the theory will be established: expression for the conjugate momenta of the particles and fields; expression for the Hamiltonian of the global system field + particles; canonical commutation relations. The results obtained in this complement give a base for the quantization process more general than the simplified approach of Chapitre XIX. The interested reader can find a more detailed description of the electrodynamic Lagrangian and Hamiltonian formalism in Chapter II of reference [16] and its complements.

1. Lagrangian with several types of variables

1-a. Lagrangian formalism with discrete and real variables

The Lagrangian L is a real function of dynamical variables composed of "generalized coordinates" $x_i(t)$ labeled by a discrete index i and of the corresponding "generalized velocities" $\dot{x}_i(t) = \mathrm{d}x_i(t)/\mathrm{d}t$. L is written:

$$L\left[x_1(t), x_2(t), ...x_N(t); \dot{x}_1(t), \dot{x}_2(t), ...\dot{x}_N(t)\right] \tag{1}$$

Consider a possible motion of the system where the coordinates $x_i(t)$ follow a certain "path" Γ between an initial time t_{in} and a final time t_2. The integral of L along the path Γ is, by definition, *the action* S_Γ associated with this path:

$$S_\Gamma = \int_{t_1}^{t_2} \mathrm{d}t \; L\left[x_1(t), x_2(t), ...x_N(t); \dot{x}_1(t), \dot{x}_2(t), ...\dot{x}_N(t)\right] \tag{2}$$

The *principle of least action* postulates that, among all the possible paths starting from the same initial conditions described by $x_i(t_{\text{in}})$ and arriving at the same final conditions described by $x_i(t_2)$, the system will follow the one for which S_Γ presents an extremum (if the path varies, S_Γ is stationary). Consider an infinitesimal variation $\delta x_i(t)$ and $\delta \dot{x}_i(t)$ of the dynamical variables around this path of extremum action, which does not change the initial and final values of the coordinates, i.e. such that:

$$\delta x_i(t_{\text{in}}) = \delta x_i(t_2) = 0 \tag{3}$$

The corresponding variation of the action

$$\delta S = \int_{t_1}^{t_2} \sum_i \left[\delta x_i(t) \frac{\partial L}{\partial x_i} + \delta \dot{x}_i(t) \frac{\partial L}{\partial \dot{x}_i} \right] \mathrm{d}t \tag{4}$$

must be zero to first order in $\delta x_i(t)$ and $\delta \dot{x}_i(t)$. We replace, in the last term of (4), $\delta \dot{x}_i(t)$ by:

$$\delta \dot{x}_i(t) = \frac{\mathrm{d}}{\mathrm{d}t} \delta x_i(t) \tag{5}$$

and integrate by parts the corresponding term. The integrated part is zero because of (3). We then get:

$$\delta S = \int_{t_1}^{t_2} \sum_i \delta x_i(t) \left[\frac{\partial L}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}_i} \right] \mathrm{d}t \tag{6}$$

As δS must be zero for any variation $\delta x_i(t)$, the path actually followed by the system must obey the N equations:

$$\frac{\partial L}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}_i} = 0 \tag{7}$$

These relations are called *Lagrange's equations*; they can be shown to be equivalent to Newton's equations (Appendix III).

The next step of the Lagrangian formalism is to introduce the *conjugate momenta* p_i of the coordinates x_i, defined by the equations:

$$p_i = \frac{\partial L}{\partial \dot{x}_i} \tag{8}$$

as well as the *Hamiltonian* H equal to:

$$H = \sum_i \dot{x}_i \, p_i - L \tag{9}$$

Let us take the differential of H

$$\begin{aligned} \mathrm{d}H &= \sum_i \left[p_i \mathrm{d}\dot{x}_1 + \dot{x}_i \mathrm{d}p_i - \frac{\partial L}{\partial x_i} \mathrm{d}x_i - \frac{\partial L}{\partial \dot{x}_1} \mathrm{d}\dot{x}_i \right] \\ &= \sum_i \left[\dot{x}_i \mathrm{d}p_i - \dot{p}_i \mathrm{d}x_i \right] \end{aligned} \tag{10}$$

To go from the first to the second line of (10), we used (7) and (8) to replace $\partial L/\partial x_i$ by $\dot{p}_i$ and $\partial L/\partial \dot{x}_i$ by p_i. We assume the $\dot{x}_i$ can be expressed as a function of the x_i and the p_i. The Hamiltonian H is then a function of the coordinates x_i and the conjugate momenta p_i, whose evolution obeys, taking (10) into account, the $2N$ equations:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i} \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial x_i} \tag{11}$$

called the *Hamilton-Jacobi equations*.

Let us finally recall the canonical quantization process. One associates with the coordinates x_i and the conjugate momenta p_i the operators $\hat{x}_i$ and $\hat{p}_i$ obeying the commutation relations:

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{12}$$

all the other commutators being equal to zero. These results are valid only if the coordinates x_i are Cartesian components (see comment in § B-5 of Chapter III).

1-b. Extension to complex variables

Let us assume, to keep things simple, that the index i takes only $N = 2$ values. With the two real coordinates $x_1(t)$ and $x_2(t)$ we build the complex variables:

$$X(t) = \frac{1}{\sqrt{2}}[x_1(t) + ix_2(t)] \qquad X^*(t) = \frac{1}{\sqrt{2}}[x_1(t) - ix_2(t)] \tag{13}$$

whose real and imaginary parts are, within a factor $1/\sqrt{2}$, equal to $x_1(t)$ and $x_2(t)$ for $X(t)$, $x_1(t)$ and $-x_2(t)$ for $X^*(t)$. Equations (13) can be inverted and yield:

$$x_1(t) = \frac{1}{\sqrt{2}}[X(t) + X^*(t)] \qquad x_2(t) = -\frac{i}{\sqrt{2}}[X(t) - X^*(t)] \tag{14}$$

Analogous equations can be written, relating $\dot{X}(t)$ and $\dot{X}^*(t)$ to $\dot{x}_1(t)$ and $\dot{x}_2(t)$ and vice versa:

$$\dot{X}(t) = \frac{1}{\sqrt{2}}[\dot{x}_1(t) + i\dot{x}_2(t)] \qquad \dot{X}^*(t) = \frac{1}{\sqrt{2}}[\dot{x}_1(t) - i\dot{x}_2(t)] \tag{15}$$

$$\dot{x}_1(t) = \frac{1}{\sqrt{2}}\left[\dot{X}(t) + \dot{X}^*(t)\right] \qquad \dot{x}_2(t) = -\frac{i}{\sqrt{2}}\left[\dot{X}(t) - \dot{X}^*(t)\right] \tag{16}$$

Inserting in the Lagrangian (1) expressions (14) and (16) for the variables, we get a Lagrangian of the form $L\left[X(t), X^*(t), \dot{X}(t), \dot{X}^*(t)\right]$ that depends on complex variables. Note however that just as in (1), this Lagrangian, even though it depends on complex variables, is still a real quantity since its time integral along a path Γ is an action (which is a real quantity). We now study what becomes of all the results established earlier using (1) when they are expressed as a function of $X, X^*, \dot{X}$ and $\dot{X}^*$.

α. *Lagrange's equations*

It is important, for what follows, to relate $\partial L/\partial X$ and $\partial L/\partial \dot{X}$ to $\partial L/\partial x_1$, $\partial L/\partial x_2$, $\partial L/\partial \dot{x}_1$ and $\partial L/\partial \dot{x}_2$. Using (14) and (16), we can write:

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial x_1}\frac{\partial x_1}{\partial X} + \frac{\partial L}{\partial x_2}\frac{\partial x_2}{\partial X} = \frac{1}{\sqrt{2}}\left[\frac{\partial L}{\partial x_1} - i\frac{\partial L}{\partial x_2}\right] \tag{17}$$

$$\frac{\partial L}{\partial \dot{X}} = \frac{\partial L}{\partial \dot{x}_1}\frac{\partial \dot{x}_1}{\partial \dot{X}} + \frac{\partial L}{\partial \dot{x}_2}\frac{\partial \dot{x}_2}{\partial \dot{X}} = \frac{1}{\sqrt{2}}\left[\frac{\partial L}{\partial \dot{x}_1} - i\frac{\partial L}{\partial \dot{x}_2}\right] \tag{18}$$

Subtracting from equation (17) the time derivative of equation (18), we get:

$$\frac{\partial L}{\partial X} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{X}} = \frac{1}{\sqrt{2}}\left(\frac{\partial L}{\partial x_1} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_1}\right) - \frac{i}{\sqrt{2}}\left(\frac{\partial L}{\partial x_2} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_2}\right) \tag{19}$$

The two parentheses on the right-hand side of this equation are zero since x_1 and x_2 obey Lagrange's equation (7). It then follows that the left-hand side is also zero, as is its complex conjugate[2] :

$$\frac{\partial L}{\partial X} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{X}} = 0 \qquad \frac{\partial L}{\partial X^*} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{X}^*} = 0 \tag{20}$$

which proves that X and X^* also obey Lagrange's equations[3].

β. Conjugate momenta

In (18) we replace $\partial L/\partial \dot{x}_1$ and $\partial L/\partial \dot{x}_2$ by p_1 and p_2 (see Eq. (8)). We get:

$$\frac{\partial L}{\partial \dot{X}} = \frac{1}{\sqrt{2}}(p_1 - ip_2) \tag{21}$$

The complex conjugate of equation (21) is written:

$$\frac{\partial L}{\partial \dot{X}^*} = \frac{1}{\sqrt{2}}(p_1 + ip_2) \tag{22}$$

To choose the definition of the conjugate momentum P of the complex variable X, it is useful to compare the way the conjugate momentum P and the velocity $\dot{X}$ are transformed upon the change of dynamical variables $x_1, x_2 \longrightarrow X, X^*$. We compare the first equation (15) and the two equations (21) and (22). The velocity $\dot{X}$ becomes $\dot{x}_1 + i\dot{x}_2$; a wise choice would be to define the associate momentum P in such a way that its transformation yields $p_1 + ip_2$. Equations (21) and (22) then clearly show that P must not be defined as $\partial L/\partial \dot{X}$, but rather as $\partial L/\partial \dot{X}^*$; the complex conjugate P^* is then equal to $\partial L/\partial \dot{X}$:

$$P = \frac{\partial L}{\partial \dot{X}^*} \qquad P^* = \frac{\partial L}{\partial \dot{X}} \tag{23}$$

This is the definition we shall use in the rest of this complement.

γ. Hamiltonian

The quantity $\dot{x}_1 p_1 + \dot{x}_2 p_2$ appears in the definition (9) of the Hamiltonian. This quantity can be rewritten by replacing $\dot{x}_1$ and $\dot{x}_2$ by their expressions (16) as a function of $\dot{X}$ and $\dot{X}^*$, as well as p_1 and p_2 by analogous expressions as a function of P and P^*:

$$\begin{aligned} \dot{x}_1 p_1 + \dot{x}_2 p_2 &= \frac{1}{\sqrt{2}}(\dot{X} + \dot{X}^*)\frac{1}{\sqrt{2}}(P + P^*) + \frac{-i}{\sqrt{2}}(\dot{X} - \dot{X}^*)\frac{-i}{\sqrt{2}}(P - P^*) \\ &= \dot{X}P^* + \dot{X}^*P \end{aligned} \tag{24}$$

[2]Since L is real, we have $(\partial L/\partial X)^* = \partial L/\partial X^*$ and an analogous equation for $\left(\partial L/\partial \dot{X}\right)^*$.

[3]These results could have been obtained directly by the variational calculation leading to (6), considering δX and δX^* to be independent variables.

We can then write:

$$H = \dot{x}_1 p_1 + \dot{x}_2 p_2 - L = \dot{X} P^* + \dot{X}^* P - L \tag{25}$$

We now take the differential of H:

$$\mathrm{d}H = \dot{X}\mathrm{d}P^* + P^*\mathrm{d}\dot{X} + \dot{X}^*\mathrm{d}P + P\mathrm{d}\dot{X}^* - \frac{\partial L}{\partial X}\mathrm{d}X - \frac{\partial L}{\partial X^*}\mathrm{d}X^* - \frac{\partial L}{\partial \dot{X}}\mathrm{d}\dot{X} - \frac{\partial L}{\partial \dot{X}^*}\mathrm{d}\dot{X}^* \tag{26}$$

Using (20) and (23), we get:

$$\mathrm{d}H = \dot{X}^*\mathrm{d}P - \dot{P}^*\mathrm{d}X + \dot{X}\mathrm{d}P^* - \dot{P}\mathrm{d}X^* \tag{27}$$

If $\dot{X}$ and $\dot{X}^*$ can be expressed in terms of the variables X, X^*, P, P^*, the Hamiltonian H only depends on those variables and we deduce from (27) the Hamilton-Jacobi equations:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\partial H}{\partial P^*} \qquad\qquad \frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial H}{\partial X^*} \tag{28}$$

and the complex conjugate equations for X^* and P^*. Note that it is the partial derivative with respect to P^* (and not with respect to P) that is equal to the total derivative of X.

δ. Canonical commutation relations

Upon quantization, the different variables become operators $\hat{x}_1$, $\hat{x}_2$, $\hat{p}_1$, $\hat{p}_2$, $\hat{X}$, $\hat{P}$, $\hat{X}^\dagger$, $\hat{P}^\dagger$. The commutation relations between the operators $\hat{X}$, $\hat{P}$, $\hat{X}^\dagger$, $\hat{P}^\dagger$ are obtained by expressing those operators in terms of $\hat{x}_1$, $\hat{x}_2$, $\hat{p}_1$, $\hat{p}_2$ and using the commutation relations (12). We can easily check that the only non-zero commutators are $[\hat{X}, \hat{P}^\dagger]$ and $[\hat{P}, \hat{X}^\dagger] = -[\hat{X}, \hat{P}^\dagger]$. Using (21), (22) and (23), we obtain:

$$\begin{aligned}[\hat{X}, \hat{P}^\dagger] &= \frac{1}{2}\left[\hat{x}_1 + i\hat{x}_2, \hat{p}_1 - i\hat{p}_2\right] \\ &= \frac{1}{2}\left([\hat{x}_1, \hat{p}_1] + [\hat{x}_2, \hat{p}_2]\right) = i\hbar\end{aligned} \tag{29}$$

1-c. Lagrangian with continuous variables

We now assume that the dynamical variables of the system depend on a continuous index, such as the point $\boldsymbol{r}$ in real space (or the point $\boldsymbol{k}$ in the reciprocal space when the real space is infinite). In other words, they constitute a field $\phi_j(\boldsymbol{r})$, where the discrete index j labels the component of the field if we are dealing with a vector field; in the reciprocal space, this field becomes $\tilde{\phi}_j(\boldsymbol{k})$. We shall only establish here Lagrange's equations for a real field. In the upcoming § 2, we shall study the radiation field described by its complex components in the reciprocal space. We will then generalize to a complex field all the results established earlier for discrete and complex variables. This will yield the expression for the free field Hamiltonian and the commutation relations for the free field, which are essential for the quantum description of the field.

The Lagrangian L of a real field is now the integral in real space of a *Lagrangian density* $\mathscr{L}$:

$$L = \int \mathrm{d}^3 r\, \mathscr{L}(\boldsymbol{r}, t) \tag{30}$$

The Lagrangian density is a function of the field $\phi_j(\boldsymbol{r})$ and of its partial derivatives with respect to t and to the components of $\boldsymbol{r}$:

$$\mathscr{L}(\boldsymbol{r},t) = \mathscr{L}\left(\phi_j(\boldsymbol{r},t), \dot{\phi}_j(\boldsymbol{r},t), \partial_i\phi_j(\boldsymbol{r},t)\right) \tag{31}$$

with the notation ($i = x, y, z$):

$$\dot{\phi}_j(\boldsymbol{r},t) = \partial\phi_j(\boldsymbol{r},t)/\partial t \tag{32}$$

$$\partial_i\phi_j(\boldsymbol{r},t) = \partial\phi_j(\boldsymbol{r},t)/\partial r_i \tag{33}$$

Consider a possible path Γ for the field, going from the value $\phi_j(\boldsymbol{r}, t_{\text{in}})$ at an initial time t_{in} to the final value $\phi_j(\boldsymbol{r}, t_2)$ at a final time t_2. The action S_Γ associated with this path is, by definition:

$$S_\Gamma = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3r\, \mathscr{L}\left(\phi_j(\boldsymbol{r},t), \dot{\phi}_j(\boldsymbol{r},t), \partial_i\phi_j(\boldsymbol{r},t)\right) \tag{34}$$

The principle of least action postulates that among all the possible paths starting from the same initial state and ending at the same final state[4], the path(s) actually followed by the system is the one (or are those) for which S_Γ presents an extremum. Let us compute the variation δS of the action for an infinitesimal variation of the path, characterized by the infinitesimal variations $\delta\phi_j(\boldsymbol{r},t)$, $\delta\left(\partial\phi_j(\boldsymbol{r},t)/\partial t\right)$ and $\delta\left(\partial\phi_j(\boldsymbol{r},t)/\partial r_i\right)$.

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3r \sum_j \left[\delta\phi_j(\boldsymbol{r},t)\frac{\partial\mathscr{L}}{\partial\phi_j} + \delta\dot{\phi}_j(\boldsymbol{r},t)\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j} + \sum_i \delta\left(\partial_i\phi_j(\boldsymbol{r},t)\right)\frac{\partial\mathscr{L}}{\partial(\partial_i\phi_j)}\right] \tag{35}$$

Using:

$$\begin{aligned} \delta\dot{\phi}_j(\boldsymbol{r},t) &= \partial\left(\delta\phi_j(\boldsymbol{r},t)\right)/\partial t \\ \delta\left(\partial_i\phi_j(\boldsymbol{r},t)\right) &= \partial_i\left(\delta\phi_j(\boldsymbol{r},t)\right) \end{aligned} \tag{36}$$

and performing an integration by parts of the terms proportional to $\partial\left(\delta\phi_j(\boldsymbol{r},t)\right)/\partial t$ and $\partial_i\left(\delta\phi_j(\boldsymbol{r},t)\right)$, we find that the integrated terms are zero because of the boundary conditions for $\delta\phi_j(\boldsymbol{r},t)$ at the initial and final times, and for $\boldsymbol{r} \to \infty$. The remaining terms are therefore all proportional to $\delta\phi_j(\boldsymbol{r},t)$. Grouping them all, we find[5]:

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3r \sum_{ij} \delta\phi_j(\boldsymbol{r},t)\left[\frac{\partial\mathscr{L}}{\partial\phi_j} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j} - \sum_i \partial_i\frac{\partial\mathscr{L}}{\partial(\partial_i\phi_j)}\right] \tag{37}$$

As δS must be zero for any time or spatial variations of $\delta\phi_j(\boldsymbol{r},t)$, we can deduce that:

$$\frac{\partial\mathscr{L}}{\partial\phi_j} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j} - \sum_i \partial_i\frac{\partial\mathscr{L}}{\partial(\partial_i\phi_j)} = 0 \tag{38}$$

which are the Lagrange equations for the field.

[4] We also assume that the Lagrangian density is zero or tends to zero fast enough when $\boldsymbol{r}$ goes to infinity.

[5] The function $\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j}$ does not directly depend on time t. It nevertheless depends indirectly on t if we replace, as in (31), the fields and their derivatives by their values for a given history of the field. By convention, we then denote by $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j}$ the time derivative of this function at each point of space. It contains the sum of the contributions of the partial derivatives of the function with respect to all the initial variables (the fields and their derivatives).

2. Application to the free radiation field

We now study the free radiation field (in the absence of sources) starting from its Lagrangian density in reciprocal space. We choose the Coulomb gauge so that the longitudinal vector potential $\boldsymbol{A}_{\parallel}$ is zero and $\boldsymbol{A}$ is reduced to $\boldsymbol{A}_{\perp}$; furthermore, the scalar potential is also zero since, in the Coulomb gauge, it would be the potential corresponding to the Coulomb field created by the charges (see relation (A-34) of Chapter XVIII), and we are assuming that there are no charges. The only fields we have to consider are thus the transverse electric field and the magnetic field related to the transverse potential vector by the following equations in real space:

$$\boldsymbol{E}_{\perp}(\boldsymbol{r},t) = -\dot{\boldsymbol{A}}_{\perp}(\boldsymbol{r},t) \qquad \boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{\nabla} \times \boldsymbol{A}_{\perp}(\boldsymbol{r},t) \tag{39}$$

and in the reciprocal space:

$$\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t) = -\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t) \qquad \tilde{\boldsymbol{B}}(\boldsymbol{k},t) = i\boldsymbol{k} \times \tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t) \tag{40}$$

2-a. Lagrangian densities in real and reciprocal spaces

The Lagrangian density most commonly used in real space is[6]:

$$\mathscr{L}(\boldsymbol{r},t) = \frac{\varepsilon_0}{2}\left[\boldsymbol{E}_{\perp}^2(\boldsymbol{r},t) - c^2\boldsymbol{B}^2(\boldsymbol{r},t)\right] \tag{41}$$

where c is the speed of light. Using (39), we see that this Lagrangian density depends both on $\boldsymbol{A}_{\perp}(\boldsymbol{r},t)$ and $\dot{\boldsymbol{A}}_{\perp}(\boldsymbol{r},t)$, as well as on the spatial derivatives of $\boldsymbol{A}_{\perp}(\boldsymbol{r},t)$.

We now go to the reciprocal space. The Lagrangian L is then written as:

$$L = \int \mathrm{d}^3k\, \tilde{\mathscr{L}}(\boldsymbol{k},t) \tag{42}$$

where the Lagrangian density $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ in the reciprocal space is obtained from (41), rewriting the fields in the reciprocal space[7]. Let us evaluate the contribution to the Lagrangian of the two terms in the bracket of (41). Using (40) and the Parseval-Plancherel equality – see relation (A-12) of Chapter XVIII – we can write:

$$\begin{aligned}\int \mathrm{d}^3r\, \boldsymbol{E}_{\perp}(\boldsymbol{r},t)\cdot\boldsymbol{E}_{\perp}(\boldsymbol{r},t) &= \int \mathrm{d}^3k\, \dot{\tilde{\boldsymbol{A}}}_{\perp}^{\,*}(\boldsymbol{k},t)\cdot\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)\\ \int \mathrm{d}^3r\, \boldsymbol{B}(\boldsymbol{r},t)\cdot\boldsymbol{B}(\boldsymbol{r},t) &= \int \mathrm{d}^3k\, \left[(-i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t))\cdot(i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t))\right]\end{aligned} \tag{43}$$

As the two vectors $\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)$ and $\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)$ are in the plane perpendicular to $\boldsymbol{k}$, we have:

$$\int \mathrm{d}^3k\, \left[(-i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)\cdot(i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\right] = \int \mathrm{d}^3k\, k^2\, \tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t) \tag{44}$$

Finally, the Lagrangian density in the reciprocal space is written as:

$$\tilde{\mathscr{L}}(\boldsymbol{k},t) = \frac{\varepsilon_0}{2}\left[\dot{\tilde{\boldsymbol{A}}}_{\perp}^{*}(\boldsymbol{k},t)\cdot\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t) - c^2k^2\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\right] \tag{45}$$

[6] This density has the advantage of being a relativistic invariant (Lorentz scalar).

[7] We use the notation $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ for this Lagrangian density even though this function is not the Fourier transform of $\mathscr{L}(\boldsymbol{r},t)$.

The generalized coordinates of the field can be seen as components of the transverse potential vector, and the generalized velocities as the time derivatives of these coordinates. As opposed to $\mathscr{L}(\boldsymbol{r},t)$, $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ depends only on $\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t)$ and $\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$, and not on the partial derivatives of $\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t)$ with respect to the $\boldsymbol{k}$ components. The computations will thus be simpler in the reciprocal space.

It will be useful for what follows to introduce the Cartesian components of $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ in the reference frame formed by $\boldsymbol{\kappa}=\boldsymbol{k}/k$ and the two polarization unit vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ in the plane perpendicular to $\boldsymbol{k}$. Since $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ is transverse, it does not have any component along $\boldsymbol{\kappa}$, and we can write:

$$\tilde{\mathscr{L}}(\boldsymbol{k},t)=\frac{\varepsilon_0}{2}\sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)-c^2k^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{46}$$

where $\sum_{\boldsymbol{\varepsilon}}$ is a simplified notation representing the summation over the two transverse polarizations $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ – see relation (B-13) of Chapter XVIII.

2-b. Lagrange's equations

Equation (38) becomes here:

$$\frac{\partial\tilde{\mathscr{L}}(\boldsymbol{k},t)}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}-\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial\tilde{\mathscr{L}}(\boldsymbol{k},t)}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=0 \tag{47}$$

We then get, taking (46) into account:

$$\ddot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=0 \tag{48}$$

This equation coincides with equations (A-30a) and (A-30b) of Chapter XVIII, which give the time evolution of the transverse vector potential of a free field in the absence of sources (we set $\tilde{\boldsymbol{j}}_\perp=\boldsymbol{0}$). We recover, as expected, the predictions of Maxwell's equations in the usual formulation of classical electrodynamics.

2-c. Conjugate momentum of the transverse potential vector

To define the conjugate momentum $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ of the complex variable $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$, we use expression (23). Note however that the velocity $\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ appears several times in the integral over $\boldsymbol{k}$ of $\tilde{\mathscr{L}}(\boldsymbol{k},t)$. Consequently, we must add all the corresponding contributions of the partial derivatives of $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ with respect to these various velocities in the definition of the conjugate momentum $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$. This situation results from the fact that the fields are real in real space. The Fourier transform properties then lead to (*cf.* relation (B-18) of Chapter XVIII):

$$\boldsymbol{A}_\perp(\boldsymbol{r},t)=\boldsymbol{A}^*_\perp(\boldsymbol{r},t)\qquad\Rightarrow\qquad\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)=\tilde{\boldsymbol{A}}^*_\perp(-\boldsymbol{k},t) \tag{49}$$

and to an equivalent relation for the time derivatives of the components of the transverse potential vector. In the integral over $\boldsymbol{k}$ of $\tilde{\mathcal{L}}(\boldsymbol{k},t)$ we get, in addition to the term $\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$, the term $\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(-\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(-\boldsymbol{k},t)$ which, according to (49), is equal to

$\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)^*$ and therefore doubles the first term. If we ignore the terms in $-\boldsymbol{k}$, we must then double the contribution of the terms in $\boldsymbol{k}$, which yields:

$$\begin{aligned}\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) &= 2\frac{\partial\tilde{\mathscr{L}}(\boldsymbol{k},t)}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)} \\ &= \varepsilon_0\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = -\varepsilon_0\tilde{E}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\end{aligned} \tag{50}$$

The conjugate momentum of the transverse potential is seen to be equal, within a factor $-\varepsilon_0$, to the transverse electric field.

Another equivalent way to obtain (50) is to use, in the Lagrangian expression, only independent variables. The reality condition (49) ensures that, if one knows the variables in half the reciprocal space, one knows them in the entire space. One can define L as the integral over only half the reciprocal space (where all the variables are independent) of a Lagrangian density, noted $\bar{\mathscr{L}}(\boldsymbol{k},t)$, equal to twice the initial density. Writing "$\fint$ " the integral over half a space (the bar indicating that the $\boldsymbol{k}$ space is divided into two parts), we get:

$$L = \fint \mathrm{d}^3k\ \bar{\mathscr{L}}(\boldsymbol{k},t) \tag{51}$$

with:

$$\bar{\mathscr{L}}(\boldsymbol{k},t) = \varepsilon_0\sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) - \omega^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] = 2\ \tilde{\mathscr{L}}(\boldsymbol{k},t) \tag{52}$$

so that one can also define the conjugate momentum of the transverse potential vector as:

$$\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = \frac{\partial\bar{\mathscr{L}}(\boldsymbol{k},t)}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)} = \varepsilon_0\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{53}$$

2-d. Hamiltonian; Hamilton-Jacobi equations

The Hamiltonian H of the free radiation field is obtained by generalizing, to continuous variables, expression (25) established for discrete variables. To only include independent variables in the integral over $\boldsymbol{k}$ of the Hamiltonian density $\bar{\mathcal{H}}(\boldsymbol{k},t)$, this integral is taken over only half the reciprocal space:

$$H = \fint \mathrm{d}^3k\ \bar{\mathcal{H}}(\boldsymbol{k},t) \tag{54}$$

where the Hamiltonian density $\bar{\mathcal{H}}(\boldsymbol{k},t)$ is equal to:

$$\bar{\mathcal{H}}(\boldsymbol{k},t) = -\bar{\mathscr{L}}(\boldsymbol{k},t) + \sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{\Pi}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + \dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{55}$$

which yields for H, taking (53) and (52) into account:

$$H = \varepsilon_0\fint \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + c^2k^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right]. \tag{56}$$

If the integral over the half-space in (56) is extended to the entire space, one must replace ε_0 by $\varepsilon_0/2$; we then get the same expression as (B-14) of Chapter XVIII for the energy of the free transverse field. The Hamiltonian obtained with the Lagrangian formalism coincides with the field energy.

We can write expression (55) for $\bar{\mathcal{H}}(\boldsymbol{k},t)$ as a function of only the variables $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$. Using (53), we get:

$$\bar{\mathcal{H}}(\boldsymbol{k},t)=\sum_{\boldsymbol{\varepsilon}}\left[\frac{1}{\varepsilon_0}\tilde{\Pi}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\varepsilon_0 c^2k^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{57}$$

Equations (28) can then be generalized and yield the Hamilton-Jacobi equations for $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$:

$$\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=\frac{\partial\bar{\mathcal{H}}(\boldsymbol{k},t)}{\partial\tilde{\Pi}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=\frac{1}{\varepsilon_0}\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{58a}$$

$$\dot{\tilde{\Pi}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=-\frac{\partial\bar{\mathcal{H}}(\boldsymbol{k},t)}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=-\varepsilon_0 c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{58b}$$

It is easy to check that the two equations (58a) and (58b) are the same as Maxwell's equations (A-30a) and (A-30b) of Chapter XVIII, which describes the evolution of the transverse fields in the absence of sources. Equation (58a) of the present complement and equations (A-30a) of Chapter XVIII are identical, and define the transverse electric field as a function of the time derivative of the transverse potential vector. As for equation (58b) of this complement and equation (A-30b) of Chapter XVIII, they are the same when $\tilde{\boldsymbol{j}}=\boldsymbol{0}$. They describe the evolution of the transverse electric field.

2-e. Field commutation relations

Generalizing the canonical commutation relation (29) to the case of continuous variables, we get:

$$\left[\hat{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}),\hat{\tilde{\Pi}}^{\dagger}_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right]=i\hbar\,\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{59}$$

all the other commutators being equal to zero. The Kronecker delta $\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}$ of the vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}'$ is equal to 1 if both these vectors are the same, and to 0 if they are different.

Comment

The canonical commutation relations only apply to independent conjugate variables, which is the case for the components of the various fields along the transverse polarization directions. Now the field components on an arbitrary fixed reference frame$\{\boldsymbol{e}_x,\ \boldsymbol{e}_y,\ \boldsymbol{e}_z\}$, $\tilde{A}_{\perp i}(\boldsymbol{k})$ with $i=x,y,z$, are not independent because of the transversality condition $\sum_i k_i\tilde{A}_{\perp i}(\boldsymbol{k})=0$. Therefore:

$$\left[\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}),\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')\right]\neq i\hbar\,\delta_{ij}\,\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{60}$$

To get the correct commutation relation between $\hat{\tilde{A}}_{\perp i}(\boldsymbol{k})$ and $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')$, we must express both quantities as functions of their components along the two polarization vectors $\boldsymbol{\varepsilon}$ and

$\boldsymbol{\varepsilon}'$, which are perpendicular both to each other and to $\boldsymbol{k}$ (here we choose a basis of linear, real, polarizations), and then use (59). As an example:

$$\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}) = \varepsilon_i \hat{\tilde{A}}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}) + \varepsilon'_i \hat{\tilde{A}}_{\perp \boldsymbol{\varepsilon}'}(\boldsymbol{k}) \tag{61}$$

where:

$$\varepsilon_i = \boldsymbol{e}_i \cdot \boldsymbol{\varepsilon} \qquad\qquad \varepsilon'_i = \boldsymbol{e}_i \cdot \boldsymbol{\varepsilon}' \tag{62}$$

We thus get:

$$\left[\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}), \hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')\right] = i\hbar \left(\varepsilon_i \varepsilon_j + \varepsilon'_i \varepsilon'_j\right) \delta(\boldsymbol{k} - \boldsymbol{k}') \tag{63}$$

This equation can be further transformed by noting that $\boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon}'$ and $\boldsymbol{k}/k$ form an orthonormal basis, so that:

$$\varepsilon_i \varepsilon_j + \varepsilon'_i \varepsilon'_j + k_i k_j / k^2 = \delta_{ij} \tag{64}$$

Finally, the correct commutation relation between $\hat{\tilde{A}}_{\perp i}(\boldsymbol{k})$ and $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')$ is written:

$$\left[\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}), \hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')\right] = i\hbar \left[\delta_{ij} - \frac{k_i k_j}{k^2}\right] \delta(\boldsymbol{k} - \boldsymbol{k}') \tag{65}$$

We multiply both sides of (65) by $e^{i\boldsymbol{k}\cdot\boldsymbol{r}} e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}'}/(2\pi)^3$ and integrate over $\boldsymbol{k}$ and $\boldsymbol{k}'$. We then get on the left-hand side the commutator of the fields in real space[8], $[\hat{A}_{\perp i}(\boldsymbol{r}), \hat{\Pi}_{\perp j}(\boldsymbol{r}')]$, and on the right-hand side the Fourier transform of the function $(\delta_{ij} - k_i k_j/k^2)$ which is the *transverse delta function*[9] $\delta^{\perp}_{ij}(\boldsymbol{r} - \boldsymbol{r}')$:

$$\begin{aligned}
&\left[\hat{A}_{\perp i}(\boldsymbol{r}), \hat{\Pi}_{\perp j}(\boldsymbol{r}')\right] \\
&\quad = \frac{i\hbar}{(2\pi)^3} \int \mathrm{d}^3 k \int \mathrm{d}^3 k' \; e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}'} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \left[\delta_{ij} - \frac{k_i k_j}{k^2}\right] \delta(\boldsymbol{k} - \boldsymbol{k}') \\
&\quad = \frac{i\hbar}{(2\pi)^3} \int \mathrm{d}^3 k \; e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \left(\delta_{ij} - \frac{k_i k_j}{k^2}\right) \equiv i\hbar \delta^{\perp}_{ij}(\boldsymbol{r} - \boldsymbol{r}')
\end{aligned} \tag{66}$$

2-f. Creation and annihilation operators

The normal variable $\boldsymbol{\alpha}(\boldsymbol{k}, t)$ introduced in equation (B-15) of Chapter XVIII has a time evolution $e^{-i\omega t}$, where $\omega = ck$ for the free field case. According to § B-1 of Chapter XVIII, this variable becomes, upon quantization, the annihilation operator $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$ of the harmonic oscillator associated with the mode $\{\boldsymbol{k}, \boldsymbol{\varepsilon}\}$. As for the complex conjugate of this normal variable, it becomes the creation operator $\hat{a}^{\dagger}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$. These two operators are therefore written as:

$$\begin{aligned}
\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t) &= \mathcal{N}(k) \left[\hat{A}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) + \frac{i}{\varepsilon_0 \omega} \hat{\Pi}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)\right] \\
\hat{a}^{\dagger}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t) &= \mathcal{N}(k) \left[\hat{A}^{\dagger}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) - \frac{i}{\varepsilon_0 \omega} \hat{\Pi}^{\dagger}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)\right]
\end{aligned} \tag{67}$$

[8] For the term in $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')$, we use the reality condition $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}') = \hat{\tilde{\Pi}}_{\perp j}(-\boldsymbol{k}')$ and change $\boldsymbol{k}'$ into $-\boldsymbol{k}'$ in the integral over $\boldsymbol{k}'$.

[9] The interested reader can find details on the properties of that function in Complement A_I of reference [16].

where we have used (53) to write $\dot{\hat{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}) = (1/\varepsilon_0)\hat{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k})$. The quantity $\mathcal{N}(k)$ is a normalization constant, arbitrary for now. It can, however, be determined by imposing for the commutator of the two operators (67) a generalization of the well-known relation $[\hat{a}, \hat{a}^\dagger] = 1$ for the harmonic oscillator. We thus use the two equations (67) to compute the commutator $\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right]$ as a function of the commutators of the fields $\hat{A}_\perp$ and $\hat{\Pi}^\dagger_\perp$ and of their adjoints. Since the only non-zero commutators are between $\hat{A}_\perp$ and $\hat{\Pi}^\dagger_\perp$ as well as between $\hat{A}^\dagger_\perp$ and $\hat{\Pi}_\perp$, we get:

$$\begin{aligned}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] &= \mathcal{N}^2(k)\frac{-i}{\varepsilon_0\omega}\left\{\left[\hat{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{\Pi}^\dagger_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] - \left[\hat{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{A}^\dagger_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right]\right\} \\ &= \mathcal{N}^2(k)\frac{-i}{\varepsilon_0\omega}\left\{2i\hbar\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}')\right\} \\ &= \mathcal{N}^2(k)\frac{2\hbar}{\varepsilon_0\omega}\,\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}') \end{aligned} \tag{68}$$

To go from the first to the second line of (68), we used (59) and its complex conjugate. The constant $\mathcal{N}(k)$ is finally determined by imposing the commutators between the annihilation and creation operators to be equal to $\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}')$, which yields:

$$\mathcal{N}(k) = \sqrt{\frac{\varepsilon_0\omega}{2\hbar}} = \sqrt{\frac{\varepsilon_0\ ck}{2\hbar}} \tag{69}$$

Inserting this relation in equality (B-25) of Chapter XVIII, we find that the contribution to the classical energy of the mode $\{\mathbf{k}, \boldsymbol{\varepsilon}\}$ is:

$$\frac{\hbar\omega}{2}\left[\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{70}$$

We shall see in Chapter XIX that it is indeed the equivalent of the expression for the quantized radiation Hamiltonian.

2-g. Discrete momentum variables

We examined, in § B-3 of Chapter XVIII, the case where the radiation is contained in a box of finite volume L^3, which leads to a discrete summation over the momenta. Relation (59) then become:

$$\left[\hat{\hat{A}}_{\perp\boldsymbol{k},\varepsilon}, \hat{\hat{\Pi}}^{\dagger}_{\perp\boldsymbol{k}',\varepsilon'}\right] = i\hbar\ \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta_{\boldsymbol{k},\boldsymbol{k}'} \tag{71}$$

Applying the substitution (B-34) or (B-35) of that chapter to both sides of (67), the two coefficients $(2\pi/L)^3$ cancel each other, and these relations remain unchanged (aside from the fact that $\mathbf{k}$ is now a discrete index rather than a continuous variable).

As for the relations (68), they become:

$$\left[\hat{a}_{\boldsymbol{k},\varepsilon}, \hat{a}^\dagger_{\boldsymbol{k}',\varepsilon'}\right] = \mathcal{N}^2(k)\frac{2\hbar}{\varepsilon_0\omega}\ \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta_{\boldsymbol{k},\boldsymbol{k}'} \tag{72}$$

With the choice (69) for $\mathcal{N}(k)$, we check that the commutator is equal to unity if $\boldsymbol{k} = \boldsymbol{k}'$ and $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}'$.

3. Lagrangian of the global system field + interacting particles

We now study the Lagrangian of the total system, including the interactions between the particles and the electromagnetic field.

3-a. Choice for the Lagrangian

We choose a Lagrangian expressed as:

$$L = L_R + L_P + L_I \tag{73}$$

where L_R depends only on the radiation variables, L_P only on the particle variables, and L_I on both types of variables as it describes the interactions between particles and radiation.

For L_R, we shall take the Lagrangian introduced above for the free field – see relations (51) and (52):

$$L_R = \oint \mathrm{d}^3k\, \bar{\mathscr{L}}_R(\boldsymbol{k},t) = \varepsilon_0 \oint \mathrm{d}^3k \left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) - \omega^2 \tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right]. \tag{74}$$

For the Lagrangian L_P of the particles, labeled by the index a, we shall use the usual Lagrangian for a system of particles, i.e. the difference between their kinetic energy and their potential interaction energy which comes from the Coulomb forces they exert on each others:

$$L_P = \sum_a \frac{1}{2} m_a \left[\dot{\boldsymbol{r}}_a(t)\right]^2 - V_{\mathrm{Coul}} \tag{75}$$

Finally, the interaction Lagrangian will be chosen as:

$$L_I = \int \mathrm{d}^3r\, \boldsymbol{j}(\boldsymbol{r},t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r},t) \tag{76}$$

where $\boldsymbol{j}(\boldsymbol{r},t)$ is the particle current density given by the expression:

$$\boldsymbol{j}(\boldsymbol{r},t) = \sum_a q_a \dot{\boldsymbol{r}}_a(t)\, \delta\left(\boldsymbol{r} - \boldsymbol{r}_a(t)\right) \tag{77}$$

and q_a is the charge of particle a. This expression does not contain terms including $\boldsymbol{A}_\parallel$, the charge density $\rho(\boldsymbol{r},t)$ or the scalar potential $U(\boldsymbol{r}_a,t)$. This is because of our choice of the Coulomb gauge in which $\boldsymbol{A}_\parallel$ is zero, so that the energy of the longitudinal electric field only depends on the particle variables; the same is true for the scalar potential, which is at the origin of the term V_{Coul} included in the particle Lagrangian (75).

For the following computations, it will be useful to give other equivalent expressions for L_I; depending on the problem we focus on, we shall use the most suitable expression. Inserting (77) into (76), we get:

$$L_I = \int \mathrm{d}^3r\, \boldsymbol{j}(\boldsymbol{r},t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r},t) = \sum_a q_a \dot{\boldsymbol{r}}_a(t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r}_a,t) \tag{78}$$

Now, using the Parseval-Plancherel identity, we get:

$$L_I = \int \mathrm{d}^3r\, \boldsymbol{j}(\boldsymbol{r},t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r},t) = \int \mathrm{d}^3k\, \tilde{\boldsymbol{j}}^*(\boldsymbol{k},t) \cdot \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{79}$$

In the integral over k in (79) we get the sum:

$$\tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)+\tilde{\boldsymbol{j}}^*(-\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t)=\tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)+\tilde{\boldsymbol{j}}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp^*(\boldsymbol{k},t) \quad (80)$$

Limiting the integral to half the reciprocal space, we can write:

$$\begin{aligned} L_I &= \int \mathrm{d}^3k\ \tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \\ &= \fint \mathrm{d}^3k\ \left[\tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)+\tilde{\boldsymbol{j}}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp^*(\boldsymbol{k},t)\right] \end{aligned} \quad (81)$$

3-b. Lagrange's equations

We now show that the Lagrange's equations associated with (73) coincide with the Maxwell-Lorentz equations, which will be a justification for the choice of L.

α. Field Lagrange's equations

The time derivative of the transverse potential vector $\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}$ only appears in the Lagrangian density $\bar{\mathscr{L}}_R(\boldsymbol{k},t)$ written in (74). Writing $\bar{\mathscr{L}}$ the Lagrangian density of the total system, we can write:

$$\frac{\partial\bar{\mathscr{L}}}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=\frac{\partial\bar{\mathscr{L}}_R}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=\varepsilon_0\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \quad (82)$$

Now $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ appears both in $\bar{\mathscr{L}}_R$ and in $\bar{\mathscr{L}}_I$ which is the function to be integrated in the last integral of (81). We get:

$$\frac{\partial\bar{\mathscr{L}}}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t))}=\frac{\partial\bar{\mathscr{L}}_R}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t))}+\frac{\partial\bar{\mathscr{L}}_I}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=-\varepsilon_0c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\tilde{j}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \quad (83)$$

The field Lagrange's equation is obtained by setting equal the time derivative of (82) and relation (83), which yields:

$$\ddot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=\frac{1}{\varepsilon_0}\tilde{j}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \quad (84)$$

We thus get the time evolution equation of the transverse potential vector in the presence of sources, which we already obtained in (A-31) of Chapter XVIII.

β. Lagrange's equations for the particles

The velocity $\dot{\boldsymbol{r}}_a$ of particle a appears in both L_P and L_I. Noting $\dot{x}_a$ the component of $\dot{\boldsymbol{r}}_a$ on the x axis, we get:

$$\frac{\partial L}{\partial\dot{x}_a}=\frac{\partial L_P}{\partial\dot{x}_a}+\frac{\partial L_I}{\partial\dot{x}_a}=m_a\dot{x}_a+q_aA_{\perp x}\left(\boldsymbol{r}_a(t),t\right) \quad (85)$$

We now compute the time derivative of this expression. The time dependence of the second term of (85) is explicit via the time dependence of the vector potential, implicit

via the time dependence of the point $\boldsymbol{r}_a(t)$ where this potential is evaluated. We therefore have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_a} = m_a\ddot{x}_a + q_a\frac{\partial A_{\perp x}(\boldsymbol{r}_a,t)}{\partial t} + q_a\dot{\boldsymbol{r}}_a\cdot\boldsymbol{\nabla}A_{\perp x}(\boldsymbol{r}_a,t) \tag{86}$$

The partial derivative of the transverse vector potential with respect to t leads to the transverse electric field:

$$q_a\frac{\partial A_{\perp x}(\boldsymbol{r}_a,t)}{\partial t} = -q_aE_{\perp x}(\boldsymbol{r}_a,t) \tag{87}$$

As for the last term of (86), it can be written:

$$q_a\dot{\boldsymbol{r}}_a\cdot\boldsymbol{\nabla}A_{\perp x}(\boldsymbol{r}_a,t) = q_a\left[\dot{x}_a\frac{\partial A_{\perp x}}{\partial x_a} + \dot{y}_a\frac{\partial A_{\perp x}}{\partial y_a} + \dot{z}_a\frac{\partial A_{\perp x}}{\partial z_a}\right] \tag{88}$$

We now compute the partial derivative of L with respect to the component x_a of $\boldsymbol{r}_a$, which appears in the term V_{Coul} of L_P as well as in L_I (via the position dependence of $\boldsymbol{A}_\perp$). We obtain:

$$\frac{\partial L}{\partial x_a} = \frac{\partial L_P}{\partial x_a} + \frac{\partial L_I}{\partial x_a} = -\frac{\partial V_{\mathrm{Coul}}}{\partial x_a} + q_a\dot{\boldsymbol{r}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial x_a} \tag{89}$$

The term $-\partial V_{\mathrm{Coul}}/\partial x_a$ is the Coulomb force exerted on the particle a, i.e. the force due to the electrostatic field created by the charge distribution and acting on the charge q_a at point $\boldsymbol{r}_a$ where the particle a is located. It can also be written as:

$$-\frac{\partial V_{\mathrm{Coul}}}{\partial x_a} = q_aE_{\parallel x}(\boldsymbol{r}_a,t) \tag{90}$$

where $\boldsymbol{E}_\parallel$ is the longitudinal electric field, since, as we saw in Chapter XVIII, the longitudinal electric field is equal to the electrostatic field created by the charge distribution. Finally, let us explicitly write the last term of (89):

$$q_a\dot{\boldsymbol{r}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial x_a} = q_a\left[\dot{x}_a\frac{\partial A_x}{\partial x_a} + \dot{y}_a\frac{\partial A_y}{\partial x_a} + \dot{z}_a\frac{\partial A_z}{\partial x_a}\right] \tag{91}$$

Lagrange's equation for particle a :

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_a} = \frac{\partial L}{\partial x_a} \tag{92}$$

can thus be written, using the previous results:

$$m_a\ddot{x}_a = q_aE_x(\boldsymbol{r}_a,t) - q_a\dot{\boldsymbol{r}}_a\cdot\boldsymbol{\nabla}A_{\perp x} + q_a\dot{\boldsymbol{r}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial x} \tag{93}$$

where the total electric field at point $\boldsymbol{r}_a$ is :

$$\boldsymbol{E}(\boldsymbol{r}_a) = \boldsymbol{E}_\perp(\boldsymbol{r}_a) + \boldsymbol{E}_\parallel(\boldsymbol{r}_a) \tag{94}$$

We can write explicitly the last term of (93) by regrouping equations (88) and (91). We then get:

$$
\begin{aligned}
&- q_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{\nabla} A_{\perp x} + q_a \dot{\boldsymbol{r}}_a \cdot \frac{\partial \boldsymbol{A}_\perp}{\partial x_a} \\
&\qquad = q_a \left[\dot{y}_a \left(\frac{\partial A_{\perp y}}{\partial x_a} - \frac{\partial A_{\perp x}}{\partial y_a} \right) - \dot{z}_a \left(\frac{\partial A_{\perp x}}{\partial z_a} - \frac{\partial A_{\perp z}}{\partial x_a} \right) \right] \\
&\qquad = q_a \left[\dot{\boldsymbol{r}}_a \times (\boldsymbol{\nabla} \times \boldsymbol{A}_\perp) \right]_x = q_a (\dot{\boldsymbol{r}}_a \times \boldsymbol{B})_x
\end{aligned} \tag{95}
$$

This yields the Lorentz magnetic force exerted by the magnetic field on the particle with velocity $\dot{\boldsymbol{r}}_a$. To sum up, Lagrange's equation for particle a is written:

$$
m_a \ddot{\boldsymbol{r}}_a = q_a \boldsymbol{E}(\boldsymbol{r}_a) + q_a \dot{\boldsymbol{r}}_a \times \boldsymbol{B}(\boldsymbol{r}_a) \tag{96}
$$

and coincides with the Lorentz equation (A-3) of ChapterXVIII. The Lagrangian we chose above therefore leads to the right equations for the field and the particles.

3-c. Conjugate momenta

The equation (82) established above can be used to compute the conjugate momenta $\tilde{\Pi}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$ of the field variables $\tilde{A}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$:

$$
\tilde{\Pi}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) = \frac{\partial \bar{\mathscr{L}}}{\partial \dot{\tilde{A}}^*_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)} = \frac{\partial \bar{\mathscr{L}}_R}{\partial \dot{\tilde{A}}^*_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)} = \varepsilon_0 \dot{\tilde{A}}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)) = -\varepsilon_0 \tilde{E}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) \tag{97}
$$

In a similar way, the computation of the conjugate momentum of the coordinate $\boldsymbol{r}_a$ of particle a is the same as the one leading to equation (85):

$$
\boldsymbol{p}_a = \frac{\partial L}{\partial \dot{\boldsymbol{r}}_a} = \frac{\partial L_P}{\partial \dot{\boldsymbol{r}}_a} + \frac{\partial L_I}{\partial \dot{\boldsymbol{r}}_a} = m_a \dot{\boldsymbol{r}}_a + q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a) \tag{98}
$$

3-d. Hamiltonian

Since $\dot{\tilde{A}}_{\perp \boldsymbol{\varepsilon}}$ and $\tilde{\Pi}^*_{\perp \varepsilon} = \varepsilon_0 \dot{\tilde{A}}^*_{\perp \boldsymbol{\varepsilon}}$ appear only in the radiation Lagrangian $\mathscr{L}_R$, which only depends on the radiation variables, the Hamiltonian of the global system must contain a term H_R identical to the expression (56) found for the free field.

To obtain the other terms coming from the particle conjugate momenta and from the subtraction of L_P and L_I, we first compute $\sum_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{p}_a$. Equation (98) yields:

$$
\sum_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{p}_a = \sum_a \left[m_a \dot{\boldsymbol{r}}_a^2 + q_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{A}_\perp(\boldsymbol{r}_a) \right] \tag{99}
$$

We must now subtract from (99) the values of L_P and L_I given by equations (75) and (78). The term coming from L_I cancels the last term of the right-hand side of (99) and we are left with:

$$
V_{\text{Coul}} + \sum_a \frac{1}{2} m_a \dot{\boldsymbol{r}}_a^2 = V_{\text{Coul}} + \sum_a \frac{\left[\boldsymbol{p}_a - q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a) \right]^2}{2 m_a} \tag{100}
$$

Finally, the global system Hamiltonian is given by:

$$H = H_R + V_{\text{Coul}} + \sum_a \frac{[\boldsymbol{p}_a - q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a)]^2}{2m_a} \tag{101}$$

where H_R has the same form as (56) for the free radiation. This result is a justification for the expression of H given in equation (A-41) of Chapter XVIII.

3-e. Commutation relations

Since the radiation variables and their conjugate momenta are the same as for the free field, the commutation relations (59) established for the free field remain valid:

$$\left[\hat{\vec{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{\vec{\Pi}}^{\dagger}_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = i\hbar\ \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\ \delta(\boldsymbol{k} - \boldsymbol{k}') \tag{102}$$

all the other commutators being equal to zero. As for the commutation relations for the positions and conjugate momenta of the particles, they are the usual relations:

$$[\hat{r}_{ai}, \hat{p}_{bj}] = i\hbar\ \delta_{ab}\ \delta_{ij} \tag{103}$$

where the indices a, b label the particles and the indices $i, j = x, y, z$ the Cartesian components of $\boldsymbol{r}$ and $\boldsymbol{p}$.

As in § 2-g, one can extend these commutation relations to the case where the momenta are discrete.

Chapter XIX

Quantization of electromagnetic radiation

A Quantization of the radiation in the Coulomb gauge 1999
A-1 Quantization rules 1999
A-2 Radiation contained in a box 2001
A-3 Heisenberg equations 2002
B Photons, elementary excitations of the free quantum field . 2004
B-1 Fock space of the free quantum field 2004
B-2 Corpuscular interpretation of states with fixed total energy and momentum 2005
B-3 Several examples of quantum radiation states 2006
C Description of the interactions 2009
C-1 Interaction Hamiltonian 2009
C-2 Interaction with an atom. External and internal variables . . 2010
C-3 Long wavelength approximation 2010
C-4 Electric dipole Hamiltonian 2011
C-5 Matrix elements of the interaction Hamiltonian; selection rules 2014

Introduction

This chapter presents a quantum description of the electromagnetic field and its interactions with an ensemble of charged particles. Such a description is necessary for interpreting certain physical phenomena such as the spontaneous emission of a photon by an excited atom, which cannot be carried out with the semiclassical treatments we have used previously[1] (classical description for the field, and quantum description for the

[1] See for example in Complement A_{XIII} the study of the interaction between an atom and an electromagnetic wave.

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

particles). Imagine, for example, that a monochromatic field with angular frequency ω is described by a classical field $\mathbf{E}_0 \cos \omega t$; its interaction with an atom is then described by the Hamiltonian $\widehat{H}_I = -\widehat{\mathbf{D}} \cdot \mathbf{E}_0 \cos \omega t$, where $\widehat{\mathbf{D}}$ is an operator (the electric dipole moment) whereas $\mathbf{E}_0$ remains a classical quantity[2]. Such a treatment is adequate for understanding how the field can excite the atom from its ground state a with energy E_a towards an excited state b of energy E_b; the processus is resonant if ω is close to the atomic Bohr frequency $\omega_0 = (E_b - E_a)/\hbar$. Imagine now that the atom is initially in the excited state b, in the absence of any incident radiation. The classical field $\mathbf{E}_0$ is then identically zero and, consequently, so is the interaction Hamiltonian $\widehat{H}_I$. The Hamiltonian of the total system is then reduced to the atomic Hamiltonian $\widehat{H}_A$. Since this operator is time-independent, its eigenstates are stationary, including, in particular, the excited state b. The semiclassical theory predicts that an atom, initially excited in a state b in the absence of incident radiation, will remain indefinitely in that state. But this is not what is experimentally observed: after a certain time, the atom spontaneously falls into a lower level a, emitting a photon whose frequency is close to $\omega_0 = (E_b - E_a)/\hbar$. This process is called spontaneous emission and happens after an average time called the *radiative lifetime* of the excited state b. This is a first example of a situation where a radiation quantum treatment is indispensable. It is far from being the only example: numerous experiments, more and more elaborate, have created situations where the quantum description of the electromagnetic field is necessary.

This chapter presents the base of this quantum description, while following an approach that is as simple as possible – a more general presentation of the quantization of the electromagnetic field is possible with the Lagrangian formulation of electrodynamics (Complement $\mathrm{A_{XVIII}}$. In the previous chapter, we underlined the analogy between the eigenmodes of the radiation field vibrations and an ensemble of harmonic oscillators. We shall use this analogy in § A of this chapter, and proceed to a simple quantization of this ensemble of oscillators. With each eigenmode i of the classical field, described by normal variables α_i and α_i^*, we shall associate annihilation $\widehat{a}_i$ and creation $\widehat{a}_i^\dagger$ operators, obeying the well-known commutation relations $\left[\widehat{a}_i, \widehat{a}_i^\dagger\right] = 1$. We shall also propose a plausible form for the quantum Hamiltonian of the system "field + particles", starting from the classical energy of that system established in the previous chapter. We will see that the equations of evolution[3] for these various quantities in the Heisenberg picture (Complement $\mathrm{G_{III}}$) are the transposition of the Maxwell-Lorentz equations to operators describing fields and particles, properly symmetrized. This will yield an a posteriori justification for the simple quantization procedure we used.

Several important properties of the free field (in the absence of sources) are described in § B. The state space of this field has the structure of a tensor product of Fock spaces, analogous to those studied in Chapter XV; the elementary excitations of the field are called photons. A few important states of the field will be described: the photon vacuum, where no photons are present (but where there exists, nonetheless, a fluctuating field throughout the entire space, with a zero average value), the one-photon states, and the quasi-classical states, which reproduce the properties of a given classical field.

Finally, § C studies the interaction Hamiltonian between an electromagnetic field and particles, in particular when those are neutral atoms (such as the Hydrogen atom

[2]For the sake of clarity, we use in the entire chapter and its complements the symbol "hat" to distinguish an operator $\widehat{G}$ from its corresponding classical quantity G.

[3]More concisely, we shall call them Heinsenberg equations.

where the positive and negative charges of the atom's constituents balance each other). It is then possible to distinguish between two types of atomic variables: the center of mass variables (external variables) and the "relative motion" variables in the center of mass frame (internal variables). We shall also study the electric dipole approximation, valid when the radiation wavelength is large compared to the atomic sizes, as well as the selection rules associated with the interaction Hamiltonian.

A. Quantization of the radiation in the Coulomb gauge

A-1. Quantization rules

In the previous chapter, we established in relation (B-26) the following expression for the energy of the classical transverse field:

$$H_{\text{trans}} = \varepsilon_0 \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\omega^2}{4\mathcal{N}^2(k)} \left[\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{A-1}$$

where $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ are the normal variables describing the transverse field, $\omega = ck$, and $\mathcal{N}(k)$ a real normalization constant that appeared in the equations defining the normal variables in terms of the transverse potential vector and its time derivative:

$$\begin{aligned}
\boldsymbol{\alpha}(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) + \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)\right] \\
\boldsymbol{\alpha}^*(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}^*_\perp(\boldsymbol{k},t)\right]
\end{aligned} \tag{A-2}$$

The analogy between the free transverse field and an ensemble of classical harmonic oscillators of frequency ω associated with the modes $\{\boldsymbol{k},\boldsymbol{\varepsilon}\}$ is clearly seen in expression (A-1).

To quantize the field, this analogy suggests replacing the normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ by annihilation and creations operators. We shall use in this § A the Schrödinger picture where these operators are time-independent and where the time dependence only appears in the evolution of the state vector. The quantization procedure will consist in replacing the $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t=0)$ by time-independent annihilation operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$, and of course the $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t=0)$ by the adjoint creation operators $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$. Once this operation is performed on (A-1), we obtain a quantum Hamiltonian identical to a sum of standard harmonic oscillator Hamiltonians, provided the factor $\omega^2/4\mathcal{N}^2(k)$ multiplying the bracket on the right-hand side of (A-1) is equal to $\hbar\omega/2\epsilon_0$. We therefore choose for $\mathcal{N}(k)$ the value:

$$\mathcal{N}(k) = \sqrt{\frac{\varepsilon_0\, ck}{2\hbar}} = \sqrt{\frac{\varepsilon_0\,\omega}{2\hbar}} \tag{A-3}$$

This relation is the same as relation (69) of Complement $\mathrm{A_{XVIII}}$, obtained from the commutation relations. We now replace in (A-1) the classical normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ by the operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ obeying the commutation relations:

$$\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\ \delta(\boldsymbol{k}-\boldsymbol{k}') \tag{A-4a}$$

$$\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = 0 \tag{A-4b}$$

This yields the Hamiltonian operator (as this operator will be frequently used, we simplify the notation and replace $\hat{H}_{\rm trans}$ by $\hat{H}_R$):

$$\hat{H}_R \equiv \hat{H}_{\rm trans} = \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\hbar\omega}{2} \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{A-5}$$

which has the expected form for the quantum Hamiltonian of the transverse field.

Extending this procedure, we now replace the classical normal variables by annihilation and creation operators in all the classical expressions established in the previous chapter for the various physical quantities. The transverse momentum – see equation (B-27) of Chapter XVIII – becomes:

$$\hat{\boldsymbol{P}}_{\rm trans} = \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\hbar\boldsymbol{k}}{2} \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{A-6}$$

As for the transverse fields, written in (B-29), (B-30) and (B-28) of Chapter XVIII, they become linear combinations of creation and annihilation operators:

$$\hat{\boldsymbol{E}}_\perp(\boldsymbol{r}) = i \int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}} \sum_{\boldsymbol{\varepsilon}} \left[\frac{\hbar\omega}{2\varepsilon_0}\right]^{1/2} \left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-7}$$

$$\hat{\boldsymbol{B}}(\boldsymbol{r}) = \frac{i}{c} \int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}} \sum_{\boldsymbol{\varepsilon}} \left[\frac{\hbar\omega}{2\varepsilon_0}\right]^{1/2} \left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-8}$$

$$\hat{\boldsymbol{A}}_\perp(\boldsymbol{r}) = \int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}} \sum_{\boldsymbol{\varepsilon}} \left[\frac{\hbar}{2\varepsilon_0\omega}\right]^{1/2} \left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-9}$$

Comment:

As in Chapter XVIII, these relations are written in the general case where the polarizations may be complex (elliptical or circular). Complex conjugate $\boldsymbol{\varepsilon}*$ of the polarization vectors are therefore associated with the creation operators. It is of course necessary to check that the quantification procedure is independent of the arbitrary choice of the polarization basis. If a quantization is performed with a given basis of polarizations, by substitution one can calculate the operators multiplying $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}^*$ in the new basis, and check that the commutation relations of these operators are indeed those of standard creation and annihilation operators. This ensures the polarization basis independence.

Finally, relation (A-48) of Chapter XVIII for the total energy of the system "particles + fields" becomes:

$$\hat{H} = \frac{1}{2m_a} \sum_a \left[\hat{\boldsymbol{p}}_a - q_a \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a)\right]^2 + \hat{V}_{\rm Coul} + \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\hbar\omega}{2} \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{A-10}$$

which is a plausible form for the quantum Hamiltonian of the system "particles + fields". The position $\hat{\boldsymbol{r}}_a$ and momentum $\hat{\boldsymbol{p}}_a$ operators defined using equation (A-47) of Chapter XVIII obey the usual commutation relations:

$$[(\hat{\boldsymbol{r}}_a)_i, (\hat{\boldsymbol{p}}_b)_j] = i\hbar\, \delta_{ab}\, \delta_{ij} \tag{A-11a}$$

$$[(\hat{\boldsymbol{r}}_a)_i, (\hat{\boldsymbol{r}}_b)_j] = [(\hat{\boldsymbol{p}}_a)_i, (\hat{\boldsymbol{p}}_b)_j] = 0 \tag{A-11b}$$

The quantization rules we just heuristically introduced have the advantage of simplicity. We are going to show in addition that the Heisenberg equations for the various operators describing the particles and the fields, deduced from the Hamiltonian (A-10) as well as from the commutation relations (A-4), (A-11a) and (A-11b), are indeed the Maxwell-Lorentz equations for operators. This result justifies a posteriori the quantization procedure exposed in this chapter.

A-2. Radiation contained in a box

If the real space is infinite, $\boldsymbol{k}$ is a continuous variable, and there exists a continuous infinity of modes. However, as we mentioned in § B-3 of Chapter XVIII, it is often more convenient to consider the field to be contained in a cube of edge length L with periodic boundary conditions; the variable $\boldsymbol{k}$ is now discrete:

$$k_{x,y,z} = 2\pi \; n_{x,y,z}/L \tag{A-12}$$

where $n_{x,y,z}$ are positive, negative or zero integers. All the physical predictions must be independent of L when it is large enough. In such an approach, we replace the Fourier integrals by Fourier series and the integrals over $\boldsymbol{k}$ by discrete summations. For a classical field, the continuous variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$ then become discrete variables $\alpha_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t)$. If the field is zero outside the box, relation (B-35) of chapter XVIII indicates the multiplicative factor that must be used to go from one type of variable to the other.

The system is then quantized as we just explained. In the Schrödinger picture, each classical coefficient $\alpha_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t=0)$ in a Fourier series becomes an annihilation operator $\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}}$; each coefficient $\alpha^*_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t=0)$ becomes a creation operator $\hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}}$. This latter operator creates a quantum in a field mode confined inside the box (instead of spreading over the entire space). The commutation relations (A-4) are then written:

$$\left[\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}}, \hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}'}\right] = \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'} \; \delta_{\boldsymbol{k},\boldsymbol{k}'} \tag{A-13a}$$

$$[\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}}, \hat{a}_{\boldsymbol{k}',\boldsymbol{\varepsilon}'}] = \left[\hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}}, \hat{a}^\dagger_{\boldsymbol{k}',\boldsymbol{\varepsilon}'}\right] = 0 \tag{A-13b}$$

Relation (B-36) of Chapter XVIII indicates that once the discrete variables have been inserted in the expressions for the fields, the following rule must be applied to go from a continuous to a discrete summation:

$$\int \mathrm{d}^3k \implies \left(\frac{2\pi}{L}\right)^{3/2} \sum_{\mathbf{k}} \tag{A-14}$$

Expressions (A-7) to (A-9) must be modified. As an example, relation (A-7) becomes:

$$\hat{\boldsymbol{E}}_\perp(\boldsymbol{r}) = i \sum_{\boldsymbol{k},\boldsymbol{\varepsilon}} \left[\frac{\hbar\omega}{2\varepsilon_0 L^3}\right]^{1/2} \left[\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}} \; \boldsymbol{\varepsilon} \; e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}} \; \boldsymbol{\varepsilon}^* \; e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-15}$$

This means that in addition to replacing the integral by a discrete summation, and multiplying by a factor $(2\pi)^{3/2}$, one must divide the field expansion by the square root of the volume L^3. Both relations (A-8) and (A-9) undergo the same changes.

A-3. Heisenberg equations

A-3-a. Heisenberg equations for massive particles

We start with the equation for the evolution of $\hat{\boldsymbol{r}}_a(t)$:

$$\dot{\hat{\boldsymbol{r}}}_a(t) = \frac{1}{i\hbar}\left[\hat{\boldsymbol{r}}_a(t), \hat{H}\right] \tag{A-16}$$

The only term in Hamiltonian (A-10) that does not commute with $\hat{\boldsymbol{r}}_a$ is the first one. Using the commutation relation deduced from (A-11a) and (A-11b) :

$$[(\hat{\boldsymbol{r}}_a)_i, f((\hat{\boldsymbol{p}}_a)_j)] = \delta_{ij}\ \hbar \frac{\partial f}{\partial (\hat{\boldsymbol{p}}_a)_i} \tag{A-17}$$

we get:

$$\begin{aligned}\dot{\hat{\boldsymbol{r}}}_a(t) &= \frac{1}{i\hbar}\left[\hat{\boldsymbol{r}}_a(t), \frac{1}{2m_a}\left(\hat{\boldsymbol{p}}_a(t) - q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t)\right)^2\right] \\ &= \frac{1}{m_a}\left[\hat{\boldsymbol{p}}_a(t) - q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t)\right]\end{aligned} \tag{A-18}$$

This equality is simply the operator form:

$$\hat{\boldsymbol{p}}_a(t) = m_a\dot{\hat{\boldsymbol{r}}}_a(t) + q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t) \tag{A-19}$$

of the classical equation relating the generalized (or canonical) momentum $\boldsymbol{p}_a$ and the mechanical momentum $m_a\dot{\boldsymbol{r}}_a$. We then define the velocity operator $\hat{\boldsymbol{v}}_a$ of particle a by:

$$\hat{\boldsymbol{v}}_a(t) = \frac{1}{m_a}\left[\hat{\boldsymbol{p}}_a(t) - q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t)\right] \tag{A-20}$$

Consider now the Heisenberg equation for the evolution of this operator. It yields the equation of motion of that particle:

$$m_a\dot{\hat{\boldsymbol{v}}}_a(t) = m_a\ddot{\hat{\boldsymbol{r}}}_a(t) = \frac{m_a}{i\hbar}\left[\hat{\boldsymbol{v}}_a(t), \hat{H}\right] \tag{A-21}$$

We shall compute below the commutator $\left[\hat{\boldsymbol{v}}_a(t), \hat{H}\right]$; it leads to the quantum equation of motion for particle a:

$$m_a\ddot{\hat{\boldsymbol{r}}}_a = q_a\hat{\boldsymbol{E}}(\hat{\boldsymbol{r}}_a) + \frac{q_a}{2}\left[\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) - \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \times \hat{\boldsymbol{v}}_a\right] \tag{A-22}$$

which is simply the quantum Lorentz equation describing the motion of particles interacting with the magnetic field $\hat{\boldsymbol{B}}$ and the total electric field $\hat{\boldsymbol{E}} = \hat{\boldsymbol{E}}_\parallel + \hat{\boldsymbol{E}}_\perp$. The special form of the magnetic force $q_a\left[\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) - \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \times \hat{\boldsymbol{v}}_a\right]/2$ comes, as shown in the computation below, from using the Heisenberg equations, and from the fact that the operator $\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a)$ is not Hermitian. To make that operator Hermitian, we must add its adjoint $\left(\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))\right)^\dagger$, which is simply $-\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \times \hat{\boldsymbol{v}}_a$, and divide the result by 2.

Demonstration of equation (A-22)

To compute the commutator of $m_a\hat{\boldsymbol{v}}_a/i\hbar$ with the first term of $\hat{H}$, it is useful to first calculate the following commutators:

$$\begin{aligned} m_a^2\left[(\hat{\boldsymbol{v}}_a)_j,(\hat{\boldsymbol{v}}_a)_l\right] &= -q_a\left[(\hat{\boldsymbol{p}}_a)_j,(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_l\right]-q_a\left[(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_j,(\hat{\boldsymbol{p}}_a)_l\right]\\ &= i\hbar q_a\left[\partial_j(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_l-\partial_l(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_j\right]\\ &= i\hbar q_a\sum_k\epsilon_{jlk}(\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))_k, \end{aligned} \tag{A-23}$$

where ϵ_{jkl} is the completely antisymmetric tensor that allows writing the cross product components of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in the form $(\boldsymbol{a}\times\boldsymbol{b})_k=\sum_{jl}\epsilon_{kjl}a_jb_l$. We then get:

$$\begin{aligned} \frac{m_a}{i\hbar}\left[(\hat{\boldsymbol{v}}_a)_j,\sum_l m_a(\hat{\boldsymbol{v}}_a)_l^2/2\right] &= \frac{m_a^2}{2i\hbar}\sum_l\left\{(\hat{\boldsymbol{v}}_a)_l\left[(\hat{\boldsymbol{v}}_a)_j,(\hat{\boldsymbol{v}}_a)_l\right]+\left[(\hat{\boldsymbol{v}}_a)_j,(\hat{\boldsymbol{v}}_a)_l\right](\hat{\boldsymbol{v}}_a)_l\right\}\\ &= \frac{q_a}{2}\sum_k\sum_l\epsilon_{jlk}\left\{(\hat{\boldsymbol{v}}_a)_l(\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))_k+(\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))_k(\hat{\boldsymbol{v}}_a)_l\right\} \end{aligned} \tag{A-24}$$

The last line in (A-24) can be rewritten in the form:

$$\frac{q_a}{2}\left[\hat{\boldsymbol{v}}_a\times\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a)-\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a)\times\hat{\boldsymbol{v}}_a\right]_j \tag{A-25}$$

and is thus the component along the j axis of the symmmetrized magnetic force.

The commutator of $m_a\hat{\boldsymbol{v}}_a/i\hbar$ with the second term of $\hat{H}$ is written:

$$\frac{m_a}{i\hbar}\left[(\hat{\boldsymbol{v}}_a)_j,V_{\text{Coul}}\right]=\frac{1}{i\hbar}\left[(\hat{\boldsymbol{p}}_a)_j,V_{\text{Coul}}\right]=-\frac{\partial}{\partial(\hat{\boldsymbol{r}}_a)_j}V_{\text{Coul}}=q_a(\hat{\boldsymbol{E}}_\parallel(\hat{\boldsymbol{r}}_a))_j \tag{A-26}$$

It describes the interaction between particle a and the longitudinal electric field.

We finally have to compute the commutator of $m_a\hat{\boldsymbol{v}}_a/i\hbar$ with the last term of $\hat{H}$. Using the commutation relations (A-4) and expressions (A-9) and (A-7) for $\hat{\boldsymbol{A}}_\perp$ and $\hat{\boldsymbol{E}}_\perp$, we get:

$$\begin{aligned} &\frac{m_a}{i\hbar}\left[(\hat{\boldsymbol{v}}_a)_j,\int \mathrm{d}^3k\sum_{\boldsymbol{\varepsilon}}\hbar\omega\left(\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+1/2\right)\right]\\ &= iq_a\int \mathrm{d}^3k\sum_{\boldsymbol{\varepsilon}}\omega\left[(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_j,\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right]\\ &= q_a(\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{r}}_a))_j \end{aligned} \tag{A-27}$$

This term describes the interaction of particle a with the transverse electric field. Finally grouping (A-25), (A-26) and (A-27) leads to (A-22).

A-3-b. Heisenberg equations for fields

As all the fields are linear combinations of the operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$, we simply have to consider the Heisenberg equation for $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$:

$$\dot{\hat{a}}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=\frac{1}{i\hbar}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t),\hat{H}\right] \tag{A-28}$$

We assume that the polarizations $\boldsymbol{\varepsilon}$ are real (linear polarizations). The commutator with the first term of $\hat{H}$ yields, with the use of(A-4a) and (A-20):

$$\frac{1}{i\hbar}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t),\sum_a \frac{m_a\hat{\boldsymbol{v}}_a^2}{2}\right] = \sum_a \frac{-q_a}{2i\hbar}\left\{\hat{\boldsymbol{v}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial\alpha^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})}+\frac{\partial \boldsymbol{A}_\perp}{\partial\alpha^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})}\cdot\hat{\boldsymbol{v}}_a\right\}$$
$$= \sum_a \frac{iq_a}{2\hbar}\sqrt{\frac{\hbar}{2\varepsilon_0\omega(2\pi)^3}}\,\boldsymbol{\varepsilon}\cdot\left[\hat{\boldsymbol{v}}_a e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}+e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}\hat{\boldsymbol{v}}_a\right] \qquad \text{(A-29)}$$

where $\partial\boldsymbol{A}_\perp/\partial a^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ denotes the coefficient of $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ in the integral (A-9) of $\hat{\boldsymbol{A}}_\perp(\boldsymbol{r})$, which is nothing but the coefficient of $\alpha^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t=0)$ in the classical expression of $\boldsymbol{A}_\perp(\boldsymbol{r})$. We introduce the current operator (symmetrized to make it Hermitian):

$$\hat{\boldsymbol{j}}(\boldsymbol{r}) = \frac{1}{2}\sum_a q_a\left[\hat{\boldsymbol{v}}_a\delta(\boldsymbol{r}-\hat{\boldsymbol{r}}_a)+\delta(\boldsymbol{r}-\hat{\boldsymbol{r}}_a)\hat{\boldsymbol{v}}_a\right] \qquad \text{(A-30)}$$

The right hand side term of equation (A-29) can then be rewritten in the form:

$$\sum_a \frac{iq_a/2}{\sqrt{2\varepsilon_0\hbar\omega(2\pi)^3}}\,\boldsymbol{\varepsilon}\cdot\left[\hat{\boldsymbol{v}}_a e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}+e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}\hat{\boldsymbol{v}}_a\right] = \frac{i}{\sqrt{2\varepsilon_0\hbar\omega(2\pi)^3}}\int \mathrm{d}^3r\; e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{j}}(\boldsymbol{r})$$
$$= \frac{i}{\sqrt{2\varepsilon_0\hbar\omega}}\,\boldsymbol{\varepsilon}\cdot\hat{\tilde{\boldsymbol{j}}}(\boldsymbol{k}) \qquad \text{(A-31)}$$

The commutator with the second term of $\hat{H}$ is zero, whereas the commutator with the third term yields, using (A-4):

$$\frac{1}{i\hbar}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t),\int \mathrm{d}^3k'\;\sum_{\boldsymbol{\varepsilon}'}\hbar\omega\left(\hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}',t)\hat{a}_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}',t)+1/2\right)\right] = -i\omega\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \qquad \text{(A-32)}$$

Finally, regrouping (A-31) and (A-32) yields:

$$\dot{\hat{a}}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+i\omega\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = \frac{i}{\sqrt{2\varepsilon_0\hbar\omega}}\,\boldsymbol{\varepsilon}\cdot\hat{\tilde{\boldsymbol{j}}}(\boldsymbol{k},t) \qquad \text{(A-33)}$$

This equation is, for the operator $\hat{a}_\varepsilon(\boldsymbol{k},t)$, an equation of motion of the same form as the equation of motion of the classical normal variables $\boldsymbol{\alpha}(\boldsymbol{k},t)$, which is given by equation (B-19) of Chapter XVIII. As this latter equation is equivalent to Maxwell's equations for the transverse fields, we may conclude that the Heisenberg equations for the quantum transverse fields are simply the usual Maxwell's equations applied to the field operators.

B. Photons, elementary excitations of the free quantum field

We now study a certain number of properties of the electromagnetic field we just quantized, starting with the simplest case: the field in the absence of charged particles.

B-1. Fock space of the free quantum field

The state space of the total system "field + particles" is the tensor product of the particle state space $\mathcal{E}_P$ and the radiation field state space $\mathcal{E}_R$. This latter space is itself

the tensor product of the state spaces of the harmonic oscillators associated with the different modes $\{\boldsymbol{k}, \boldsymbol{\varepsilon}\}$:

$$\mathcal{E}_R = \mathcal{E}_{\mathbf{k}_1,\varepsilon_1} \otimes \mathcal{E}_{\mathbf{k}_2,\varepsilon_2} \otimes .... \otimes \mathcal{E}_{\mathbf{k}_i,\varepsilon_i} \otimes ... \tag{B-1}$$

where $\mathcal{E}_{\mathbf{k}_i,\varepsilon_i}$ is the state space of the harmonic oscillator associated with the mode $\{\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i\}$, with frequency ω_i.

As in § A-2, we assume the radiation to be contained in a box of edge length L. The operators $\hat{a}_\varepsilon(\boldsymbol{k})$ depending on the variables $\boldsymbol{k}$ are then transformed into operators $\hat{a}_{\boldsymbol{k}_i,\boldsymbol{\varepsilon}_i}$ depending only on discrete variables. We can even use a more compact notation $\hat{a}_i$, where the index i labels[4] the whole set of indices $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$; the operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ are now simply written $\hat{a}_i$. In this section, it is convenient to use the Heisenberg picture; the time dependence of the $\hat{a}_i$ and $\hat{a}_i^\dagger$ is then particularly simple, since we have:

$$\hat{a}_i(t) = \exp(i\hat{H}_R t/\hbar)\, \hat{a}_i\, \exp(-i\hat{H}_R t/\hbar) = \hat{a}_i\, e^{-i\omega_i t} \tag{B-2}$$

as well as the Hermitian conjugate relation.

Once the discrete variables have been inserted in the continuous expressions of the fields, we must use rule (A-14) to transform the continuous integrals into discrete summations. The expansions of these fields in term of normal variables are then:

$$\hat{\boldsymbol{E}}_\perp(\boldsymbol{r}, t) = i \sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} \left[\hat{a}_i \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)} - \hat{a}_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)}\right] \tag{B-3}$$

$$\hat{\boldsymbol{B}}(\boldsymbol{r}, t) = i \sum_i \left(\frac{\hbar k_i}{2\epsilon_0 c L^3}\right)^{1/2} \left[\hat{a}_i \boldsymbol{\kappa}_i \times \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)} - \hat{a}_i^\dagger \boldsymbol{\kappa}_i \times \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)}\right] \tag{B-4}$$

$$\hat{\boldsymbol{A}}_\perp(\boldsymbol{r}, t) = \sum_i \left(\frac{\hbar}{2\epsilon_0 \omega_i L^3}\right)^{1/2} \left[\hat{a}_i \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)} + \hat{a}_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)}\right] \tag{B-5}$$

$$\hat{H}_R = \sum_i \frac{\hbar\omega_i}{2} \left[\hat{a}_i^\dagger \hat{a}_i + \hat{a}_i \hat{a}_i^\dagger\right] = \sum_i \hbar\omega_i \left[\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2}\right] \tag{B-6}$$

$$\hat{\boldsymbol{P}}_{\text{trans}} = \sum_i \frac{\hbar\boldsymbol{k}_i}{2} \left[\hat{a}_i^\dagger \hat{a}_i + \hat{a}_i \hat{a}_i^\dagger\right] = \sum_i \hbar\boldsymbol{k}_i\, \hat{a}_i^\dagger \hat{a}_i \tag{B-7}$$

Note that the last term in (B-7) does not contain the factor 1/2, since $\sum_i \boldsymbol{k}_i = \boldsymbol{0}$.

B-2. Corpuscular interpretation of states with fixed total energy and momentum

Consider first the mode i. The eigenvalues of the operator $\hat{a}_i^\dagger \hat{a}_i$ appearing in expressions (B-6) and (B-7) for $\hat{H}_R$ and $\hat{\boldsymbol{P}}_{\text{trans}}$ are all the positive or zero integers n_i :

$$\hat{a}_i^\dagger \hat{a}_i |n_i\rangle = n_i |n_i\rangle, \quad n_i = 0, 1, 2... \tag{B-8}$$

[4] For each $\boldsymbol{k}_i$, there exists two polarization vectors $\boldsymbol{\varepsilon}_{i1}$ and $\boldsymbol{\varepsilon}_{i2}$ perpendicular to $\boldsymbol{k}_i$ and perpendicular to each other. The compact notation $\sum_i$ must be interpreted as a summation over $\boldsymbol{k}_i$ and, for each value of $\boldsymbol{k}_i$, as a sum over $\boldsymbol{\varepsilon}_{i1}$ and $\boldsymbol{\varepsilon}_{i2}$.

Remember the well-known actions of operators $\hat{a}_i^\dagger$ and $\hat{a}_i$ on the states $|n_i\rangle$:

$$\begin{aligned}\hat{a}_i^\dagger|n_i\rangle &= \sqrt{n_i+1}|n_i\rangle \\ \hat{a}_i|n_i\rangle &= \sqrt{n_i}|n_i-1\rangle \\ \hat{a}_i|0_i\rangle &= 0\end{aligned} \tag{B-9}$$

As $\hat{a}_i^\dagger\hat{a}_i$ commutes with $\hat{a}_j^\dagger\hat{a}_j$, the eigenstates of $\hat{H}_R$ and $\hat{\boldsymbol{P}}_{\text{trans}}$ are the tensor products of the eigenstates $|n_1...n_i...\rangle = |n_1\rangle\otimes...\otimes|n_i\rangle...$ of the creation and annihilation operators $\hat{a}_1^\dagger\hat{a}_1$,....$\hat{a}_i^\dagger\hat{a}_i$ ...:

$$\hat{H}_R|n_1...n_i...\rangle = \sum_i \left(n_i+\frac{1}{2}\right)\hbar\omega_i|n_1...n_i...\rangle \tag{B-10a}$$

$$\hat{\boldsymbol{P}}_{\text{trans}}|n_1...n_i...\rangle = \sum_i n_i\hbar\boldsymbol{k}_i|n_1...n_i...\rangle \tag{B-10b}$$

The field's ground state corresponds to all the n_i equal to zero, and will be noted $|0\rangle$:

$$|0\rangle = |0_1...0_i...\rangle \tag{B-11}$$

the states $|n_1...n_i...\rangle$ being obtained by the action of a certain number of creation operators on this $|0\rangle$ state:

$$|n_1...n_i...\rangle = \frac{(\hat{a}_1^\dagger)^{n_1}}{\sqrt{n_1!}}...\frac{(\hat{a}_i^\dagger)^{n_i}}{\sqrt{n_i!}}...|0\rangle \tag{B-12}$$

With respect to the field ground state, the state $|n_1...n_i...\rangle$ has an energy $\sum_i n_i\hbar\omega_i$ and a momentum $\sum_i n_i\hbar\boldsymbol{k}_i$. It can be interpreted as describing an ensemble of n_1 particles of energy $\hbar\omega_1$ and momentum $\hbar\boldsymbol{k}_1$,....., n_i particles of energy $\hbar\omega_i$ and momentum $\hbar\boldsymbol{k}_i$, *etc...* . These particles characterize the elementary excitations of the quantum field and are called *photons*. The quantum number n_i is therefore the number of photons occupying the mode i, so that the ground state $|0\rangle$, corresponding to all the n_i equal to zero, can be called the photon *vacuum*.

Whereas there exists for photons eigenstates of momentum and energy, there are no quantum states of the electromagnetic field where the position can be perfectly known; no position operator is associated with this field. This is a different situation from what we encounter with massive particles, which have both a position and a momentum operator; the wave functions in the two representations are related by a simple Fourier transform. This non-existence of a position operator is linked to the impossibility of building, by a linear superposition of transverse electromagnetic waves, a vector wave perfectly localized at a point in space. The relativistic and transverse character of the electromagnetic field yields commutation relations between its components that involve the transverse delta function (Complement A_{XVIII}, § 2-e) instead of the usual delta function.

B-3. Several examples of quantum radiation states

We now study several examples of states of quantum radiation.

B-3-a. Photon vacuum

The presence of the 1/2 term in the parenthesis on the right-hand side of equation (B-10a) shows that the vacuum state energy is not zero, but equal to $\sum_i \hbar\omega_i/2$; this sum is an infinite quantity. We encounter here a first example of the difficulties linked to the divergences appearing in quantum electrodynamics. They can be resolved by *renormalization* techniques, whose presentation is outside the scope of this book. We shall avoid this difficulty by only considering energy differences with respect to the vacuum.

If we consider a single mode i of the field, the energy $\hbar\omega_i/2$ of the vacuum state for this mode is finite, and reminiscent of the zero-point energy of a harmonic oscillator of frequency ω_i. As you may recall, this zero-point energy is due to the impossibility of having simultaneously zero values for the position x and momentum p of that oscillator, because of the Heisenberg relations. The lowest energy state of the oscillator results from a compromise between the kinetic energy, proportional to p^2, and the potential energy, proportional to x^2 (this problem is discussed in § D-2 of Chapter V). The same arguments can be presented for the contribution, at a given point $\boldsymbol{r}$, of mode i to the electric $\hat{\boldsymbol{E}}_\perp(\boldsymbol{r},t)$ and magnetic $\hat{\boldsymbol{B}}(\boldsymbol{r},t)$ fields; according to (B-3) and (B-4), those fields are represented by two different linear superpositions of operators $\hat{a}_i$ and $\hat{a}_i^\dagger$, which thus do not commute. Consequently, one cannot have simultaneously a zero value for the electric energy proportional to $\hat{\boldsymbol{E}}_\perp^2$, and for the magnetic energy proportional to $\hat{\boldsymbol{B}}^2$.

One can further calculate the average value and variance of the contribution of mode i to the electric field $\hat{\boldsymbol{E}}_\perp(\mathbf{r},t)$ at point $\boldsymbol{r}$. Since $\hat{a}_i$ and $\hat{a}_i^\dagger$ change n_i by ± 1, a simple calculation yields:

$$\langle 0|\hat{\boldsymbol{E}}_\perp(\boldsymbol{r},t)|0\rangle_{\text{mode i}} = 0 \tag{B-13a}$$

$$\langle 0|\hat{\boldsymbol{E}}_\perp^2(\boldsymbol{r},t)|0\rangle_{\text{mode i}} = \frac{\hbar\omega_i}{2\varepsilon_0 L^3} \tag{B-13b}$$

Similar calculations can be done for the magnetic field. They show that in the photon vacuum state, the average value of both the electric and magnetic fields is zero, but not their variance. Since result (B-13b) is proportional to $\hbar$, the non-zero variance of the fields in the vacuum is a quantum effect.

Comments

(i) The summation over all the modes of expressions (B-13) yields, once we have replaced the discrete sum by an integral:

$$\langle 0|\hat{\boldsymbol{E}}_\perp(\boldsymbol{r},t)|0\rangle = 0 \tag{B-14a}$$

$$\langle 0|\hat{\boldsymbol{E}}_\perp^2(\boldsymbol{r},t)|0\rangle = \sum_i \frac{\hbar\omega_i}{2\varepsilon_0 L^3} = \frac{\hbar c}{2\varepsilon_0\pi^2}\int_0^{k_M} k^3 \mathrm{d}k \tag{B-14b}$$

This means that the variance of the electric field diverges as the fourth power of the upper boundary of the integral over k appearing in the summation of the modes of frequency $\omega = ck$. This divergence is the same as that mentioned above.

(ii) To characterize the dynamics of these field fluctuations, it is possible to compute the field correlation functions in vacuum[5]. This calculation shows that the electric and magnetic fields fluctuate very rapidly around their zero average value. These fluctuations are called the *vacuum fluctuations*. Certain radiative corrections, such as the "Lamb shift"

[5] See for example § III-C-3-c and Complement C_{III} of reference [16].

in atoms, can be interpreted from a physical point of view, as resulting from the vibration of the atom's electron caused by its interaction with this fluctuating electric field. This vibration leads the electron to explore the nucleus Coulomb potential over the range of its vibrational motion. The corresponding correction to its binding energy depends on the energy level it occupies; this explains why the degeneracy between the $2s_{1/2}$ and $2p_{1/2}$ states of the hydrogen atom, predicted by the Schrödinger and Dirac equations, can be lifted by the interaction with the vacuum fluctuations [6].

B-3-b. Field quasi-classical states

The state and observables of a classical field are characterized by the normal variables $\{\alpha_i\}$ introduced in § B-2-b of Chapter XVIII. The coherent states of a one-dimensional harmonic oscillator studied in Complement G_V, can be used to build the field quantum states whose properties are closest to those of the classical field $\{\alpha_i\}$.

The coherent state, supposed to be normalized, of a one-dimensional harmonic oscillator is the eigenstate of the annihilation operator $\hat{a}$, with eigenvalue α:

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \tag{B-15}$$

The eigenvalue α may be a complex number since operator $\hat{a}$ is not Hermitian. Equation (B-15) leads to:

$$\langle\alpha|\hat{a}|\alpha\rangle = \alpha \qquad \langle\alpha|\hat{a}^\dagger|\alpha\rangle = \alpha^* \tag{B-16}$$

More generally, the average value of any function of $\hat{a}$ and $\hat{a}^\dagger$, once put in the *normal* order, i.e. where all the annihilation operators are positioned to the right of the creation operators (Complement B_{XVI}, § 1-a-α), is equal to the expression obtained by replacing operator $\hat{a}$ by α and operator $\hat{a}^\dagger$ by α^*. As an example:

$$\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = \alpha^*\alpha \tag{B-17}$$

Consider then the field quantum state:

$$|\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \ldots |\alpha_i\rangle \ldots = |\alpha_1, \alpha_2 \ldots \alpha_i \ldots\rangle \tag{B-18}$$

where each mode i is in the coherent state $|\alpha_i\rangle$ corresponding to the classical normal variable α_i. Using equations (B-16) and (B-17), we can obtain the average values of the various field operators (B-3), (B-4) and (B-5) in the state (B-18); they coincide with the values of these various physical quantities for a classical field described by the normal variables $\{\alpha_i\}$. The same is true for the observables (B-6) and (B-7) corresponding to the energy and momentum of the transverse field. This is why the quantum state (B-18), which yields average values identical to all the properties of a classical field, is called a *quasi-classical state*[7]. We shall see later that the correlation functions of the quantum and classical fields involved in various photodetection signals also coincide when the field state is a quasi-classical state.

[6] See for example [17].

[7] For more details on the properties of the radiation quasi-classical states, see § III-C-4 of reference [16].

B-3-c. Single photon state

Consider the state vector:

$$|\Psi\rangle = \sum_i c_i |1_i\rangle \prod_{\otimes j \neq i} |0_j\rangle \tag{B-19}$$

which is a linear superposition of kets where a mode i contains one photon, whereas all the other modes $j \neq i$ are empty. Such a ket is an eigenket of the operator total number of photons $\hat{N} = \sum_i \hat{a}_i^\dagger \hat{a}_i$ with an eigenvalue equal to 1. It is therefore a single photon state. However, except in special cases, it is not a stationary state since it is not an eigenstate of the field energy $\hat{H}_R$. It describes a single photon propagating in space with velocity c. We shall see later (Complement $\mathrm{D_{XX}}$) that, when the field is in the state (B-19), a photodetector placed in a small region of space yields a signal corresponding to the passage, in that region, of a wave packet.

C. Description of the interactions

C-1. Interaction Hamiltonian

The Hamiltonian $\hat{H}$ of the system "particles + field" has been given above. In its expression (A-10), we now separate the terms that depend only on the particle variables or only on the field variables, and those that depend on both. We can then write $\hat{H} = \hat{H}_P + \hat{H}_R + \hat{H}_I$, where the particle Hamiltonian is:

$$\hat{H}_P = \sum_a \frac{\hat{\boldsymbol{p}}_a^2}{2m_a} + \hat{V}_{\mathrm{Coul}} \tag{C-1}$$

whereas the radiation one is:

$$\hat{H}_R = \sum_i \hbar\omega_i \left(\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2} \right) \tag{C-2}$$

Finally, the interaction Hamiltonian is the sum:

$$\hat{H}_I = \hat{H}_{I1} + \hat{H}_{I2}, \tag{C-3}$$

with:

$$\hat{H}_{I1} = -\sum_a \frac{q_a}{2m_a} \left[\hat{\boldsymbol{p}}_a \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a) + \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a) \cdot \hat{\boldsymbol{p}}_a \right] \tag{C-4}$$

$$\hat{H}_{I2} = \sum_a \frac{q_a^2}{2m_a} \left[\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a) \right]^2 \tag{C-5}$$

(we have separated the linear and quadratic terms with respect to the fields).

To that interaction Hamiltonian, we must further add the term:

$$\hat{H}_{I1}^s = -\sum_a \hat{\boldsymbol{M}}_a^s \cdot \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \tag{C-6}$$

describing the interaction of the spin magnetic moments of the various particles with the magnetic field of the radiation (Complement A_{XIII}, § 1-d):

$$\hat{\boldsymbol{M}}_a^s = g_a \frac{q_a}{2m_a} \hat{\boldsymbol{S}}_a \tag{C-7}$$

where g_a is the "Landé g-factor" of particle a whose spin is noted $\hat{\boldsymbol{S}}_a$.

Comment

Even with this additional term, all the possible interactions are not contained in that Hamiltonian: missing for example are the electron spin-orbit coupling, the hyperfine interaction between the electron and the nucleus, etc. – see comment (iii) of § C-5. The Hamiltonian we wrote is however sufficient in a great number of cases.

C-2. Interaction with an atom. External and internal variables

Consider the case where the particle system is a single atom, assumed to be neutral, formed by an electron e and a nucleus N which have opposite charges ($q_e = -q_N = q$) and whose masses are noted m_e and m_N. This is the case for example of the hydrogen atom. It is standard practice (see for example § B of Chapter VII) to separate the variables $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{P}}$ of the system's center of mass and the variables $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{p}}$ of the relative motion. These two types of variables commute with each other and are given by equations:

$$\begin{cases} \boldsymbol{R} = \dfrac{m_e\, \boldsymbol{r}_e + m_N\, \boldsymbol{r}_N}{M} \\ \boldsymbol{r} = \boldsymbol{r}_e - \boldsymbol{r}_N \end{cases} \qquad \begin{cases} \boldsymbol{P} = \boldsymbol{p}_e + \boldsymbol{p}_N \\ \dfrac{\boldsymbol{p}}{m} = \dfrac{\boldsymbol{p}_e}{m_e} - \dfrac{\boldsymbol{p}_N}{m_N} \end{cases} \tag{C-8}$$

where we have noted M the total mass of the system, and m its reduced mass:

$$M = m_e + m_N \qquad ; \qquad m = \frac{m_e m_N}{M} \tag{C-9}$$

Expressed as a function of these new variables, the particle Hamiltonian is written:

$$\hat{H}_P = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{\hat{\boldsymbol{p}}^2}{2m} + \hat{V}_{\text{Coul}}(\hat{r}) \tag{C-10}$$

The center of mass variables, also called *external variables*, describe the global motion of the atom, whereas the variables of the relative motion, also called *internal variables*, describe the motion in the center of mass reference frame.

C-3. Long wavelength approximation

The interaction Hamiltonians (C-4), (C-5) and (C-6) contain fields evaluated at the electron $\boldsymbol{r}_e$ and nucleus $\boldsymbol{r}_N$ positions. These positions can be described with respect to the position of the center of mass and we can write for example:

$$\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_e) = \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}} + \hat{\boldsymbol{r}}_e - \hat{\boldsymbol{R}}) \tag{C-11}$$

In an atom, the distance between the position of the electron or the nucleus and the atom's center of mass is of the order of the atom's size, i.e. just a fraction of a nanometer. Now

the radiation wavelengths that can have a resonant interaction with the atom are of the order of a fraction of a micron, much larger than the atomic dimension. One can thus neglect the variation of the fields over distances of the order of $|\boldsymbol{r}_e - \boldsymbol{R}|$ (or $|\boldsymbol{r}_N - \boldsymbol{R}|$) and write:

$$\begin{aligned} \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_e) &\simeq \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \\ \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_N) &\simeq \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \end{aligned} \tag{C-12}$$

Such an approximation is called the long wavelength approximation (or dipole approximation).

Using this approximation in the interaction Hamiltonian $\hat{H}_{I1}$, yields:

$$\begin{aligned} \hat{H}_{I1} &= -\frac{q_e}{m_e}\hat{\boldsymbol{p}}_e \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_e) - \frac{q_N}{m_N}\hat{\boldsymbol{p}}_N \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_N) \\ &\simeq -q\left[\frac{\hat{\boldsymbol{p}}_e}{m_e} - \frac{\hat{\boldsymbol{p}}_N}{m_N}\right] \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \\ &= -\frac{q}{m}\hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \end{aligned} \tag{C-13}$$

We used the relation $q_e = -q_N = q$ as well as definition (C-8) for the relative momentum.

As for Hamiltonian $\hat{H}_{I2}$, it becomes with this approximation:

$$\begin{aligned} \hat{H}_{I2} &= \frac{q_e^2}{2m_e}\hat{\boldsymbol{A}}_\perp^2(\hat{\boldsymbol{r}}_e) + \frac{q_N^2}{2m_N}\hat{\boldsymbol{A}}_\perp^2(\hat{\boldsymbol{r}}_N) \\ &\simeq \frac{q^2}{2m}\hat{\boldsymbol{A}}_\perp^2(\hat{\boldsymbol{R}}) \end{aligned} \tag{C-14}$$

Comment

When we include the Hamiltonian describing the spin magnetic coupling $\hat{H}_{I1}^s$ written in (C-6), we also replace all the $\hat{\boldsymbol{r}}_a$ by $\hat{\boldsymbol{R}}$. This is however insufficient: we must add other terms of the same order, obtained by including first order terms in $\boldsymbol{k}\cdot(\hat{\boldsymbol{r}}_e - \hat{\boldsymbol{R}})$ in $\hat{H}_{I1}$ and $\hat{H}_{I2}$, and representing corrections to the long wavelength approximation. This is because a computation analogous to the one in § 1-d of Complement AXIII shows that these corrections yield new interaction terms of the same order as $\hat{H}_{I1}^s$: interaction between the atomic orbital momentum $\boldsymbol{L}$ and the radiation magnetic field; electric quadrupole interaction.

C-4. Electric dipole Hamiltonian

Using the long wavelength approximation, the global Hamiltonian for the system atom + field is written:

$$\hat{H} = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{1}{2m}\left[\hat{\boldsymbol{p}} - q\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})\right]^2 + \hat{V}_{\text{Coul}} + \sum_i \frac{\hbar\omega_i}{2}\left[\hat{a}_i^\dagger \hat{a}_i + \hat{a}_i \hat{a}_i^\dagger\right] \tag{C-15}$$

We are going to perform a unitary transformation on this Hamiltonian, leading to a new interaction Hamiltonian, composed of a single term of the form $-\hat{\boldsymbol{D}} \cdot \hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$, where $\hat{\boldsymbol{D}}$ is the electric dipole moment of the atom:

$$\hat{\boldsymbol{D}} = q\,\hat{\boldsymbol{r}} \tag{C-16}$$

and $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$, the quantum field given by expression (B-3). This new interaction Hamiltonian is called the *electric dipole Hamiltonian*.

To find this unitary transformation, it is useful to start with the simpler case where the radiation field is treated classically.

C-4-a. Electric dipole Hamiltonian for a classical field

When the radiation field is treated classically, as an external field whose dynamic is externally imposed and hence has a fixed time dependence, the last term of relation (C-15) does not exist; operator $\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})$, which appears in the second term, must be replaced by the external field $\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)$. The system Hamiltonian is then written:

$$\hat{H} = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{1}{2m}\left[\hat{\boldsymbol{p}} - q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)\right]^2 + \hat{V}_{\text{Coul}} \tag{C-17}$$

We are looking for a unitary transformation that performs a translation of $\hat{\boldsymbol{p}}$ by a quantity $q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)$, so that the second term in (C-17) is reduced to $\hat{\boldsymbol{p}}^2/2m$. Such a transformation reads:

$$\hat{T}(t) = \exp\left[-\frac{i}{\hbar} q\,\hat{\boldsymbol{r}} \cdot \boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)\right] \tag{C-18}$$

We can check this since, using $[\hat{\boldsymbol{p}}, f(\hat{\boldsymbol{r}})] = -i\hbar\,\partial f/\partial\hat{\boldsymbol{r}}$ and the fact that the internal variable $\hat{\boldsymbol{r}}$ commutes with the external variable $\hat{\boldsymbol{R}}$, we have:

$$\hat{T}(t)\,\hat{\boldsymbol{p}}\,\hat{T}^\dagger(t) = \hat{\boldsymbol{p}} + q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t) \tag{C-19}$$

As they do not depend on $\hat{\boldsymbol{p}}$, the other terms of (C-17) are unchanged by the transformation. On the other hand, since this transformation has an explicit time dependence via the term $q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)$, the new Hamiltonian H' that governs the evolution of the new state vector:

$$|\Psi'(t)\rangle = \hat{T}(t)|\Psi(t)\rangle \tag{C-20}$$

is given by:

$$\hat{H}'(t) = \hat{T}(t)\hat{H}(t)\hat{T}^\dagger(t) + i\hbar\left[\frac{\mathrm{d}\hat{T}(t)}{\mathrm{d}t}\right]\hat{T}^\dagger(t) \tag{C-21}$$

As we have in addition:

$$i\hbar\left[\frac{\mathrm{d}\hat{T}(t)}{\mathrm{d}t}\right]\hat{T}^\dagger(t) = q\hat{\boldsymbol{r}} \cdot \frac{\partial \boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)}{\partial t} = -\hat{\boldsymbol{D}} \cdot \boldsymbol{E}_{\perp e}(\hat{\boldsymbol{R}}, t) \tag{C-22}$$

where $\hat{\boldsymbol{D}} = q\,\hat{\boldsymbol{r}}$ is the electric dipole moment of the atom, we finally obtain:

$$\hat{H}'(t) = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{\hat{\boldsymbol{p}}^2}{2M} + V_{\text{Coul}} - \hat{\boldsymbol{D}} \cdot \boldsymbol{E}_{\perp e}(\hat{\boldsymbol{R}}, t) \tag{C-23}$$

where the last term has the expected form for an electric dipole Hamiltonian.

C-4-b. Electric dipole Hamiltonian for a quantum field

The results we just obtained suggest using the unitary transformation:

$$\hat{T} = \exp\left[-\frac{i}{\hbar} q\, \hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \right] \tag{C-24}$$

where it is now the operator $\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})$ that appears in the exponential. One can check that this operator is still a translation operator for $\hat{\boldsymbol{p}}$, so that the second term in (C-15) is now simply of the form $\hat{\boldsymbol{p}}^2/2m$.

As the transformation (C-24) no longer has an explicit time dependence, the term analogous to (C-22) does not exist anymore. On the other hand, we must study the transformation of the last term of (C-15), which represents the energy $\hat{H}_R$ of the transverse quantum field. We therefore rewrite expression (C-24) using the expansion (B-5) of $\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})$ as a function of $\hat{a}_i$ and $\hat{a}_i^\dagger$:

$$\hat{T} = \exp\left[\sum_i \left(\lambda_i^* \hat{a}_i - \lambda_i \hat{a}_i^\dagger \right) \right] \tag{C-25}$$

with:

$$\lambda_i = \frac{i}{\sqrt{2\varepsilon_0 \hbar \omega_i L^3}}\, \boldsymbol{\varepsilon}_i \cdot \hat{\boldsymbol{D}}\; e^{-i\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}} \tag{C-26}$$

In this form, operator $\hat{T}$ does appear as a translation operator (Complement $\mathrm{G_V}$, § 2-d); it obeys the equations:

$$\hat{T}\, \hat{a}_j\, \hat{T}^\dagger = \hat{a}_j + \lambda_j \qquad \hat{T}\, \hat{a}_j^\dagger\, \hat{T}^\dagger = \hat{a}_j^\dagger + \lambda_j^* \tag{C-27}$$

To prove relations (C-27), one can use (Complement $\mathrm{B_{II}}$ $\mathrm{B_{II}}$, § 5-d) the identity:

$$e^{(A+B)} = e^A e^B e^{-[A,B]/2} \tag{C-28}$$

valid if A and B commute with their commutator $[A, B]$, as well as the commutation relation $\left[\hat{a}, f(\hat{a}^\dagger)\right] = \partial f / \partial \hat{a}^\dagger$. The transformation of the last term in (C-15) then yields:

$$\hat{T} \hat{H}_R \hat{T}^\dagger = \sum_i \frac{\hbar\omega_i}{2} \left[(\hat{a}_i + \lambda_i)(\hat{a}_i^\dagger + \lambda_i^*) + (\hat{a}_i^\dagger + \lambda_i^*)(\hat{a}_i + \lambda_i) \right] \tag{C-29}$$

The terms on the right-hand side of (C-29) that are independent of λ_i and λ_i^* yield again $\hat{H}_R$. The terms linear in λ_i and λ_i^* yield:

$$\begin{aligned} \sum_i \hbar\omega_i \left(\lambda_i \hat{a}_i^\dagger + \lambda_i^* \hat{a}_i \right) &= -\sum_i i \sqrt{\frac{\hbar\omega_i}{2\epsilon_0 L^3}} \left(\hat{a}_i \boldsymbol{\varepsilon}_i e^{i\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}} - \hat{a}_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}} \right) \cdot \hat{\boldsymbol{D}} \\ &= -\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) \cdot \hat{\boldsymbol{D}} \end{aligned} \tag{C-30}$$

where we have used (B-3). We thus get the expected electric dipole form for the interaction Hamiltonian:

$$\hat{H}_I = -\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) \cdot \hat{\boldsymbol{D}} \tag{C-31}$$

Finally, the terms quadratic in λ_i and λ_i^* introduce a term we shall note $\hat{h}_{\text{dip}}$:

$$\hat{h}_{\text{dip}} = \sum_i \hbar\omega_i \lambda_i^* \lambda_i = \sum_i \frac{1}{2\varepsilon_0 L^3} (\boldsymbol{\varepsilon}_i \cdot \hat{\boldsymbol{D}})(\boldsymbol{\varepsilon}_i^* \cdot \hat{\boldsymbol{D}}) \tag{C-32}$$

It represents a dipolar energy intrinsic to the atom.

To sum up, regrouping all the previous terms, we get for the transformed Hamiltonian:

$$\hat{H}' = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{\hat{\boldsymbol{p}}^2}{2M} + V_{\text{Coul}} + \hat{H}_R - \hat{\boldsymbol{D}} \cdot \hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) + \hat{h}_{\text{dip}} \tag{C-33}$$

This is a form similar to (C-23), with an additional term $\hat{h}_{\text{dip}}$.

Comments

(i) The same mathematical operator does not describe the same physical quantity in two different representations, deduced from one another by a unitary transformation. As an example, the operator $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ appearing in (C-31), does not represent the transverse electric field in the new point of view, which should be $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ transformed by $\hat{T}$, written as $\hat{T}\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})\hat{T}^\dagger$, and hence different from $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$. Actually, one can show that the operator $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ represents in the new point of view the physical quantity $\hat{\mathfrak{D}}(\hat{\boldsymbol{R}})/\varepsilon_0$ where $\hat{\mathfrak{D}}(\hat{\boldsymbol{R}})$ is called the *electric displacement field* (see Complement A_{IV} of [16]).

(ii) The intrinsic dipolar energy $\hat{h}_{\text{dip}}$ is given by an integral over k, which diverges at infinity. This integral must however be limited to values of k for which the long wavelength approximation is still valid.

C-5. Matrix elements of the interaction Hamiltonian; selection rules

Consider an initial state where the atom is described by $|\psi_{\text{in}}^{\text{int}}\rangle$ for its internal state, $|\psi_{\text{in}}^{\text{ext}}\rangle$ for its external state, and where the radiation is in the state $|\psi_{\text{in}}^R\rangle$. The interaction Hamiltonian (C-31) couples this initial state to a final state where the atomic internal and external variables, as well as the radiation variables are respectively in the states $|\psi_{\text{fin}}^{\text{int}}\rangle$, $|\psi_{\text{fin}}^{\text{ext}}\rangle$, and $|\psi_{\text{fin}}^R\rangle$. As the operator $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ appearing in (C-31) is a linear superposition of annihilation $\hat{a}_i$ and creation $\hat{a}_i^\dagger$ operators, the matrix element of $\hat{H}_I$ describes two types of processes: the absorption processes associated with operator $\hat{a}_i$ where one photon disappears, and the emission processes associated with operator $\hat{a}_i^\dagger$ where a new photon appears. This matrix element can be factored into a product of three matrix elements concerning the three types of variables; they are written, for the absorption processes:

$$i\sqrt{\frac{\hbar\omega_i}{2\varepsilon_0 L^3}} \left\langle \psi_{\text{fin}}^{\text{int}} \right| \boldsymbol{\varepsilon}_i \cdot \hat{\boldsymbol{D}} \left| \psi_{\text{in}}^{\text{int}} \right\rangle \left\langle \psi_{\text{fin}}^{\text{ext}} \right| \exp(i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}) \left| \psi_{\text{in}}^{\text{ext}} \right\rangle \left\langle \psi_{\text{fin}}^R \right| \hat{a}_i \left| \psi_{\text{in}}^R \right\rangle \tag{C-34}$$

and for the emission processes:

$$-i\sqrt{\frac{\hbar\omega_i}{2\varepsilon_0 L^3}} \left\langle \psi_{\text{fin}}^{\text{int}} \right| \boldsymbol{\varepsilon}_i^* \cdot \hat{\boldsymbol{D}} \left| \psi_{\text{in}}^{\text{int}} \right\rangle \left\langle \psi_{\text{fin}}^{\text{ext}} \right| \exp(-i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}) \left| \psi_{\text{in}}^{\text{ext}} \right\rangle \left\langle \psi_{\text{fin}}^R \right| \hat{a}_i^\dagger \left| \psi_{\text{in}}^R \right\rangle \tag{C-35}$$

The central term in these expressions is a matrix element concerning the external atomic variables; it expresses the conservation of the global momentum as we now

show. Operator $\exp(\pm i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}})$ translates the momentum by a quantity $\pm\hbar\boldsymbol{k}_i$. If the atom's center of mass has an initial momentum $\hbar\boldsymbol{K}_{\text{in}}$, once it absorbs a photon, its final momentum will be $\hbar\boldsymbol{K}_{\text{fin}} = \hbar\boldsymbol{K}_{\text{in}} + \hbar\boldsymbol{k}_i$; the momentum $\hbar\boldsymbol{k}_i$ of the absorbed photon is therefore transferred to the atom during the absorption process. In a similar way, one can show that the atom's momentum decreases by the quantity $\hbar\boldsymbol{k}_i$ when a photon is emitted.

In the first matrix element of (C-34), which concerns the internal atomic variables, operator $\hat{\boldsymbol{D}}$ is an odd operator. The matrix element will be different from zero only if the initial and final internal atomic states have opposite parity, as for instance the $1s$ and $2p$ states of the hydrogen atom. We rediscover here a second conservation law, the conservation of parity. In addition, as the operator $\hat{\boldsymbol{D}}$ is a vector operator, it leads to selection rules on the internal angular momentum which will be studied in Complement C_{XIX}.

Comments

(i) The conservation of the total momentum comes from the central matrix elements in expressions (C-34) and (C-35). One may wonder whether this result is only valid for the approximate form (C-31) of the interaction Hamiltonian used to establish those equations. Actually it can be shown, using the commutation relations $[\boldsymbol{p}_a, F(\boldsymbol{r}_a)] = -i\hbar\partial F/\partial \boldsymbol{r}_a$ and $[\hat{a}_i^\dagger \hat{a}_i, \hat{a}_i] = -\hat{a}_i$, that the interaction Hamiltonian $\hat{H}_{I1}$ written is (C-4) (without the long wavelength approximation) commutes with the system total momentum $\sum_a \hat{\boldsymbol{p}}_a + \sum_i \hbar\boldsymbol{k}_i \hat{a}_i^\dagger \hat{a}_i$. The same result is true for all the terms of the interaction Hamiltonian. Consequently, the exact (without approximation) interaction Hamiltonian has non-zero matrix elements only between states having the same total momentum. The fact that the total momentum commutes with all the terms in the Hamiltonian is related to the system invariance with respect to spatial translation. The properties of the system are unchanged upon the translation by the same quantity of the particles and the fields. Similar considerations apply to the rotational invariance and cause the interaction Hamiltonian to only connect states with the same total angular momentum. These results are important for understanding in a simple fashion the exchanges of linear and angular momenta between atoms and photons, which will be discussed in Complements A_{XIX} and C_{XIX}.

(ii) Conservation of total momentum during the absorption process, combined with total energy conservation, shows that the energy of the absorbed photon is different from the energy separating the two internal levels involved in the transition. Two effects account for this difference: the Doppler effect, and the recoil effect (Complement A_{XIX}); they play an important role in laser cooling methods.

(iii) If we continue the calculations beyond the long wavelength approximation, we find additional terms for the interaction Hamiltonian, describing the interaction between the radiation magnetic field and the atomic orbital or spin magnetic moments (Complement A_{XIII}, § 1). Some of these terms have already been written in (C-6). Transitions, called magnetic dipole transitions, may occur between levels having the same parity, as opposed to the electric dipole transitions studied above. Other types of transitions may also be observed at higher orders, such as the quadrupole transitions.

Note finally that, if the initial radiation state already contains n_i photons, the last two matrix elements of (C-34) and (C-35) are equal to $\langle n_i - 1|\hat{a}_i|n_i\rangle = \sqrt{n_i}$ and $\langle n_i + 1|\hat{a}_i^\dagger|n_i\rangle = \sqrt{n_i + 1}$. In the presence of n_i incident photons, the probability of the absorption process is thus proportional to n_i, whereas the emission probability is

proportional to $n_i + 1$. We shall see in Chapter XX that this difference is linked to the existence of two types of emission, the stimulated emission and the spontaneous emission.

With the knowledge of the various Hamiltonians $\hat{H}_A$, $\hat{H}_R$ and $\hat{H}_I$, as well as their matrix elements, we can now solve Schrödinger's equation to compute the transition amplitude between an initial state and a final state of the system "atom + field". This will be done in the next chapter, where we study various processes, such as the absorption or emission of photons for an incident radiation either monochromatic or having a large spectral band, the photoionization phenomenon, multiphoton processes and photon scattering.

COMPLEMENTS OF CHAPTER XIX, READER'S GUIDE

The processes of photon absorption and emission by atoms must obey conservation laws for the total linear or angular momentum. This has implications that are highlighted in the three complements of this chapter. When an atom absorbs (or emits) a photon, it gains (or loses) an energy, a momentum and an angular momentum equal to the energy, momentum and angular momentum of the photon. This allows "manipulating" several properties of the atoms. It is, for example, at the base of optical pumping and laser cooling methods.

A_{XIX} : MOMENTUM EXCHANGE BETWEEN ATOMS AND PHOTONS	Momentum exchange between atoms and photons plays an important role in determining, for example, the Doppler effect and the shape of the spectral lines emitted or absorbed by gases. As an atom continually absorbs and re-emits photons, its momentum can be greatly affected. The atom can be decelerated, which allows slowing down and even bringing to rest an atomic beam over short distances. Other uses of the Doppler effect include the introduction of a friction force that slows down atoms to form ultra-cold gases. When the atoms are confined in a trap, the nature of the Doppler shift may be greatly changed and even completely disappear (Mössbauer effect). Multiphoton processes, in which the total momentum of the absorbed photons is zero, are also discussed.
B_{XIX} : ANGULAR MOMENTUM OF RADIATION	This complement is more technical than the previous one. It shows how the photon can be seen as a spin 1 particle; this particle also has an orbital angular momentum. The photon thus possesses two types of angular momentum. These concepts are useful for reading the next complement.
C_{XIX} : ANGULAR MOMENTUM EXCHANGE BETWEEN ATOMS AND PHOTONS	This complement studies the exchanges between the photon spin angular momentum (related to the light beam polarization) and the internal degrees of freedom of the atoms. These exchanges obey selection rules for the atomic transitions. Such rules are essential for many experimental methods in atomic physics, such as "optical pumping" where a polarized light beam allows accumulating the atoms of a gas in specific Zeeman sublevels. A number of applications of these methods are briefly reviewed in this complement.

Complement A_{XIX}

Momentum exchange between atoms and photons

1 Recoil of a free atom absorbing or emitting a photon 2020
1-a Conservation laws 2020
1-b Doppler effect, Doppler width 2022
1-c Recoil energy 2023
1-d Radiation pressure force in a plane wave 2024
2 Applications of the radiation pressure force: slowing and cooling atoms 2025
2-a Deceleration of an atomic beam 2025
2-b Doppler laser cooling of free atoms 2026
2-c Magneto-optical trap 2034
3 Blocking recoil through spatial confinement 2036
3-a Assumptions concerning the external trapping potential . . . 2036
3-b Intensities of the vibrational lines 2037
3-c Effect of the confinement on the absorption and emission spectra 2038
3-d Case of a one-dimensional harmonic potential 2039
3-e Mössbauer effect 2040
4 Recoil suppression in certain multi-photon processes 2040

We studied, in § C of Chapter XIX, the matrix elements of the interaction Hamiltonian between an atom and a field. This led us to establish selection rules based on the conservation of the total momentum of the system "atom + field" during the absorption or emission of photons by the atom. The study in Chapter XX of the absorption and emission amplitudes will show that the global energy of the system is also conserved during these processes. The goal of this complement is to show how these conservation laws[1] provide an interesting view on many aspects of momentum exchange between atoms and photons.

We start in § 1 with the case of a free atom, whose center of mass is not submitted to any external potential. We shall establish the expressions for the Doppler shift and the recoil energy that appear in the equation yielding the frequency of the absorbed or emitted photons. In a gas containing a large number of atoms with different velocities, the dispersion of these velocities yields a Doppler broadening of the emission and absorption lines. This broadening, as well as the shift related to the recoil energy, introduce perturbations in the lines observed using high resolution spectroscopy, hence limiting its precision. In addition, when the atom constantly absorbs and re-emits lots of photons, its momentum change per unit time can become very large. This results in a force generated

[1] The first study along this line concerned the Compton effect (1922), where the scattering of a photon by an electron is considered as a collision between two particles. Assigning the photon an energy $h\nu$ and a momentum $\hbar\boldsymbol{k}$, with $|\boldsymbol{k}| = 2\pi\nu/c$, one writes the conservation equations for the total momentum and energy during the collision. This yields the change of frequency of the photon as it is scattered in a given direction, in complete agreement with experimental observations.

by the radiation pressure. We shall calculate the order of magnitude of that force and show that it can produce an acceleration or deceleration of the atom a hundred thousand times larger than the one due to gravity.

In § 2, we will show that this force is able to slow down and immobilize a beam of atoms propagating in the direction opposed to that of the light beam. The velocity dependence of the radiation pressure force, due to the Doppler effect, is also very interesting. It allows, using two light beams in opposite directions, but with the same intensity and frequency, to generate a friction force on the atom, provided the light frequency is lower than the atomic frequency: the radiation pressure force is zero when the atomic velocity v is zero, but opposite to v when it is different from zero, therefore producing a damping of that velocity. This is the principle of one of the first laser cooling mechanisms observed experimentally. The very low temperatures obtained, millions of times lower than room temperature, explain the increasing number of application of the ultra-cold atoms thus obtained. We also explain in § 2-c the principle of the magneto-optical trap, which involves a position dependence of the radiation pressure force.

We describe in §§ 3 and 4 of this complement a number of methods developed to suppress or circumvent the effect related to the recoil. Confining the atom in a trap, as studied in § 3, may prevent the atom's recoil if the trapping is strong enough. If the transition is multi-photonic, for example if two photons having the same energy but opposite momenta are absorbed in the transition, no recoil is experienced by the atom, and there is no Doppler shift. An important example of this method is the Doppler-free two-photon spectroscopy (§ 4).

1. Recoil of a free atom absorbing or emitting a photon

Consider first an atom that is not subjected to any external potential (free atom). The Hamiltonian $\hat{H}_{\text{ext}}$ for the external variables is reduced to the kinetic energy term $\hat{\boldsymbol{P}}^2/2M$, where $\hat{\boldsymbol{P}}$ is the momentum of the center of mass and M the total mass of the atom. The eigenstates of $\hat{H}_{\text{ext}}$ can be chosen as states having well defined momentum $\boldsymbol{P}$ and energy $P^2/2M$.

1-a. Conservation laws

Let us express the radiation field as a function of plane waves of wave vector $\boldsymbol{k}$ and frequency $\omega = ck$. The eigenstates of the radiation Hamiltonian $\hat{H}_R$ can be described in terms of photons having an energy $\hbar\omega$ and a momentum $\hbar\boldsymbol{k}$. The interaction Hamiltonian $\hat{H}_I$ studied in § C of Chapter XIX is invariant[2] under spatial translation of the total system "atom + field". Consequently, it commutes with the system's total momentum, and can induce transitions only between states having the same total momentum. Furthermore, the transition amplitude associated with an interaction lasting a time T can only connect states of the total system having the same total energy, within $\hbar/T$ (this point will be further discussed in the next chapter). These two conservation laws can be used to study the influence of the motion of the center of mass on the frequencies of the photons it can absorb or emit.

[2]The interaction Hamiltonian involves field values at points where particles are located. It is therefore invariant when both fields and particles are shifted (by the same quantity).

Consider first an absorption process where the atom goes from an internal state a to another internal state b by absorbing a photon of energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$. We shall note:

$$\hbar\omega_0 = E_b - E_a \tag{1}$$

the energy difference between the two internal states. The initial (final) momentum of the center of mass before (after) the photon absorption is noted $\boldsymbol{P}_{\rm in}$ ($\boldsymbol{P}_{\rm fin}$). The conservation of the total momentum leads to:

$$\boldsymbol{P}_{\rm fin} = \boldsymbol{P}_{\rm in} + \hbar\boldsymbol{k} \tag{2}$$

The total energy in the initial state is the sum of the photon energy, the internal energy of the atom and its translation energy $P_{\rm in}^2/2M$:

$$E_{\rm in} = \hbar\omega + E_a + \frac{P_{\rm in}^2}{2M} \tag{3}$$

The total energy in the final state reduces to the atom's energy since the photon has been absorbed:

$$E_{\rm fin} = E_b + \frac{P_{\rm fin}^2}{2M} \tag{4}$$

The conservation of the total energy means that $E_{\rm in} = E_{\rm fin}$. Using (1) and (2), this last relation reads:

$$\hbar\omega = \hbar\omega_0 + \frac{\hbar\boldsymbol{k}\cdot\boldsymbol{P}_{\rm in}}{M} + \frac{\hbar^2k^2}{2M} \tag{5}$$

The last two terms in (5) represent the variation of the external energy of the atom during the transition – i.e. the variation of the atom's center of mass kinetic energy between the final state where it is equal to $(\boldsymbol{P}_{\rm in} + \hbar\boldsymbol{k})^2/2M$, and the initial state, where it is equal to $\boldsymbol{P}_{\rm in}^2/2M$. This equation can be rewritten as:

$$\omega = \omega_0 + \boldsymbol{k}\cdot\boldsymbol{v}_{\rm in} + \frac{E_{\rm rec}}{\hbar} \tag{6}$$

where $\boldsymbol{v}_{\rm in} = \boldsymbol{P}_{\rm in}/M$ is the initial velocity of the center of mass, and where:

$$E_{\rm rec} = \frac{\hbar^2k^2}{2M} = \hbar\omega_{\rm R} \tag{7}$$

is the *recoil energy*; it is the energy an atom, initially at rest, will acquire upon the absorption of a photon having a momentum $\hbar k$.

The same type of calculation holds for the emission process during which an atom, whose center of mass has an initial momentum $\boldsymbol{P}_{\rm in}$, goes from the internal state b to the internal state a by emitting a photon of energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$. Equation (6) must now be replaced by:

$$\omega = \omega_0 + \boldsymbol{k}\cdot\boldsymbol{v}_{\rm in} - \frac{E_{\rm rec}}{\hbar} \tag{8}$$

where only one sign is changed with respect to (6).

1-b. Doppler effect, Doppler width

The term $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ in equations (6) and (8) is simply the Doppler shift, due to the motion of the atom, of the frequencies it absorbs or emits. Setting:

$$\omega - \omega_0 = 2\pi(\nu - \nu_0) = 2\pi\Delta\nu \qquad \text{and} \qquad k = 2\pi\frac{\nu}{c} \tag{9}$$

we get for the frequency variation $\Delta\nu$ due to the Doppler effect:

$$\frac{\Delta\nu}{\nu} = \frac{\boldsymbol{\kappa} \cdot \boldsymbol{v}_{\text{in}}}{c} \tag{10}$$

where $\boldsymbol{\kappa} = \boldsymbol{k}/k$ is the unit vector defining the propagation direction of the photon. Note that a radiation quantum theory is not needed to account for this frequency shift, which can be predicted by a classical theory. This was to be expected since, among the last two terms of (6) and (8), $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ is the only one that does not go to zero when $\hbar$ tends toward zero, as opposed to $E_{\text{rec}}/\hbar = \hbar k^2/2M$ (which is proportional to $\hbar$).

For an ensemble of atoms in a dilute gas at thermal equilibrium at temperature T, the velocities are distributed according to the Maxwell-Boltzmann law, and the velocity dispersion Δv is of the order of $\sqrt{k_B T/M}$, where k_B is the Boltzmann constant. The shifts $\Delta\nu$ of the frequencies emitted or absorbed by the atoms are distributed following a Gaussian curve[3] whose width $\Delta\nu_D$ (standard deviation of the frequency distribution, equal to the square root of the variance), called the *Doppler width*, is given by:

$$\frac{\Delta\nu_D}{\nu_0} = \sqrt{\frac{k_B T}{Mc^2}} \tag{11}$$

In general, in the optical domain and for temperatures around 300 K, the Doppler width $\Delta\nu_D$ is of the order of 1GHz=10^9 Hz, much smaller than the frequency ν_0 (of the order of 10^{15} Hz), but much larger than the natural width Γ, of the order of 10^7 Hz. In this domain, the resolution of spectroscopic measurements of line frequencies emitted by a dilute gas is generally limited by the Doppler broadening of the lines.

Relativistic Doppler effect

The previous calculations are only valid in the non-relativistic limit ($v \ll c$). The Doppler shift expression can be generalized to any value of v by noting that the four quantities $\{k_x, k_y, k_z, \omega/c\}$ are the four components of a four-vector. Let us assume the atom is at rest in a reference frame R and emits a photon of frequency ω along the x axis (we ignore here the recoil energy). An observer, in a reference frame R' moving with velocity v along the x axis, sees the atom moving away at velocity $-v$ and measures a frequency ω' for the photon emitted by the atom. According to the relativistic expressions for the transformations of four-vector components, we have:

$$\omega' = \frac{\omega - kv}{\sqrt{1 - v^2/c^2}} \tag{12}$$

To first order in v/c, replacing $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ by $-kv$, we again find the Doppler shift of equation (8) – the relativistic correction being (in relative value) of the order of v^2/c^2 for $v \ll c$.

[3] Actually, this distribution is the convolution of a Gaussian and a curve of width Γ, where Γ is the natural width (due to the spontaneous emission) of the line emitted or absorbed by the atoms. For a gas at room temperature, Γ is much smaller than the Doppler width.

1-c. Recoil energy

Imagine the atom is initially at rest, so that the terms in $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ in (6) and (8) are zero. When the atom absorbs a photon, its momentum increases by a quantity $\hbar\boldsymbol{k}$ equal to the momentum of the absorbed photon. Consequently, the atom recoils with a velocity $v_{\text{rec}} = \hbar k/M$ in the direction of the incident photon, and its kinetic energy becomes $E_{\text{fin}} = Mv_{\text{rec}}^2/2 = E_{\text{rec}}$. The energy $\hbar\omega$ of the incident photon is used both to increase the internal energy of the atom by $\hbar\omega_0$ (since the atom goes from a to b), and to increase its kinetic energy from 0 to E_{rec}. We then have $\hbar\omega = \hbar\omega_0 + E_{\text{rec}}$, which is a particular case of (6) for a zero initial velocity. However, for the emission process where the atom goes from b to a by emitting a photon, this relation is modified. As the photon leaves the atom with a momentum $\hbar\boldsymbol{k}$, the atom recoils with the same momentum but in the opposite direction, and acquires a kinetic energy E_{rec}. The loss of internal energy of the atom, equal to $\hbar\omega_0$, must now be used both to increase the radiation energy by $\hbar\omega$ (the energy of the emitted photon), and to increase the kinetic energy of the atom from zero to E_{rec}. We then have $\hbar\omega_0 = \hbar\omega + E_{\text{rec}}$, that is $\hbar\omega = \hbar\omega_0 - E_{\text{rec}}$, which coincides with (8) with the minus sign on the right-hand side.

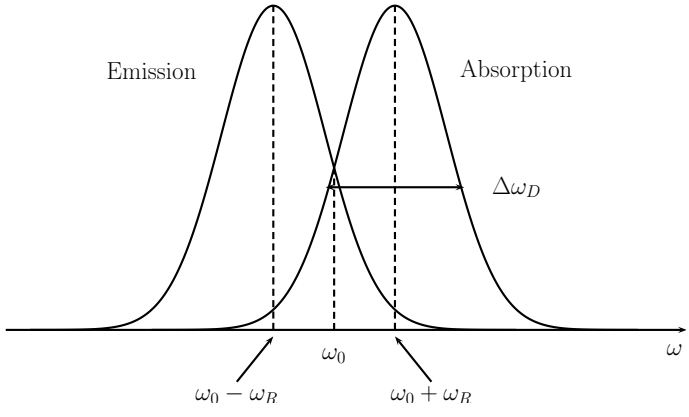


Figure 1: Because of the recoil effect of the atom, the absorption and emission lines do not coincide, but form a doublet called the recoil doublet*; they are centered at $\omega + \omega_R$ for the absorption line and $\omega - \omega_R$ for the emission line. In a gas, their width is the Doppler width $\Delta\omega_D$.*

The recoil of the atom with a kinetic energy E_{rec} causes the centers of the absorption and emission lines to be different (Figure 1); their position is $\omega_0 + \omega_R$ for the absorption line and $\omega_0 - \omega_R$ for the emission line, with $\omega_R = E_{\text{rec}}/\hbar$. For a gas containing a large number of moving atoms, the velocity dispersion around a zero average gives each of these two lines a Doppler width $\Delta\omega_D$. In the optical domain, the recoil frequency ω_R is of the order of a few kHz, much smaller than the Doppler width at room temperature, and also smaller than the natural width. Consequently the recoil doublet shown in Figure 1 is not resolved: the distance between the centers of the two lines is less than their width. However, when one studies the lines emitted by nuclei in the X-ray or γ-ray domains, the recoil frequency (which increases as k^2) becomes comparable to or

even larger than the Doppler width (which only increases as k). The two lines in Figure 1 then only overlap far out on their wings. In this case, the photon emitted by a nucleus in an excited state b has very little chance of being absorbed by another identical nucleus in the lower state a. We shall see later how the recoil of a nucleus can be blocked when the atom having that nucleus is sufficiently strongly bound to other atoms in a crystal (Mössbauer effect).

1-d. Radiation pressure force in a plane wave

Each time an atom absorbs a photon, it gains a momentum $\hbar\mathbf{k}$. If $\dot{N}_{\text{abs}}$ is the number of absorptions per unit time, the atom gains a momentum $\dot{N}_{\text{abs}}\hbar\mathbf{k}$, per unit time. In a steady state, the number $\dot{N}_{\text{abs}}$ of absorptions per unit time is equal to the number of emissions per unit time $\dot{N}_{\text{em}}$. This latter number is equal to $\Gamma\sigma_{bb}$, where Γ is the natural width of the atom's excited state, and where the diagonal element σ_{bb} of the atom's density operator in b is the occupation probability of that state. This gain in momentum per unit time of the atom can be considered as coming from the action of a force, associated with the radiation pressure exerted by the light beam on the atom. One often calls this force the "radiation pressure force". According to what we just saw, it is equal to[4]:

$$\mathbf{F} = \dot{N}_{\text{abs}}\hbar\mathbf{k} = \Gamma\sigma_{bb}\hbar\mathbf{k} \tag{13}$$

To get an idea of the order of magnitude of this force, let us assume the light intensity is very high; the atomic transition is then saturated, meaning the occupation probabilities σ_{bb} and σ_{aa} of the higher level b and the lower level a are both equal to $1/2$. We then have:

$$\mathbf{F} = \hbar\mathbf{k}\frac{\Gamma}{2} \tag{14}$$

Such a force can communicate an acceleration $\mathbf{A}$ equal to $\mathbf{F}/M$ to the atom of mass M. Taking (14) into account, this acceleration is equal to:

$$\mathbf{A} = \frac{\hbar\mathbf{k}}{M}\frac{\Gamma}{2} = \frac{\mathbf{v}_{\text{rec}}}{2\tau} \tag{15}$$

where $\mathbf{v}_{\text{rec}} = \hbar\mathbf{k}/M$ is the recoil velocity of an atom absorbing or emitting a photon, and $\tau = 1/\Gamma$ is the radiative lifetime of the excited state b.

Let us calculate the value of this acceleration for a sodium atom. The recoil velocity is of the order of 3×10^{-2}m/s and the radiative lifetime of the order of 16.2×10^{-9}s, meaning the acceleration A is of the order of 10^6m/s^2, which is 10^5 times larger (!) than the acceleration due to gravity (of the order of 10m/s^2). This high value for the acceleration A arises because the velocity change of the atom v_{rec} for each absorption-emission cycle, though very small by itself, accumulates during the very large number of cycles, $1/2\tau$, occurring in one second.

[4]To compute the momentum change, we only considered here the photon absorption processes. The spontaneous emission processes also change the atom's momentum, as the atom recoils when it emits a photon. However, the spontaneously emitted photons can have any direction in space, and the momentum change of the atom has a zero average. This phenomenon, however, gives rise to a diffusion of momentum, hence increasing the velocity dispersion of the atoms. We shall see in § 2-b-γ that this momentum diffusion must be taken into account when evaluating the limits of laser cooling.

2. Applications of the radiation pressure force: slowing and cooling atoms

We now study three applications of the radiation pressure force using either one laser beam, or two laser beams with the same intensity and same frequency, propagating in opposite directions along the x axis. We shall see in § 2-a how, with just a single laser beam, the radiation pressure force exerted by the beam on the atoms in an atomic beam propagating in the opposite direction can be used to slow down, and even bring to rest, the atoms of the beam. With two laser beams propagating in opposite directions, interesting effects occur if one can introduce an imbalance between the two radiation pressure forces, depending either on the atom's velocity v along the x axis, or on its position x. We study in § 2-b how to get this velocity dependence, and how the sum F of the two forces exerted by the two waves, zero for $v = 0$, becomes, for $v \neq 0$ but sufficiently small, a linear function of v; it can then be expressed as $F \simeq \gamma v$. For a proper detuning of the lasers' frequency, the coefficient γ can be negative, so that the resulting force F acts as a friction force that damps the velocities of all the atoms in the beam, and hence cools them down. This is the principle of laser cooling. In § 2-c, we shall see how a position dependence can be obtained, and how the resulting force, zero for $x = 0$, becomes different from zero if $x \neq 0$ and equal to βx if x is sufficiently small. If β is negative, this force becomes a restoring force that can trap the atoms around $x = 0$. This is the principle of the magneto-optical trap.

2-a. Deceleration of an atomic beam

Imagine that an atomic beam is irradiated by a resonant laser beam propagating in the direction opposite to that of the beam. Due to the radiation pressure force, the atoms of the beam will slow down. Is it possible to bring them to rest? It is important to notice that even if the laser beam is initially resonant, it will not be when the atoms' velocities change, since the Doppler effect takes the atoms out of resonance; this will significantly lower the radiation pressure force, and hence the slowing down effect. For the sake of simplicity, we shall first ignore the Doppler effect following the change in the atoms' velocities; we shall see later how it can actually be circumvented.

We first assume the radiation pressure force does not change in the course of the deceleration process, and that the laser intensity is high enough for the atomic transition to be saturated; we can then use the orders of magnitude calculated in the previous paragraph. We found that for sodium atoms, the deceleration A is of the order of $10^6 \mathrm{m/s^2}$. If the atoms of the beam have an initial velocity of the order of $10^3 \mathrm{m/s}$, their velocity will be zero after a time T of the order of $10^{-3}\mathrm{s}$, after they traveled a distance $AT^2/2$ of the order of 0.5m, which shows that the size of such an experiment is not, a priori, excessive.

To solve the problem of the atoms going out of resonance because of the Doppler effect which changes in the course of the slowing down, an ingenious method was proposed and demonstrated [18]. It is based on the propagation of the atoms in a spatially inhomogeneous magnetic field. More precisely, the atomic beam travels along the axis of a spatially varying solenoid coil (Figure 2). The magnetic field produced by the solenoid is parallel to the beam direction, and its intensity varies along the beam axis. As an atom propagates along this field, it undergoes a variable Zeeman shift of its resonance frequency. One can then adjust the profile of the field so that as the atom is slowed down, the Zeeman shift of the atomic frequency balances the Doppler shift of the laser

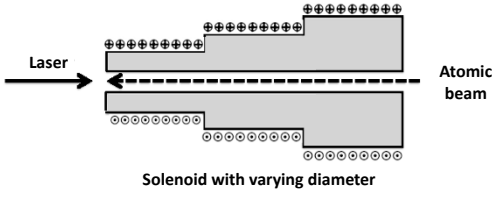


Figure 2: Schematic diagram of a Zeeman slower. The atomic beam is cooled by a laser beam propagating in the opposite direction. It travels along the axis of a solenoid composed of a set of magnetic coils with decreasing diameters, whose cross sections are shown in the figure (the current in the coils flows perpendicularly to the figure). While they propagate, the atoms are submitted to a larger and larger magnetic field. The Zeeman shift of their resonance frequency can thus follow the Doppler shift of the apparent laser frequency in their own reference frame. Consequently, instead of going off resonance, they can be slowed down during their entire propagation through the solenoid, and even come to rest.

frequency: at each point x, the field is calculated so that both Doppler and Zeeman shifts exactly balance each other. This type of apparatus is called a "Zeeman slower".

2-b. Doppler laser cooling of free atoms

α. Doppler laser cooling principle

The "slower" described above concerns the mean velocity of atoms, which can be brought down to zero. However, the root mean square of the atomic velocities remains non-zero, as does the temperature which is characterized not by the mean velocity but by its root mean square. We now describe a method using the velocity dependence of the radiation pressure force, due to the Doppler effect, and which permits reducing the dispersion of the atomic velocities around their mean value, and hence really cooling down the atoms. As this method uses the Doppler effect, it is called "Doppler laser cooling". It was proposed in 1975 for free atoms [19] and for trapped ions [20]. Here, we shall only study the case of free atoms.

The idea is to use two laser waves 1 and 2, having the same angular frequency ω and the same intensity, counterpropagating along the Ox axis, wave 1 toward negative x, and wave 2 toward positive x (Figure 3). Imagine an atom also propagating along the Ox axis with, for example, a positive velocity v. We call ω_0 the angular frequency of the atomic transition $a \leftrightarrow b$ excited by the laser. We assume the lasers are "red-detuned", meaning that:

$$\omega < \omega_0 \tag{16}$$

In the reference frame where the atom is at rest, the apparent frequency of wave i (with $i = 1, 2$) is Doppler shifted, and equal to $\omega - \mathbf{k_i} \cdot \mathbf{v}$. As v is positive, the atom and wave 1 propagate in opposite directions, so that $\mathbf{k_1} \cdot \mathbf{v}$ is negative. The apparent frequency

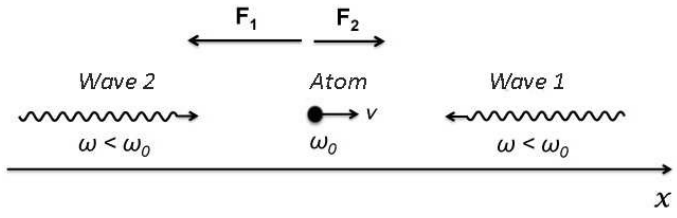


Figure 3: Principle of Doppler laser cooling. An atom moving with velocity v along the Ox axis interacts with two laser waves 1 and 2 having the same angular frequency ω; this frequency is "red-detuned", meaning $\omega < \omega_0$. The two waves have the same intensity but propagate in opposite directions along the Ox axis. The direction of the velocity v is opposite to that of wave 1. In the atom's frame of reference and because of the Doppler effect, the frequency of wave 1 is shifted closer to resonance, whereas the frequency of wave 2 is shifted away from it. Consequently, the modulus of force F_1 exerted by wave 1 increases, whereas that of the force F_2 exerted by wave 2 decreases. The resulting force, zero for $v = 0$, is opposed to v when $v \neq 0$, and proportional to v for small enough values of v. It is therefore equivalent to a friction force.

$\omega - \mathbf{k}\cdot\mathbf{v}$ of wave 1 is therefore increased. The Doppler effect brings the apparent frequency of wave 1 and the atomic resonance frequency closer together; consequently, the modulus of the radiation pressure force F_1 exerted by wave 1 on the atom and directed, like wave 1, towards the negative x, increases with respect to its value for $v = 0$. The conclusions are just the opposite for wave 2, whose frequency is shifted away from resonance by the Doppler effect; the force F_2 it creates along the positive x is weaker than its value for $v = 0$.

The sum of the two forces[5] is zero for $v = 0$ (both forces have the same modulus and opposite directions), but no longer zero when $v \neq 0$. For positive values of v, it has the same direction as F_1 since the modulus of F_1 is larger than that of F_2 (Figure 3); for negative values of v, it is just the opposite since the two waves have switched their roles. The sum $F = F_1 + F_2$ of the two radiation pressure forces exerted by the two waves is thus in the direction opposite to that of the velocity v. For small values of v, it is proportional to v and can be written:

$$F = -\gamma v \tag{17}$$

where γ is a positive friction coefficient.

Under the effect of this friction force, the atomic velocities are constantly reduced. Their dispersion, however, does not tend towards zero for long interaction times, because of the unavoidable fluctuations in the emission and absorption processes. There is actually a competition between the friction effect we just described, which tends to cool down the atoms, and the momentum diffusion that tends to heat them up. We evaluate in the next paragraphs the effect of these two processes to estimate the order of magnitude of the temperatures that can be obtained by Doppler laser cooling.

[5]We shall see below that the effects of interference between the two waves can be neglected.

β. Estimation of the friction coefficient

We will now seek an estimation of the friction coefficient, in order to calculate the evolution of the momentum of the atom, as well as of its energy. In a first step and for the sake of simplicity, we will limit ourselves to a calculation of the average effect of the photon absorption and emission cycles by the atom. This will allow us to determine the evolution of its average momentum $\overline{p}$. Nevertheless, the absorption and emission processes of photons by an atom are actually fluctuating processes, as we will discuss below. Ssuch a calculation is theferore not sufficient to obtain the evolution of the average $\overline{p^2}$ of the square of the momentum, that is of the average kinetic energy, since the average value of a square differs in general from the square of the average value. In a second step, we will use a calculation that takes the fluctuations of the momentum transfers betwee photons and atoms into account.

We set:

$$\delta = \omega - \omega_0 + \omega_R \tag{18}$$

where $\hbar\omega_R$ is the recoil energy defined in (7); δ is the detuning between the laser frequency ω and the atomic frequency $\omega_0 - \omega_R$. We assume from now on that the intensity of the two lasers is weak; the atomic transition is not saturated and consequently the population σ_{bb} of the excited level b remains low. Its variation as a function of the detuning δ follows a Lorentzian curve with a total width at half maximum equal to Γ:

$$\sigma_{bb}(\delta) = \sigma_{bb}(0)\frac{(\Gamma/2)^2}{[(\Gamma/2)^2 + \delta^2]} \tag{19}$$

We shall not need the expression for $\sigma_{bb}(0)$, since it cancels out in the expressions for the friction and diffusion coefficients we shall obtain; it can however be found in Chapter V of reference [21] (optical Bloch equations).

In a perturbative approach to the problem, two types of terms must be considered: the "square" terms, coming from the interaction between the dipole induced by wave i with that same wave i; the "cross" terms coming from the interaction of the dipole induced by wave i with the other wave $j \neq i$. The cross terms correspond to interference between the effects of the two waves. However, as these two waves do not have the same spatial dependence (they propagate in opposite directions), these interference effects vary rapidly in space as $\exp(\pm 2ikx)$. They consequently vanish when the forces are averaged over distances of the order of the laser wavelengths, as we shall assume. It is then possible to consider that the radiation pressure force acting on the atom is simply the sum of the radiation pressure forces exerted by each wave, in the absence of the other.

When the atom is at rest, the two waves have the same frequency in the atom's reference frame, and hence the same detuning δ; remember that δ is supposed to be negative in a laser cooling experiment – see (16). If the atom moves with a velocity $v > 0$, we just saw that the apparent frequency of wave 1 is increased by a quantity kv (where k and v are positive), so that the detuning for the interaction of the atom with wave 1 is equal to:

$$\delta_1 = \delta + kv \tag{20}$$

while it is $\delta_2 = \delta - kv$ for wave 2. For this atom, the population of the excited state is the sum of two contributions: that of wave 1, obtained by replacing δ with $\delta + kv$ in

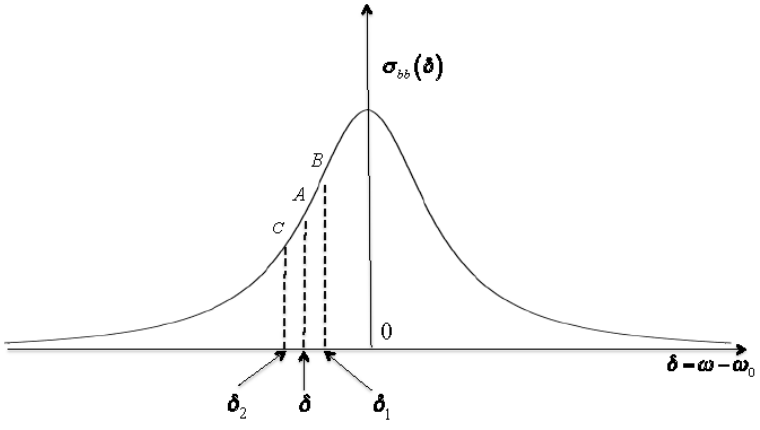


Figure 4: Populations $\sigma_{\rm bb}(\delta_1)$ and $\sigma_{\rm bb}(\delta_2)$ excited in the upper state b by wave 1, with detuning $\delta_1 = \delta + kv$, and wave 2, with detuning $\delta_2 = \delta - kv$. As the detuning δ for $v = 0$ is assumed to be negative, the ordinate $\sigma_{\rm bb}(\delta_1)$ of point B is higher than the ordinate $\sigma_{\rm bb}(\delta_2)$ of point C.

expression (19), and which is proportional to the ordinate of point B whose abscissa is $\delta_1 = \delta + kv$ in Figure 4; that of wave 2, which requires to replace δ with $\delta - kv$ in (19), and which is therefore proportional to the the ordinate of point C having the abscissa $\delta - kv$ in that same figure. Computations similar to those of § 1-d allows computing the total force F acting on the atom, as the sum of the two forces exerted by each wave, added independently of each other:

$$F = -\Gamma\hbar k\sigma_{bb}(\delta + kv) + \Gamma\hbar k\sigma_{bb}(\delta - kv) \tag{21}$$

where $\sigma_{bb}(\delta)$ is the function defined on the right-hand side of (19).

When kv is small compared to the width Γ of the curve in Figure 4, one can expand $\sigma_{bb}(\delta \pm kv)$ to first order in kv and obtain:

$$F = -2kv\ \Gamma\hbar k\ \frac{\mathrm{d}}{\mathrm{d}\delta}\sigma_{bb}(\delta) \tag{22}$$

The last factor on the right-hand side of (22), which is the slope at point A (with abscissa δ) of the curve representing $\sigma_{bb}(\delta)$, can be computed from (19). For the point $\delta = -\Gamma/2$ where the slope is maximum, we find:

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\sigma_{bb}(\delta) = \frac{2\sigma_{bb}(\delta)}{\Gamma} \tag{23}$$

Inserting (23) into (22), we get:

$$F = -\gamma v \tag{24}$$

where the friction coefficient γ is given by:

$$\gamma = 4\hbar k^2\sigma_{bb}(\delta) \tag{25}$$

Using equation (24), we can also compute the damping of the momentum $\overline{p}$ and of its square. As $F = \mathrm{d}\overline{p}/\mathrm{d}t$ and $v = \overline{p}/M$, we have:

$$\frac{\mathrm{d}\overline{p}}{\mathrm{d}t} = -\gamma\frac{\overline{p}}{M} \tag{26}$$

We can also write:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{p}^2 = 2\overline{p}\frac{\mathrm{d}\overline{p}}{\mathrm{d}t} = -2\gamma\frac{\overline{p}^2}{M} \tag{27}$$

Remember, however, that the average value of the kinetic energy is proportional to $\overline{p^2}$, the average value of the square of the momentum, and not to the square of its average value $\overline{p}^2$. We should therefore not conclude from relation (27) that the ultimate temperature that can be reached by Doppler laser cooling when $t \to \infty$ is zero. Equation (27) was obtained by considering only the average effect on the atom's velocity of the light beams and of the successive spontaneous emission processes, which introduces a continuous average evolution. But, in reality, the absorption and emission processes fluctuate and yield photons emitted in random directions. Even though their total momentum has a zero average (meaning $\overline{p}$ is not affected), these fluctuations will change $\overline{p^2}$. This effect can be considered as a source of noise (also called momentum diffusion) that increases $\overline{p^2}$, and acts in the direction opposite to the friction. It is therefore the competition between these two opposed mechanisms that leads to an equilibrium state, whose energy $\overline{p^2}/2M$ determines the ultimate temperature that can be obtained.

γ. *Momentum diffusion*

We now study the diffusion of the momentum of one atom, due to the spontaneously emitted photons; the ensemble of particles discussed above then reduces to one single atom ($P = p$). Let us consider a time interval $\mathrm{d}t$ whose value will be specified later. We call $\mathrm{d}N_1$ and $\mathrm{d}N_2$ the photon numbers from waves 1 and 2 that are absorbed during that time interval. As we assume the friction has acted long enough to cancel the average velocity, the detuning δ has become the same for the two beams. We then have, on the average:

$$\mathrm{d}\overline{N_1} = \mathrm{d}\overline{N_2} = \mathrm{d}N \tag{28}$$

Each absorbed photon then yields a spontaneously emitted photon. We use here a simple one-dimension model: each photon is emitted spontaneously in a random way, either in the positive x direction (the atom's recoil is then negative), or in the negative x direction (the recoil is then positive). The variation of the atom's momentum is then $\xi_i\hbar k$, with $\xi_i = -1$ in the first case and $\xi_i = +1$ in the second. There are no correlations in the directions of two consecutive photons. The total momentum $\mathrm{d}P$ gained by the atom during the absorption and re-emission of $\mathrm{d}N_1 + \mathrm{d}N_2 = 2\ \mathrm{d}N$ photons is equal to the sum of the momentum gained by the absorption of photons from beam 1, of the momentum gained by the absorption from beam 2, and finally of the momentum coming from the spontaneous emissions:

$$\mathrm{d}\overline{p} = \hbar k\left[\mathrm{d}N_2 - \mathrm{d}N_1 + \sum_{i=1}^{2\ \mathrm{d}N}\xi_i\right] \tag{29}$$

(i) To begin with, we neglect the fluctuations of the number of photons absorbed from each wave (we shall discuss this point later). The numbers $\mathrm{d}N_1$ and $\mathrm{d}N_2$ are then equal to their average values, and we have:

$$\mathrm{d}\overline{p} = \hbar k \sum_{i=1}^{2\ \mathrm{d}N} \xi_i \tag{30}$$

The summation over i in (29) yields a zero contribution to $\mathrm{d}\overline{p}$, since the sum of the ξ_i is zero on the average, but this is not the case for the contribution to $\mathrm{d}\overline{p^2}$ that we now compute. Taking the square of the summation over i in (30), the cross terms $\overline{\xi_i \xi_j}$ with $j \neq i$ are zero on the average, because there are no correlations between the signs of the ξ_i and the ξ_j. We are left with the square terms $\overline{\xi_i^2}$ which are all equal to 1. We obtain:

$$\mathrm{d}\overline{p^2} = 2\ \mathrm{d}N\ \hbar^2 k^2 \tag{31}$$

which is obviously non-zero.

Let us specify the time interval $\mathrm{d}t$, which should be long enough for the number dN of absorptions and spontaneous emissions to be large during this time, but sufficiently small for the variations of σ_{bb} to stay negligible. We can then write:

$$\mathrm{d}N = \dot{N}\mathrm{d}t = \Gamma\sigma_{bb}(\delta)\mathrm{d}t \tag{32}$$

where we used in the last equality the fact that the average number of photons absorbed per unit time in each wave is equal to $\Gamma\sigma_{bb}(\delta)$, as we have seen above. Inserting this result in (31), we get the increase of $\langle P^2 \rangle$ during the time interval $\mathrm{d}t$:

$$\mathrm{d}\overline{p^2} = 2\hbar^2 k^2 \mathrm{d}N = 2\hbar^2 k^2 \Gamma\sigma_{bb}(\delta)\mathrm{d}t \tag{33}$$

We finally get:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{sp}} = 2\hbar^2 k^2 \Gamma\sigma_{bb}(\delta) = D_{\mathrm{sp}} \tag{34}$$

The subscript "sp" of the parenthesis reminds us that this increase of $\langle P^2 \rangle$ per unit time is due to spontaneous emission processes. This expression is often called the "diffusion coefficient" D_{sp}.

(ii) We now study the effect of fluctuations on the numbers $\mathrm{d}N_1$ and $\mathrm{d}N_2$ of absorbed photons in each wave during the time interval $\mathrm{d}t$. If these fluctuations are no longer neglected, we must write:

$$\mathrm{d}N_i = \mathrm{d}\overline{N_i} + \delta n_i = \mathrm{d}N + \delta n_i \tag{35}$$

with $i = 1, 2$. In this equation, δn_i is the fluctuation of the number of absorbed photons in wave i (by definition, the average value of δn_i is zero). The total momentum $\mathrm{d}P$ the atom receives during the absorption of these photons is equal to:

$$\mathrm{d}p = \hbar k(\mathrm{d}N_2 - \mathrm{d}N_1) = \hbar k(\delta n_2 - \delta n_1) \tag{36}$$

Since the average values of δn_1 and δn_2 are zero, the fluctuations in the absorption process do not change the average value of $\overline{p}$, but this is not true for the average value

$\overline{p^2}$. Taking the square of (36) and using the fact that there are no correlations between the fluctuations of δn_1 and δn_2 (and hence the average value of their product is zero), we get:

$$\mathrm{d}\overline{p^2} = \left[\overline{\delta n_1^2} + \overline{\delta n_2^2}\right]\hbar^2 k^2 \tag{37}$$

To compute $\overline{\delta n_i^2}$, we take the square of equation (35). This leads to:

$$\mathrm{d}N_i^2 = \mathrm{d}N^2 + 2\ \mathrm{d}N\delta n_i + \delta n_i^2 \tag{38}$$

We now take the average value of each side of this equation. Using the fact that the average value of δn_i is zero, we get:

$$\overline{\delta n_i^2} = \mathrm{d}\overline{N_i^2} - \mathrm{d}N^2 \tag{39}$$

The quantity $\mathrm{d}\overline{N_i^2} - \mathrm{d}N^2$ is simply the variance of the number of photons absorbed from the wave. In general, for Poisson statistics, we have[6]:

$$\mathrm{d}\overline{N_i^2} - \mathrm{d}\overline{N_i}^2 = \mathrm{d}\overline{N_i} = \mathrm{d}N \tag{40}$$

We deduce that $\overline{\delta n_i^2}$ is simply equal to $\mathrm{d}N$, so that equation (37) is simply written:

$$\mathrm{d}\overline{p^2} = 2\,\mathrm{d}N\ \hbar^2 k^2 \tag{41}$$

We therefore obtain for the increase of $\overline{p^2}$ due to fluctuations in the absorption processes a result identical to (31). The computations leading from (31) to (34) can be repeated and we get the following result for the increase, per unit time, of $\overline{p^2}$ due to the absorption processes:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{abs}} = 2\hbar^2 k^2 \Gamma \sigma_{bb}(\delta) = D_{\mathrm{abs}} \tag{42}$$

where the diffusion coefficient D_{abs} due to the absorption processes has the same value as D_{sp} for the spontaneous emission processes:

$$D_{\mathrm{abs}} = D_{\mathrm{sp}} = 2\hbar^2 k^2 \Gamma \sigma_{bb}(\delta) \tag{43}$$

To evaluate the global rate of variation of $\overline{p^2}$, we must add to the rates of variation (34) and (42) the one due to the cooling process $(\mathrm{d}\overline{p^2}/\mathrm{d}t)_{\mathrm{cool}}$. This variation is simply the variation of $\overline{p}^2$ in the absence of fluctuations during the momentum exchanges, so that it can be obtained by assuming that $\overline{p^2}$ and $\overline{p}^2$ are simply equal. Using (27), we can write this variation as:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{cool}} = -2\gamma\frac{\overline{p^2}}{M} \tag{44}$$

[6] One could evaluate the effects of the deviations from Poisson statistics, but it will not be done here to keep things simple; this is legitimate for the low laser intensities (unsaturated transition) we assumed here.

Adding (34), (42) and (44), we finally obtain a total rate of change equal to:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{tot}} = -2\gamma\frac{\overline{p^2}}{M} + D_{\mathrm{sp}} + D_{\mathrm{abs}} \tag{45}$$

This rate of change goes to zero when :

$$\overline{p^2} = (D_{\mathrm{sp}} + D_{\mathrm{abs}})\frac{M}{2\gamma} \tag{46}$$

Dividing both sides of this equation by $2M$ and using expressions (25) and (43) for the friction and diffusion coefficients, we finally get[7] the expression for the average kinetic energy in the stationary regime:

$$\frac{\overline{p^2}}{2M} = \frac{\hbar\Gamma}{4} \tag{47}$$

This result indicates that the ultimate temperature is directly related to the natural width of the excited level.

δ. Doppler temperature

In a Doppler laser cooling experiment, and for a one-dimensional treatment of the problem, the average kinetic energy of the velocity fluctuations around its average value is related to the Doppler temperature T_D :

$$\frac{\overline{p^2}}{2M} = \frac{k_B T_D}{2} \tag{48}$$

where k_B is the Boltzmann constant. Using (47) and (48), we find that the equilibrium temperature that can be reached by Doppler laser cooling is given by:

$$T_D = \frac{\hbar\Gamma}{2k_B} \tag{49}$$

For sodium atoms, this temperature is equal to 235×10^{-6} K, that is 10^6 times lower than room temperature (of the order of 300 K)!

Our treatment of Doppler laser cooling is based on several approximations, as for example the one-dimensional treatment and the simplified description of spontaneous emission occurring only in two opposite directions. Nevertheless, more precise calculations lead to the conclusion that the average kinetic energy reached in a stationary regime is indeed of the order of $\hbar\Gamma$, within numerical coefficients of a few units.

Finally, equation (47) yields an estimation of the velocity $\bar{v}_D$ of the cold atoms thus obtained:

$$\bar{v}_D \simeq \sqrt{\hbar\Gamma/M} \tag{50}$$

This means that the Doppler shifts $\pm k\bar{v}_D$ of the apparent frequencies of waves 1 and 2 are such that the separations of the dotted lines in Figure 4 is of the order of $\pm k\sqrt{\hbar\Gamma/M}$.

[7] As the diffusion and friction coefficients are all proportional to $\sigma_{bb}(\delta)$, this factor disappears from (46) and is no longer present in (47).

We can compare these Doppler shifts with the width Γ of the curve plotted in Figure 4 and obtain:

$$\frac{k\sqrt{\hbar\Gamma/M}}{\Gamma} \simeq \sqrt{\frac{\hbar k^2/M}{\Gamma}} \simeq \sqrt{\frac{E_{\rm rec}}{\hbar\Gamma}} \tag{51}$$

where $E_{\rm rec}$ is the recoil energy given in (7). The ratio $E_{\rm rec}/\hbar\Gamma$ is in general small for the atomic resonance lines used in laser cooling; it is of the order of 1/100 for sodium, which means that $k\bar{v}_D$ remains small compared to Γ and shows the validity of the limited series expansion used in equation (22).

Other laser cooling methods

Until now we have described laser cooling methods using only the Doppler effect. Other methods have been proposed and demonstrated, such as "Sisyphus cooling" (§ 4 of Complement D_{XX} and [22]), the "subrecoil cooling" and "evaporative cooling" [23]. The reader interested in the two latter methods may read for instance § 13.3 of [24].

2-c. Magneto-optical trap

We wish to introduce an imbalance, depending on the atom's position x, between two laser beams 1 and 2, with the same frequency but propagating in opposite directions along the Ox axis. This requires achieving detunings (between the lasers' frequency and the atomic frequency) that depend on the position x of the atom along the Ox axis; these detunings must be the same for both lasers when $x = 0$, and different when $x \neq 0$. We must then necessarily use an atom with several Zeeman sublevels and different polarizations for the two beams 1 and 2. The principle of the method, suggested for the first time by Jean Dalibard in 1986, is schematized in Figure 5. We assume the atomic transition is between a ground state with a zero angular momentum ($J_g = 0$) and an excited state with an angular momentum equal to 1 ($J_e = 1$). The solid lines in Figure 5 plot the energy of the three Zeeman sublevels e_{+1}, e_0 and e_{-1} of the excited state, and of the sublevel g_0 of ground state, in an inhomegeneous magnetic field B applied along the Ox axis. This field is zero at $x = 0$ and varies linearly with x around $x = 0$; it can be created, for example, by two circular coils having the same axis Ox, placed symmetrically with respect to $x = 0$, and carrying currents of opposite directions. The quantization axis is chosen along Ox and allows defining the magnetic quantum numbers m_e and m_g of the sublevels in the excited and ground states. The two laser waves 1 and 2 propagate in opposite directions and have opposite circular polarizations with respect to the quantization axis [8]. Wave 1, with polarization σ_+, excites the transition $g_0 \leftrightarrow e_{+1}$ whereas wave 2, with polarization σ_-, excites the transition $g_0 \leftrightarrow e_{-1}$. The energy $\hbar\omega_0$ of the zero-field atomic transition is equal to the difference between the energies of states e_0 and g_0 (solid horizontal lines in the figure). The detuning $\hbar\delta = \hbar\omega - \hbar\omega_0$ between the energy $\hbar\omega$ of the laser photons and that of the zero-field atomic transition is shown as the difference in height between the dashed and solid horizontal lines in the figure. We

[8] One generally defines the right-hand and left-hand circular polarization with respect to the propagation direction of the photon. In that case, the polarization of the photons of wave 1 and wave 2 in Figure 5 would be σ_+ (in both cases the electric field of the wave turns around the direction of propagation following the right-hand rule). When studying the selection rules of the various transitions resulting from the conservation of spin angular momentum (see Complements B_{XIX} and C_{XIX}), it is best to define both the quantum numbers and the polarizations of all the beams with respect to the same axis, chosen here to be the quantization axis Ox.

assume here that the detuning is negative (the laser's frequency ω is shifted towards the red of the atomic transition ω_0).

At point $x = 0$, the energies of states e_{+1} and e_{-1} are equal, as are the detunings of waves 1 and 2 which excite the transitions $g_0 \leftrightarrow e_{+1}$ and $g_0 \leftrightarrow e_{-1}$. The radiation pressure forces exerted by the two waves are equal in intensity and opposite in direction, so that the resulting force is zero. This balance does not hold as soon as we move away from $x = 0$. For example, at $x = x_+$, wave 1 which excites transition $g_0 \rightarrow e_{+1}$, is at resonance with this transition and the force it exerts is at its maximum. On the other hand, at that same point, wave 2, which excites transition $g_0 \rightarrow e_{-1}$, is way off-resonance and hence exerts a much weaker force. Consequently, the balance between the two waves is broken in favor of wave 1 and the resulting force exerted by both waves is non-zero and directed towards the right. The conclusions are inverted at point $x = x_-$, where the total force is non-zero and directed towards the left. We obtain a restoring force, proportional to x in the vicinity of $x = 0$, which traps the atoms around $x = 0$. Such a trap is called a "magneto-optical trap" or "MOT".

For the sake of simplicity, we only considered a one-dimensional model, but the extension to three dimensions is possible. Note in particular that the field created by

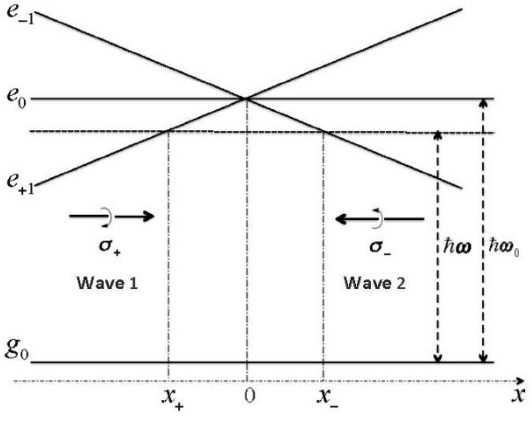


Figure 5: Principle of the magneto-optical trap. The transition used is a $J_g = 0 \rightarrow J_e = 1$ transition, excited by two laser waves 1 and 2 propagating in opposite directions along the Ox axis, with polarizations (σ_+) and (σ_-) with respect to the quantization axis Ox. A magnetic field gradient is applied along the Ox axis, the magnetic field being equal to zero at $x = 0$. Wave 1, which excites the transition $g_0 \leftrightarrow e_{+1}$, is resonant for this transition at point $x = x_+$ and the radiation pressure force it exerts on the atoms is then maximal. At that point, wave 2, which excites the transition $g_0 \leftrightarrow e_-$, is off-resonance for the corresponding atomic frequency, and, consequently, exerts a much weaker force. The two forces are unbalanced, in favor of the wave 1 force. The resulting force is non-zero, directed towards $x > 0$. The conclusions are reversed at point $x = x_-$, where the resulting force is non-zero, directed towards $x < 0$. Finally at $x = 0$ both waves are off-resonance by the same amount, and the resulting force is zero. We obtain a restoring force that traps the atoms around $x = 0$.

two circular coils centered around the Ox axis, placed on each side of point $x = 0$ and carrying currents in opposite directions, is zero at $x = 0$ and exhibits non-zero gradients along the Oy and Oz axis. The detuning towards the red of the laser frequency with respect to the atomic frequency also has the advantage of providing a Doppler laser cooling effect. The magneto-optical trap is nowadays a basic tool of cold atom physics[9].

Other trapping methods

Other methods for trapping atoms with light beams exist, for instance laser dipole trapping methods (Complement D_{XX}, §§ 1, 2 and 3).

3. Blocking recoil through spatial confinement

We now assume the atom or the ion under study is subjected to an external potential that traps it in a region of space. The energy spectrum of the external variables is no longer a continuous spectrum (as would be the case for a free atom), but a spectrum including a discrete part corresponding to the atomic bound states. Furthermore, because of this external potential, the atomic Hamiltonian is no longer translation invariant, and hence the total momentum is no longer a good quantum number. In this section we study how the absorption and emission spectra of the atom are modified by its confinement within the potential, and how the recoil of the atom can be blocked in certain cases.

3-a. Assumptions concerning the external trapping potential

We will assume the external potential acts only on the external variables, and not on the internal variables. This is the case for example for an atomic ion trapped by electric and magnetic fields, which only act on the center of mass via the global ionic charge, but does not act on the internal variables [10]. Figure 6 plots the trapping potentials for an atom or an ion in the internal states a or b. The two potential curves are identical, deduced from one another by a vertical translation of amplitude $\hbar\omega_0$ where ω_0 is the frequency of the internal atomic transition $b \leftrightarrow a$. The spectrum of the vibrational levels, of energies E_v, $E_{v'}$, $E_{v''}$, ... is the same for the two potentials. The atomic states are labeled by two quantum numbers: a quantum number a or b for the internal variables; a vibrational quantum number v, v', v'', ... for the external variables. The atomic transitions between the two internal states a and b now present a structure due to the vibrational motion of the center of mass. The frequency $\omega_{v',v}$ of the transition $b, v' \leftrightarrow a, v$ is given by:

$$\hbar\omega_{v',v} = \hbar\omega_0 + E_{v'} - E_v \tag{52}$$

The quantity $E_{v'} - E_v$ gives the variation of the external energy of the atom during the $b, v' \to a, v$ transition.

[9] See § 14-7 of [24] for a description of the first experimental realizations of such a trap and for a more quantitative study of its performances.

[10] The ion is generally confined in the center of the trap, a region where the electric and magnetic fields are very weak. It is then legitimate to neglect the Stark or Zeeman shifts of the internal states.

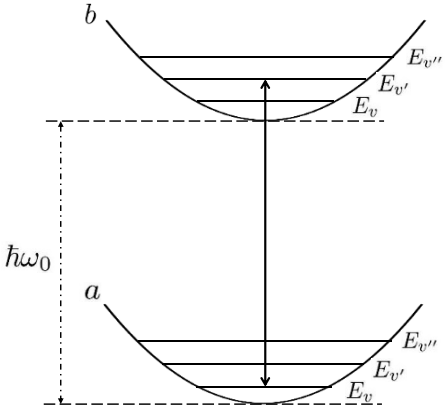


Figure 6: The external potential trapping the ion is the same when the ion is in either of its internal states a or b, separated by an energy difference $\hbar\omega_0$. Consequently, the spectrum of the vibrational levels of the center of mass in the external potential is the same for both internal states.

3-b. Intensities of the vibrational lines

We showed in § C-5 of Chapter XIX that the matrix elements of the interaction Hamiltonian could be factored into three terms pertaining to the three types of variables, the internal and external atomic variables, and the radiation variables – see relations (C-34) and (C-35) of that chapter. The part relative to the external variables is equal to $\langle\psi_{\text{fin}}^{\text{ext}}|\exp(i\,\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\psi_{\text{in}}^{\text{ext}}\rangle$ for an absorption process, where $|\psi_{\text{fin}}^{\text{ext}}\rangle$ and $|\psi_{\text{in}}^{\text{ext}}\rangle$ are the external final and initial states of the transition, equal here to $|\varphi_{v'}\rangle$ and $|\varphi_v\rangle$. This leads to an intensity of the vibrational line $a, v \leftrightarrow b, v'$ proportional to:

$$I_{v',v} = |\langle\varphi_{v'}|\exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\varphi_v\rangle|^2 \tag{53}$$

The $I_{v',v}$ obey the sum rule (obtained by the closure relation on the states $|\varphi_{v'}\rangle$):

$$\sum_{v'} I_{v',v} = \sum_{v'}\langle\varphi_v|\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\varphi_{v'}\rangle\langle\varphi_{v'}|\exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\varphi_v\rangle = 1 \tag{54}$$

It follows that the relative weight of the transition $a, v \to b, v'$ compared to all the transitions starting from a, v is precisely equal to $I_{v',v}$.

Another sum rule

The relative weights $I_{v',v}$ obey another sum rule:

$$\langle E\rangle = \sum_{v'} I_{v',v}(E_{v'} - E_v) = \frac{\hbar^2 k^2}{2M} = E_{\text{rec}} \tag{55}$$

which indicates that the average energy gained by an atom going from a given level a, v to another level b, v' is equal to the recoil energy, whatever the value of v. To prove relation (55), we rewrite the sum over v' in (55) in the form:

$$\sum_{v'} \langle\varphi_v| \left[\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}), \hat{H}_{\text{ext}}\right] |\varphi_{v'}\rangle \langle\varphi_{v'}| \exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) |\varphi_v\rangle \tag{56}$$

where $\hat{H}_{\text{ext}} = \hat{\boldsymbol{P}}^2/2M + V(\hat{\boldsymbol{R}})$ is the external variable Hamiltonian. The only term in H_{ext} that does not commute with $\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})$ is the kinetic energy term. We can replace the commutator appearing in (56) by $\left[\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}), \hat{\boldsymbol{P}}^2/2M\right]$. We now develop this commutator and use the closure relation on the $|\varphi_{v'}\rangle$ states, as well as relation:

$$\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) \left(\frac{\hat{\boldsymbol{P}}^2}{2M}\right) \exp(+i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) = \frac{1}{2M}(\hat{\boldsymbol{P}} + \hbar\boldsymbol{k})^2 \tag{57}$$

This yields:

$$\begin{aligned}\langle\varphi_v| \exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) \frac{\hat{\boldsymbol{P}}^2}{2M} \exp(+i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) - \frac{\hat{\boldsymbol{P}}^2}{2M} |\varphi_v\rangle \\ = E_{\text{rec}} + \langle\varphi_v|\hbar\boldsymbol{k}\cdot\hat{\boldsymbol{P}}/M|\varphi_v\rangle = E_{\text{rec}}\end{aligned} \tag{58}$$

since, taking Ehrenfest's theorem into account (Chapter III, § D-1-d-β), the average value of operator $\hat{\boldsymbol{P}}$ (equal to $M\ \mathrm{d}\langle\boldsymbol{R}\rangle/\mathrm{d}t$) in the stationary state $|\varphi_v\rangle$ is zero.

3-c. Effect of the confinement on the absorption and emission spectra

The absorption and emission spectra of an atom are significantly modified by the confinement of its center of mass.

As we saw in § 1, when a free atom has a well defined initial momentum $\boldsymbol{P}_{\text{in}}$, the absorption of a photon with momentum $\hbar\boldsymbol{k}$ places it in a well defined momentum state $\boldsymbol{P}_{\text{fin}} = \boldsymbol{P}_{\text{in}} + \hbar\boldsymbol{k}$. The conservation of the global momentum means that there is only one final state, $|b, \boldsymbol{P}_{\text{fin}}\rangle$, corresponding to an initial state $|a, \boldsymbol{P}_{\text{in}}\rangle$, and hence a single absorption line $a \leftrightarrow b$.

On the other hand, when the global momentum is no longer conserved because of the external potential's trapping of the atom, we get several lines going from an initial given state a, v to several possible final states b, v', whose frequencies $\omega_{v',v}$ are given by (52). One can then ask which of these lines is the strongest.

To answer that question, we go back to expression (53) for the relative weight of the transition $a, v \to b, v'$. In this equation, operator $\exp(i\,\boldsymbol{k}\cdot\hat{\boldsymbol{R}})$ represents a translation operator, in momentum space, of the quantity $\hbar\boldsymbol{k}$. The quantity $I_{v',v}$ is thus proportional to the squared modulus of the scalar product of the vibrational wave function $\varphi_{v'}(\boldsymbol{r})$ and the wave function $\varphi_v(\boldsymbol{r})$ translated in momentum space by the quantity $\hbar\boldsymbol{k}$. Let us assume the atom is trapped in a region of spatial extension Δx, very small compared to the wavelength $\lambda = 2\pi/k$ of the incident photon. The momentum spread Δp of the wave function is then larger than $\hbar k$ since:

$$\Delta x \ll \lambda \tag{59}$$

leads to:

$$\Delta p \simeq \frac{\hbar}{\Delta x} \gg \frac{\hbar}{\lambda} = \frac{\hbar k}{2\pi} \tag{60}$$

This means that, as long as the excited levels' energy is not too large, the translation in momentum space of the wave function $\varphi_v(\boldsymbol{r})$ by a quantity $\hbar\boldsymbol{k}$ much smaller than its width leaves that wave function practically unchanged, and it consequently remains orthogonal to $\varphi_{v'}(\boldsymbol{r})$ for $v' \neq v$. The strongest line in the absorption spectrum is therefore the line with no change in the vibrational quantum number (a line sometimes called the *zero-phonon line*), whose frequency remains unchanged, equal to ω_0. A strong confinement suppresses the Doppler shift and the atom's recoil.

Comment on momentum conservation

One may wonder what happened to the momentum of the absorbed photon in the zero-phonon transition. Remember that we have treated the trapping potential of the atom as an external potential, which breaks the translation invariance of the problem under study: the Hamiltonian no longer commutes with the total momentum, which is no longer conserved. It is thus not surprising that we cannot follow what becomes of the photon momentum. We can, however, describe the trapping potential, not as an externally given potential, but rather coming from the interaction of the atom with another physical object whose dynamics must be taken into account. A quantum treatment of that device and its interaction with the atom permits introducing for the global system, "atom + trapping device", a Hamiltonian that commutes with the total momentum; it is the global system that absorbs the photon momentum. As this momentum is microscopic, whereas the mass of the device is macroscopic, the recoil velocity is so weak that the corresponding frequency shifts are totally undetectable.

3-d. Case of a one-dimensional harmonic potential

We now assume the external trapping potential is harmonic, and we call $\omega_{\rm osc}$ the oscillation frequency of the atoms in this potential. The energies E_v of the vibrational levels are equal to $(v + 1/2)\hbar\omega_{\rm osc}$, where v is an integer, positive or zero. The spatial extension Δx_0 of the ground state wave function $v = 0$ is equal to $\sqrt{\hbar/2M\omega_{\rm osc}}$. To characterize the confinement, we introduce the dimensionless parameter[11]:

$$\eta = k\Delta x_0 = 2\pi\frac{\Delta x_0}{\lambda} \tag{61}$$

If $\eta \ll 1$, the atoms are confined in a region small compared to the radiation wavelength. The square of the parameter η has a simple physical significance since:

$$\eta^2 = \frac{\hbar k^2}{2M\omega_{\rm osc}} = \frac{E_{\rm rec}}{\hbar\omega_{\rm osc}} \tag{62}$$

is the ratio between the recoil energy $E_{\rm rec}$ and $\hbar\omega_{\rm osc}$, which is the energy difference between vibrational levels in the potential well.

It is instructive to compute, as a function of η, the intensities $I_{0,0}$ and $I_{0,1}$ of the vibrational lines $a, 0 \to b, 0$ and $a, 0 \to b, 1$. Assume the photon wave vector $\boldsymbol{k}$ is parallel to the x axis. Exponential $\exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})$ appearing in equation (53) can be replaced by

[11] This parameter is often called the *Lamb Dicke parameter*, after the names of the physicists who first introduced the idea of recoil-free absorption in a trapped system. To get a historical overview of the various studies on the suppression of recoil due to confinement, the interested reader may consult § 6-4-4 of [24] as well as the references cited in that §.

$\exp(ik\hat{X})$. We now use the expression of operator $\hat{X}$ in terms of the annihilation and creation operators of the harmonic oscillator associated with the external potential:

$$\hat{X} = \sqrt{\hbar/2M\omega_{\text{osc}}}\,(\hat{a} + \hat{a}^\dagger) = \Delta x_0\,(\hat{a} + \hat{a}^\dagger) \tag{63}$$

We then get, in the limit $\eta \ll 1$, using (61):

$$\begin{aligned}\exp(ik\hat{X}) &= \exp\left[i\eta(\hat{a} + \hat{a}^\dagger)\right] \\ &= 1 + i\eta(\hat{a} + \hat{a}^\dagger) - \frac{\eta^2}{2}(\hat{a} + \hat{a}^\dagger)^2 + ...\end{aligned} \tag{64}$$

The series expansion (64) used in (53) yields, to order 2 in η:

$$I_{0,0} = 1 - \eta^2 \tag{65a}$$

$$I_{1,0} = \eta^2 \tag{65b}$$

All the other transitions $a, 0 \to b, v$ with $v \geqslant 2$ have relative intensities of a higher order, in η^{2v}. The transition with no change in the vibrational state, and hence with no recoil, is predominant for a strong confinement. The transition $a, 0 \to b, 1$ has a much lower probability, by a factor $E_{\text{rec}}/\hbar\omega_{\text{osc}}$; when it occurs, it increases the atomic energy by a quantity $\hbar\omega_{\text{osc}}$ much larger than E_{rec}. The sum rule (55) shows that, on the average, the energy gained by the atom remains equal to E_{rec}.

3-e. Mössbauer effect

In 1958, Rudolf Mössbauer observed very narrow lines in the resonant absorption spectrum of γ rays by the atomic nuclei in a crystal. Building on the previous work of Lamb [25] on the suppression of the recoil in the resonant absorption of slow neutrons (and not of photons) by the atomic nuclei in a crystalline network, he attributed the narrow spectral structures he observed to a suppression of the recoil. This suppression can occur if, in the crystal phonon spectrum, there are frequencies larger than the recoil frequency $\omega_R = E_{\text{rec}}/\hbar$. The interest of the Mössbauer effect comes from the high value of the frequency of the internal γ transition, which can reach 10^{18} Hz, or even much higher frequencies. If the Doppler width and the recoil shift are suppressed by the confinement, and if the natural width remains of the order of 10^6 to 10^7 Hz (as for an optical transition), the quality factor of the γ transition (ratio between frequency and the spectral width of the resonance) can reach values of the order of 10^{12}. Such a resolution allowed measuring, already in 1960 [26], and for the first time in a laboratory, the gravitational shift predicted by general relativity between the frequencies of an emitter and a receptor, both located in the earth's gravitational field but separated by an altitude of roughly twenty meters.

4. Recoil suppression in certain multi-photon processes

Until now, we only considered one-photon processes. We shall see in Chapter XX that there are multi-photon processes in which the atom goes from an internal state a to another state b by absorbing or emitting several photons. During such processes, the total energy and momentum must of course be conserved[12].

[12] We now consider again free atoms, without an external potential.

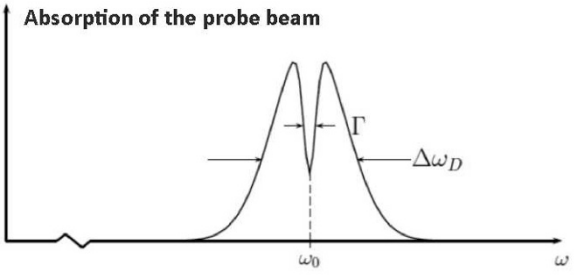


Figure 7: Saturated absorption spectroscopy. This figure plots the absorption profile of the probe beam when scanning the frequency ω of the two laser beams. A narrow hole, of width Γ, appears in the middle of a much larger Doppler profile, of width $\Delta\omega_D$.

Imagine a two-photon process, where the two photons have the same frequency and opposite wave vectors $+\boldsymbol{k}$ and $-\boldsymbol{k}$. The total radiation momentum is zero in this case. If the atom absorbs those two photons, its momentum does not change. Its external energy is not modified, meaning there is neither a Doppler effect, nor a recoil energy. This possibility can be extended to p-photon processes as long as the sum of the wave vectors of the p photons is zero:

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 ... + \boldsymbol{k}_p = \boldsymbol{0} \tag{66}$$

This idea was proposed independently by two groups, in Russia [27] and in France [28]. It led to significant experimental advances in high resolution spectroscopy where the line width is no longer limited by the Doppler width, but rather by the often much smaller natural width. A particularly interesting example is the study of the transition $1s \leftrightarrow 2s$ of the hydrogen atom by Doppler-free two-photon spectroscopy [29]. The upper state $2s$ of this transition is metastable, with a long lifetime (around 120 ms). Consequently, its natural width is very small and the two-photon line very narrow, which allows extremely precise measurements of fundamental constants such as the Rydberg constant. Note must be taken however that the interaction with the laser radiation inducing the two-photon transition leads to shifts of the energy levels[13], proportional to the light intensity; these must be taken precisely into account to determine the non-perturbed frequency of the two-photon transition.

Doppler-free saturated absorption spectroscopy

Nonlinear effects also appear in experimental set-ups where the atom interacts with two counter-propagating light beams, with the same frequency ω, one having a high intensity (pump beam) and the other a weaker one (probe beam). Contrary to the two-photon transitions considered in the previous paragraph, we assume here that the transitions induced by each beam are one-photon transitions between the two atomic internal states a and b; the laser frequency ω is therefore close to the atomic transition ω_0 (and not close to $\omega_0/2$).

We neglect here the recoil energy, in general very small in the optical domain compared with the natural width Γ of the upper state b. However, the Doppler shifts of the absorp-

[13] See Complement B_{XX} , §2-b

tion lines of the various atoms play an important role, as they are different for the pump beam and the probe beam which propagate in opposite directions. The pump beam interacts with an atom of velocity v_{pump} if its apparent frequency $\omega - kv_{\text{pump}}$ for that atom coincides with the atomic frequency ω_0 (within Γ), i.e. if $kv_{\text{pump}} = \omega - \omega_0$ within Γ. In the same way, the probe beam interacts with atoms of velocity v_{probe} if $\omega + kv_{\text{probe}} = \omega_0$ within Γ, i.e. if $kv_{\text{probe}} = -(\omega - \omega_0)$ within Γ. When ω is different from ω_0, we have $v_{\text{pump}} \neq v_{\text{probe}}$: the two beams do not interact with the same atoms in the velocity distribution, so that the absorption of the probe beam is not perturbed by the presence of the pump beam. However, this perturbation becomes important when $\omega = \omega_0$ (within Γ), since the two beams interact with the same sub-set of atoms (those belonging to the same "velocity group" along the beams' axis).

The high intensity pump beam lowers the difference in populations between the a and b levels of the atomic transition, and tends to equalize these populations. The absorption of the probe beam is thus diminished when the two beams interact with the same velocity group, i.e in the vicinity of $\omega = \omega_0$. When scanning the frequency ω of the two laser beams, the absorption of the probe beam varies according to a Doppler profile centered around ω_0, with width $\Delta\omega_D$, in the middle of which (Figure 7) appears a hole with a much smaller width Γ. This method, called *saturated absorption*, allows the determination of the atomic frequency ω_0 with a much better resolution than when using a single beam.

Conclusion

In this complement, we showed how the analysis of the momentum exchanges between atoms and photons allows introducing and interpreting several important physical phenomena. These phenomena include Doppler width, recoil energy, radiation pressure forces, Doppler laser slowing down and cooling of atoms, suppression of the Doppler effect due to confinement or in two-photon transitions, and the Mössbauer effect.

Thanks to these various methods, spectacular improvements in the resolution of spectroscopic measurements have been obtained. This led to high precision measurements and improvements of atomic clocks, which now have a relative stability of the order of 10^{-16}. Placing such a clock in the international spatial station and comparing its frequency with that of a similar clock on the Earth, one hopes to be able to test the value of the gravitational shift predicted by general relativity with a precision better, by a factor close to 100, than all the other existing tests. Another conclusion we can draw is that atom-photon interactions are useful tools for controlling and manipulating atoms.

We shall see in Complement C_{XIX} how the exchanges of angular momentum between atoms and photons allows controlling the angular momentum of the atoms, polarizing them via optical pumping. Such achievements have opened new fields of research, such as atomic interferometry and the study of degenerate quantum gases.

Complement B_{XIX}

Angular momentum of radiation

1 Quantum average value of angular momentum for a spin 1 particle 2044
1-a Wave function, spin operator 2044
1-b Average value of the spin angular momentum 2045
1-c Average value of the orbital angular momentum 2046
2 Angular momentum of free classical radiation as a function of normal variables 2047
2-a Calculation in position space 2047
2-b Reciprocal space 2048
2-c Difference between the angular momenta of massive particles and of radiation 2050
3 Discussion 2050
3-a Spin angular momentum of radiation 2050
3-b Experimental evidence of the radiation spin angular momentum 2051
3-c Orbital angular momentum of radiation 2052

Introduction

Radiation angular momentum plays an important role in many situations, in particular in atomic physics experiments. As will be explained in Complement C_{XIX}, the exchange of angular momentum between atoms and photons is the base of many experimental methods, such as optical pumping, which illustrated for the first time the manipulation of atoms by light.

In Chapter XVIII, a spatial Fourier transform of the classical fields led to the introduction of the field normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$, which are the field components in a basis of transverse plane waves. Upon quantization, these normal variables became the annihilation $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and creation $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ operators of a photon in a mode $\boldsymbol{k}, \boldsymbol{\varepsilon}$. Such a plane wave basis is particularly useful for studying the radiation energy and momentum, since the photons of mode $\boldsymbol{k}, \boldsymbol{\varepsilon}$ have a well defined energy $\hbar\omega = \hbar ck$ and momentum $\hbar\boldsymbol{k}$. On the other hand, the expansion of the field angular momentum in terms of the normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ is not as simple, since the photons of mode $\boldsymbol{k}, \boldsymbol{\varepsilon}$ do not have a well determined angular momentum. The aim of this complement is to find another expansion better adapted to the study of the radiation angular momentum, and establish a number of useful results.

In the classical description of radiation, the normal variable $\boldsymbol{\alpha}(\boldsymbol{k}) = \sum_{\boldsymbol{\varepsilon}} \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\boldsymbol{\varepsilon}$ is a vector function of $\boldsymbol{k}$ presenting a certain analogy with a wave function in the reciprocal space, and which could be seen as the wave function of the radiation (in momentum space). A physical quantity, such as the radiation total energy or the total momentum, does appear as the average value in that wave function of a one-particle operator

representing the energy or the momentum of a photon. We will see other examples of this analogy in this complement. As it is a vector function, this wave function can be regarded as the wave function of a particle of spin 1 whose total angular momentum $\boldsymbol{J}$ would be the sum of the orbital angular momentum $\boldsymbol{L}$ and the spin angular momentum $\boldsymbol{S}$. We will present, in § 1, a quantum mechanical calculation of the average values, in the state of a spin 1 particle described by the vector wave function $\boldsymbol{\Psi}(\boldsymbol{k})$, of the orbital and spin angular momentum of that particle. Returning to classical physics, in § 2 we will establish the expression for the total angular momentum $\boldsymbol{J}_R$ of free radiation; we shall first write it as a function of the fields in reciprocal space, then as a function of the field normal variables $\boldsymbol{\alpha}(\boldsymbol{k})$. The expression thus obtained has the same form as that obtained in § 1, provided we replace the $\boldsymbol{\Psi}(\boldsymbol{k})$ by the $\boldsymbol{\alpha}(\boldsymbol{k})$. This will lead us to the expansion in terms of normal variables of not only the field total angular momentum, but also of its orbital and spin angular momenta. The physical interpretation of these results is discussed in § 3, which highlights, in particular, some important characteristics of these two types of angular momenta.

1. Quantum average value of angular momentum for a spin 1 particle

We first study a spin 1 particle that has a mass, and is therefore not a photon. Our results will be useful as a point of comparison for the next paragraph's computations, where we return to the electromagnetic field.

1-a. Wave function, spin operator

The conclusions of this § 1 will be compared with those of the following § 2 concerning the radiation angular momentum expressed in terms of normal variables. As normal variables characterize the field in reciprocal space, it is important here to describe the state of the spin 1 particle in that same space. The particle state vector can be expanded on a basis $\{|\boldsymbol{k}, \nu\rangle\}$, where $\boldsymbol{k}$ represents the wave vector and ν the spin state:

$$|\boldsymbol{\Psi}\rangle = \sum_{\nu} \int \mathrm{d}^3k \; \psi_\nu(\boldsymbol{k}) \, |\boldsymbol{k}, \nu\rangle \tag{1}$$

with:

$$\psi_\nu(\boldsymbol{k}) = \langle \boldsymbol{k}, \nu | \boldsymbol{\Psi} \rangle \tag{2}$$

In general, we choose for the states $|\nu\rangle$ the eigenstates $|+1\rangle, |0\rangle, |-1\rangle$ of the S_z spin component. Here, we shall choose another basis:

$$\begin{aligned} |x\rangle &= (1/\sqrt{2})\Big[|-1\rangle - |+1\rangle\Big] \\ |y\rangle &= (i/\sqrt{2})\Big[|-1\rangle + |+1\rangle\Big] \\ |z\rangle &= |0\rangle \end{aligned} \tag{3}$$

The action of the $\boldsymbol{S}$ components S_x, S_y, S_z on these basis vectors can easily be computed. We must use $S_x = (S_+ + S_-)/2$ and $S_y = -i\,(S_+ - S_-)/2$, as well as the action of

S_+, S_-, S_z on the states $|+1\rangle, |0\rangle, |-1\rangle$ (Chapter VI, relations (C-50)). As an example, we obtain:

$$S_x|x\rangle = \frac{1}{2}(S_+ + S_-)\frac{1}{\sqrt{2}}\Big[|-1\rangle - |+1\rangle\Big] = \frac{\hbar}{2\sqrt{2}}\Big[\sqrt{2}|0\rangle - \sqrt{2}|0\rangle\Big] = 0$$
$$S_x|y\rangle = \frac{i}{2}(S_+ + S_-)\frac{1}{\sqrt{2}}\Big[|-1\rangle + |+1\rangle\Big] = \frac{i\hbar}{2\sqrt{2}}\Big[\sqrt{2}\,|0\rangle + \sqrt{2}\,|0\Big] = i\hbar|z\rangle$$
$$S_x|z\rangle = \frac{1}{2}(S_+ + S_-)|0\rangle = \frac{\hbar}{2}\Big[\sqrt{2}\,|+1\rangle + \sqrt{2}\,|-1\rangle\Big] = -i\hbar|y\rangle \tag{4}$$

These equations, and those similar for the action of S_y and S_z, can be written in a more compact way as:

$$S_a|b\rangle = i\hbar \sum_c \epsilon_{abc}\,|c\rangle \tag{5}$$

where a, b, c are the indices x, y, z and ϵ_{abc} is the three-dimensional completely antisymmetric tensor[1]. Equation (5) also leads to:

$$\langle c|S_a|b\rangle = i\hbar\,\epsilon_{abc} \tag{6}$$

1-b. Average value of the spin angular momentum

Taking (1) into account, the average value of S_a is written:

$$\langle\boldsymbol{\Psi}|S_a|\boldsymbol{\Psi}\rangle = \int \mathrm{d}^3k \int \mathrm{d}^3k' \sum_{b,c} \psi_c^*(\boldsymbol{k}')\langle\boldsymbol{k}', c|S_a|\boldsymbol{k}, b\rangle\psi_b(\boldsymbol{k}) \tag{7}$$

As S_a does not act on the orbital degrees of freedom, described by the variable $\boldsymbol{k}$, we get, taking (6) into account :

$$\langle\boldsymbol{k}', c|S_a|\boldsymbol{k}, b\rangle = i\hbar\,\epsilon_{abc}\,\delta(\boldsymbol{k} - \boldsymbol{k}') \tag{8}$$

Inserting this result in (7) then yields:

$$\langle\boldsymbol{\Psi}|S_a|\boldsymbol{\Psi}\rangle = i\hbar\int \mathrm{d}^3k \sum_{b,c} \epsilon_{abc}\,\psi_c^*(\boldsymbol{k})\psi_b(\boldsymbol{k}) = -i\hbar\int \mathrm{d}^3k \sum_{b,c} \epsilon_{abc}\,\psi_b^*(\boldsymbol{k})\psi_c(\boldsymbol{k}) \tag{9}$$

(using the fact that ϵ_{abc} is antisymmetric). As the a component of the cross product of two vectors $\boldsymbol{V}$ and $\boldsymbol{W}$ is written:

$$(\boldsymbol{V}\times\boldsymbol{W})_a = \sum_{b,c} \epsilon_{abc}\,V_b\,W_c \tag{10}$$

we get:

$$\langle\boldsymbol{\Psi}|S_a|\boldsymbol{\Psi}\rangle = -i\hbar\int \mathrm{d}^3k\,(\boldsymbol{\Psi}^*\times\boldsymbol{\Psi})_a \tag{11}$$

[1] By definition $\epsilon_{abc} = +1$ if abc are xyz or can be deduced from xyz by an even permutation; $\epsilon_{abc} = -1$ if abc is deduced from xyz by an odd permutation; finally $\epsilon_{abc} = 0$ if two of the three indices abc (or all three) are equal.

1-c. Average value of the orbital angular momentum

The orbital angular momentum is written:

$$\boldsymbol{L} = \boldsymbol{R} \times \boldsymbol{P} \tag{12}$$

Its L_a component acts in position space as:

$$L_a = (\boldsymbol{R} \times \boldsymbol{P})_a = \sum_{b,c} \epsilon_{abc}\, R_b\, P_c \Rightarrow -i\hbar \sum_{b,c} \epsilon_{abc} r_b\, \partial_c \qquad \text{with} \qquad \partial_c = \partial/\partial r_c \tag{13}$$

Going to the reciprocal space amounts to performing a spatial Fourier transform. The operators which in $\boldsymbol{r}$ space correspond to a multiplication by r_b and a derivation with respect to r_c, respectively become, in $\boldsymbol{k}$ space, the operators derivation with respect to k_b and multiplication by k_c (multiplied by appropriate factors):

$$r_b \underset{\text{FT}}{\leftrightarrow} i\partial/\partial k_b = i\tilde{\partial}_b \qquad\qquad \partial/\partial r_c = \partial_c \underset{\text{FT}}{\leftrightarrow} ik_c \tag{14}$$

with the notation:

$$\tilde{\partial}_b \equiv \partial/\partial k_b \tag{15}$$

In reciprocal space, the action of the L_a component of the orbital angular momentum is therefore:

$$\begin{aligned} L_a &\Rightarrow -i\hbar \sum_{b,c} \epsilon_{abc}\, (i\tilde{\partial}_b)\, (ik_c) = i\hbar \sum_{b,c} \epsilon_{abc}\, k_c\, \tilde{\partial}_b \\ &= -i\hbar(\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})_a \end{aligned} \tag{16}$$

In the last equality of the first line, we have moved $\tilde{\partial}_b$ to the right of k_c, which is allowed since the presence of ϵ_{abc} means that all the terms with equal indices b and c must be zero. To obtain the equality in the second line of (16), we use the anstisymetry of ϵ_{abc} under the exchange of b and c, and relation (10) of the vector product; $\boldsymbol{\nabla}_{\boldsymbol{k}}$ is the gradient with respect of $\boldsymbol{k}$.

We finally compute the average value of L_a in the state $\boldsymbol{\Psi}$. Taking (1) into account, we get:

$$\langle \boldsymbol{\Psi} | L_a | \boldsymbol{\Psi} \rangle = \int \mathrm{d}^3k \int \mathrm{d}^3k' \sum_{d,e} \psi_d^*(\boldsymbol{k}') \langle \boldsymbol{k}', d | L_a | \boldsymbol{k}, e \rangle \psi_e(\boldsymbol{k}) \tag{17}$$

As L_a does not act on the spin degrees of freedom, we must have $d = e$ in the matrix element on the right-hand side, which yields, using the differential form (16) for L_a:

$$\begin{aligned} \langle \boldsymbol{\Psi} | L_a | \boldsymbol{\Psi} \rangle &= \int \mathrm{d}^3k \int \mathrm{d}^3k' \sum_{d} \psi_d^*(\boldsymbol{k}') \langle \boldsymbol{k}', d | L_a | \boldsymbol{k}, d \rangle \psi_d(\boldsymbol{k}) \\ &= -i\hbar \int \mathrm{d}^3k \sum_{b,c,d} \psi_d^*(\boldsymbol{k})\; \epsilon_{abc}\; k_b\; \tilde{\partial}_c\; \psi_d(\boldsymbol{k}) \\ &= -i\hbar \int \mathrm{d}^3k \sum_{d} \psi_d^*(\boldsymbol{k})\, (\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})_a\, \psi_d(\boldsymbol{k}) \end{aligned} \tag{18}$$

Finally, adding (11) and (18), we obtain the following expression for the average value of the particle total angular momentum in the state $\boldsymbol{\Psi}$:

$$\langle \boldsymbol{\Psi} | \boldsymbol{J} | \boldsymbol{\Psi} \rangle = -i\hbar \int \mathrm{d}^3 k \left\{ \underbrace{\boldsymbol{\Psi}^*(\boldsymbol{k}) \times \boldsymbol{\Psi}(\boldsymbol{k})}_{\text{spin}} + \sum_d \underbrace{\psi_d^*(\boldsymbol{k})\, (\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})\, \psi_d(\boldsymbol{k})}_{\text{orbital}} \right\} \tag{19}$$

2. Angular momentum of free classical radiation as a function of normal variables

We now show that the classical calculation of the radiation angular momentum presents a certain analogy with the results of § 1.

2-a. Calculation in position space

Relation (A-53) of Chapter XVIII yields for the total angular momentum of free radiation (in the absence of particles):

$$\boldsymbol{J} = \varepsilon_0 \int \mathrm{d}^3 r \Big[\boldsymbol{r} \times [\boldsymbol{E}_\perp(\boldsymbol{r}) \times \boldsymbol{B}(\boldsymbol{r})] \Big] \tag{20}$$

Let us replace $\boldsymbol{B}$ by $\boldsymbol{\nabla} \times \boldsymbol{A}$ and use the triple product expansion $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\, \boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\, \boldsymbol{c}$. We obtain, keeping the right order between $\boldsymbol{\nabla}$ and $\boldsymbol{A}$:

$$\boldsymbol{E} \times [\boldsymbol{\nabla} \times \boldsymbol{A}] = \sum_d E_d \boldsymbol{\nabla} A_d - (\boldsymbol{E} \cdot \boldsymbol{\nabla}) \boldsymbol{A} \tag{21}$$

where E_d is the component of $\boldsymbol{E}_\perp(\boldsymbol{r})$ on the Ox, Oy or Oz axis, labeled by the index d. Inserting (21) in (20) leads to:

$$\boldsymbol{J} = \epsilon_0 \int \mathrm{d}^3 r \left\{ \sum_d E_d (\boldsymbol{r} \times \boldsymbol{\nabla}) A_d - \boldsymbol{r} \times (\boldsymbol{E} \cdot \boldsymbol{\nabla}) \boldsymbol{A} \right\} \tag{22}$$

Consider first the contribution $\boldsymbol{J}^{(1)}$ of the second term in (22). Its $J_a^{(1)}$ component is written:

$$\begin{aligned} J_a^{(1)} &= -\epsilon_0 \int \mathrm{d}^3 r \left\{ \boldsymbol{r} \times (\boldsymbol{E} \cdot \boldsymbol{\nabla}) \boldsymbol{A} \right\}_a \\ &= -\epsilon_0 \int \mathrm{d}^3 r \sum_{b,c,d} \epsilon_{abc}\, r_b\, E_d\, \partial_d\, A_c \end{aligned} \tag{23}$$

We now move r_b to the right of ∂_d, using:

$$r_b\, \partial_d = \partial_d\, r_b - \delta_{bd} \tag{24}$$

The contribution of the term $-\delta_{bd}$ to (23) yields:

$$\epsilon_0 \int \mathrm{d}^3 r \sum_{b,c} \epsilon_{abc}\, E_b\, A_c = \epsilon_0 \int \mathrm{d}^3 r\, (\boldsymbol{E} \times \boldsymbol{A})_a \tag{25}$$

As for the contribution of the term $\partial_d\, r_b$ to (23), we perform an integration by parts. The contribution of the integrated term yields a zero surface integral if the fields decrease fast enough at infinity. We obtain for the contribution of the term $\partial_d\, r_b$ to (23) :

$$-\epsilon_0 \int \mathrm{d}^3 r \sum_{b,c,d} \epsilon_{abc}\, E_d\, \partial_d\, (r_b\, A_c) = 0 + \epsilon_0 \int \mathrm{d}^3 r \sum_{b,c,d} \epsilon_{abc}\, r_b\, A_c\, \partial_d\, E_d \tag{26}$$

In the last term of (26), we note the quantity

$$\sum_d \partial_d\, E_d = \boldsymbol{\nabla} \cdot \boldsymbol{E} \tag{27}$$

which is equal to zero since the electric field, in the absence of sources, is purely transverse (and hence of zero divergence). This term therefore disappears.

Finally, the average value of $\boldsymbol{J}$ is the sum of (25) and of the first term of (22):

$$\boldsymbol{J} = \epsilon_0 \int \mathrm{d}^3 r \left\{ \underbrace{\boldsymbol{E}(\boldsymbol{r}) \times \boldsymbol{A}(\boldsymbol{r})}_{\text{spin}} + \sum_d \underbrace{E_d(\boldsymbol{r})(\boldsymbol{r} \times \boldsymbol{\nabla}) A_d(\boldsymbol{r})}_{\text{orbital}} \right\} \tag{28}$$

2-b. Reciprocal space

Expression (28) for $\boldsymbol{J}$ can be rewritten as a function of the Fourier transforms of the fields $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{A}(\boldsymbol{r})$. Using the Parseval-Plancherel equality (Appendix I, § 2-c) and relations (14), we get:

$$\boldsymbol{J} = \epsilon_0 \int \mathrm{d}^3 k \left\{ \underbrace{\tilde{\boldsymbol{E}}^*(\boldsymbol{k}) \times \tilde{\boldsymbol{A}}(\boldsymbol{k})}_{\text{spin}} + \sum_d \underbrace{\tilde{E}_d^*(\boldsymbol{k})(\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}}) \tilde{A}_d(\boldsymbol{k})}_{\text{orbital}} \right\} \tag{29}$$

We now use expressions (B-22a) and (B-22b) of Chapter XVIII to obtain the fields $\tilde{\boldsymbol{E}}(\boldsymbol{k})$ and $\tilde{\boldsymbol{A}}(\boldsymbol{k})$ as a function of the normal variables:

$$\tilde{\boldsymbol{E}}(\boldsymbol{k}) = \frac{i\omega}{2\mathcal{N}(k)} \left[\boldsymbol{\alpha}(\boldsymbol{k}) - \boldsymbol{\alpha}^*(-\boldsymbol{k})\right] \tag{30a}$$

$$\tilde{\boldsymbol{A}}(\boldsymbol{k}) = \frac{1}{2\mathcal{N}(k)} \left[\boldsymbol{\alpha}(\boldsymbol{k}) + \boldsymbol{\alpha}^*(-\boldsymbol{k})\right] \tag{30b}$$

where $\mathcal{N}(k)$ was determined by relation (A-3) of Chapter XIX:

$$\mathcal{N}(k) = \sqrt{\frac{\epsilon_0 \omega}{2\hbar}} \tag{31}$$

Inserting (30) into (29), we get:

$$\begin{aligned} J_a = -\, i\frac{\hbar}{2} \int \mathrm{d}^3 k \sum_{b,c} \Big\{ & \left[\alpha_b^*(\boldsymbol{k}) - \alpha_b(-\boldsymbol{k})\right] \epsilon_{abc} \left[\alpha_c(\boldsymbol{k}) + \alpha_c^*(-\boldsymbol{k})\right] \\ & + \sum_{b,c,d} \left[\alpha_d^*(\boldsymbol{k}) - \alpha_d(-\boldsymbol{k})\right] \epsilon_{abc}\, k_b\, \tilde{\partial}_c \left[\alpha_d(\boldsymbol{k}) + \alpha_d^*(-\boldsymbol{k})\right] \Big\} \end{aligned} \tag{32}$$

Each line in (32) contains four terms: two of these terms include either $\boldsymbol{\alpha}^*$ twice for the first one, or $\boldsymbol{\alpha}$ twice for the second – both will be shown below to be equal to zero; the other two terms contain either $\boldsymbol{\alpha}^*$ once or $\boldsymbol{\alpha}$ once – we show below that they are equal. We finally obtain:

$$\boldsymbol{J} = -i\hbar \int \mathrm{d}^3k \left\{ \underbrace{\boldsymbol{\alpha}^*(\boldsymbol{k}) \times \boldsymbol{\alpha}(\boldsymbol{k})}_{\text{spin}} + \sum_d \underbrace{\alpha_d^*(\boldsymbol{k})\, (\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})\, \alpha_d(\boldsymbol{k})}_{\text{orbital}} \right\} \tag{33}$$

This expression has the same form as (19): the angular momentum is the sum of a spin term and an orbital term involving spatial derivatives. This result confirms that the normal variable $\boldsymbol{\alpha}(\boldsymbol{k})$ can be regarded as the wave function in reciprocal space of the photon field, and that the photon is indeed a spin 1 particle. This result also gives the explicit expressions of the radiation spin angular momentum (first term in the bracket of (33)) and orbital angular momentum (second term).

For a massive particle, we know (Chapitre VI, § D-1-a) that in spherical (or cylindrical) coordinates, the action of the Oz angular momentum component corresponds to a derivation with respect to the azimuthal angle φ:

$$L_z \Rightarrow -i\hbar(k_x\, \tilde{\partial}_y - k_y\, \tilde{\partial}_x) = \frac{\hbar}{i} \frac{\partial}{\partial \varphi} \tag{34}$$

This result simply comes from a calculation of partial derivatives; it is thus also valid for a field.

Computation of the various terms appearing in equation (32)

Consider, in the first line of (32), the terms involving a product of two $\boldsymbol{\alpha}^*$ or two $\boldsymbol{\alpha}$, for example:

$$-i(\hbar/2) \int \mathrm{d}^3k \sum_{b,c} [\alpha_b^*(\boldsymbol{k})\, \epsilon_{abc}\, \alpha_c^*(-\boldsymbol{k})] \tag{35}$$

Changing $\boldsymbol{k}$ into $-\boldsymbol{k}$, inverting the indices b and c, and using $\epsilon_{acb} = -\epsilon_{abc}$, we can show that (35) is equal to its opposite, and hence must be zero. The same approach, followed for the term:

$$+i(\hbar/2) \int \mathrm{d}^3k \sum_{b,c} [\alpha_b(-\boldsymbol{k})\, \epsilon_{abc}\, \alpha_c^*(-\boldsymbol{k})] \tag{36}$$

shows that this term is equal to:

$$-i(\hbar/2) \int \mathrm{d}^3k \sum_{b,c} [\alpha_c(\boldsymbol{k})\, \epsilon_{abc}\, \alpha_b^*(\boldsymbol{k})] \tag{37}$$

which is identical to the other term on the first line of (32) containing one $\boldsymbol{\alpha}^*$ and one $\boldsymbol{\alpha}$, provided we change the relative order in which the term $\alpha_b^*(\boldsymbol{k})$ and $\alpha_c(\boldsymbol{k})$ appear. In classical theory, these quantities are numbers, and hence commute: their order does not matter. It is however useful to keep track of that order in order to obtain an expression still valid when, upon quantization, the $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ will be replaced by the non-commuting creation and annihilation operators.

This computation can be extended to the terms on the second line of (32). In addition to changing $\boldsymbol{k}$ into $-\boldsymbol{k}$, we must also perform an integration by parts. The integrated term,

which yields a surface integral, is zero if the fields tend to zero fast enough at infinity. Added to it is a contribution that shows that the terms containing two $\boldsymbol{\alpha}^*$ or two $\boldsymbol{\alpha}$ are equal to their opposite, and hence equal to zero. On the other hand, the integration by parts shows that the two terms containing one $\boldsymbol{\alpha}^*$ and one $\boldsymbol{\alpha}$ are equal if the order of the $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ can be switched. In the case where the order of the $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ is not taken into account, we obtain expression (33).

2-c. Difference between the angular momenta of massive particles and of radiation

In spite of the strong analogy between equations (19) and (33), we should not forget an important difference between the two angular momenta, arising from the fact that the normal variables $\boldsymbol{\alpha}(\boldsymbol{k})$ are transverse. Maxwell's equation $\boldsymbol{\nabla}\cdot\boldsymbol{E}=0$ for the free field does require $\boldsymbol{\alpha}(\boldsymbol{k})$ to always be perpendicular to the wave vector $\boldsymbol{k}$:

$$\boldsymbol{k}\cdot\boldsymbol{\alpha}(\boldsymbol{k})=0 \qquad\qquad \forall \quad \boldsymbol{k} \tag{38}$$

while the wave function $\boldsymbol{\Psi}(\boldsymbol{k})$ of a massive particle in the reciprocal space is not necessarily perpendicular to $\boldsymbol{k}$. Another difference, of course, is that the norm of this wave function does not have any particular physical meaning (it can arbitrarily be put equal to unity), while changing the norm of the normal variables of the field changes its amplitude.

3. Discussion

3-a. Spin angular momentum of radiation

The spin angular momentum, first term on the right-hand side of (33), can be written as:

$$(S_R)_a=-i\hbar\int \mathrm{d}^3k\ \sum_{b,c}\alpha_b^*(\boldsymbol{k})\,\varepsilon_{b,c}\,\alpha_c(\boldsymbol{k})=-i\hbar\int \mathrm{d}^3k\ \left[\boldsymbol{\alpha}^*(\boldsymbol{k})\times\boldsymbol{\alpha}(\boldsymbol{k})\right]_a \tag{39}$$

Instead of using the components $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ on a basis of three vectors $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, $\boldsymbol{e}_z$ independent of the wave vector $\boldsymbol{k}$ direction, we can choose a basis of three vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$, $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$, $\boldsymbol{\varepsilon}_3(\boldsymbol{k})=\boldsymbol{\kappa}=\boldsymbol{k}/k$, including the unit vector $\boldsymbol{\kappa}$ along $\boldsymbol{k}$ and two other vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$, orthogonal to each other and to $\boldsymbol{\kappa}$, and forming a right-handed reference frame. As the normal variables $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ are transverse, their components on $\boldsymbol{\kappa}$ are zero. In addition, we introduce the two complex linear combinations of $\boldsymbol{e}_1(\boldsymbol{k})$ and $\boldsymbol{e}_2(\boldsymbol{k})$:

$$\begin{aligned}\boldsymbol{\varepsilon}_+(\boldsymbol{k})&=-\left[\boldsymbol{\varepsilon}_1(\boldsymbol{k})+i\boldsymbol{\varepsilon}_2(\boldsymbol{k})\right]/\sqrt{2}\\ \boldsymbol{\varepsilon}_-(\boldsymbol{k})&=+\left[\boldsymbol{\varepsilon}_1(\boldsymbol{k})-i\boldsymbol{\varepsilon}_2(\boldsymbol{k})\right]/\sqrt{2}\end{aligned} \tag{40}$$

corresponding to right and left circular polarizations with respect to the $\boldsymbol{k}$ direction. The transverse normal variables $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ can be expanded on these two vectors:

$$\begin{aligned}\boldsymbol{\alpha}(\boldsymbol{k})&=\alpha_+(\boldsymbol{k})\boldsymbol{e}_+(\boldsymbol{k})+\alpha_-(\boldsymbol{k})\boldsymbol{e}_-(\boldsymbol{k})\\ \boldsymbol{\alpha}^*(\boldsymbol{k})&=\alpha_+^*(\boldsymbol{k})\boldsymbol{e}_+^*(\boldsymbol{k})+\alpha_-^*(\boldsymbol{k})\boldsymbol{e}_-^*(\boldsymbol{k})\end{aligned} \tag{41}$$

Using these two expansions, we compute the cross product $\boldsymbol{\alpha}^*(\boldsymbol{k}) \times \boldsymbol{\alpha}(\boldsymbol{k})$. Since:

$$\begin{aligned}
\boldsymbol{\varepsilon}_+^* \times \boldsymbol{\varepsilon}_+ &= \frac{1}{2}(\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) \times (\boldsymbol{\varepsilon}_1 + i\boldsymbol{\varepsilon}_2) = \frac{i}{2}(\boldsymbol{\varepsilon}_1 \times \boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_2 \times \boldsymbol{\varepsilon}_1) = i\boldsymbol{\kappa} \\
\boldsymbol{\varepsilon}_-^* \times \boldsymbol{\varepsilon}_- &= \frac{1}{2}(\boldsymbol{\varepsilon}_1 + i\boldsymbol{\varepsilon}_2) \times (\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) = \frac{-i}{2}(\boldsymbol{\varepsilon}_1 \times \boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_2 \times \boldsymbol{\varepsilon}_1) = -i\boldsymbol{\kappa} \\
\boldsymbol{\varepsilon}_+^* \times \boldsymbol{\varepsilon}_- &= \frac{-1}{2}(\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) \times (\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) = 0 = \boldsymbol{\varepsilon}_-^* \times \boldsymbol{\varepsilon}_+
\end{aligned} \tag{42}$$

we get:

$$\boldsymbol{S}_R = \int \mathrm{d}^3k \left\{ \left[\alpha_+^*(\boldsymbol{k})\alpha_+(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} - \left[\alpha_-^*(\boldsymbol{k})\alpha_-(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} \right\} \tag{43}$$

The form of this expression, "diagonal with respect to the spin variables", has a clear physical significance: to each plane wave with wave vector $\boldsymbol{k}$ and a right (left) polarization with respect to $\boldsymbol{k}$, correspond photons of momentum $\hbar\boldsymbol{k}$ and spin angular momentum $+\hbar$ $(-\hbar)$ along the direction $\boldsymbol{\kappa}$ of $\boldsymbol{k}$. Upon quantization, when the normal variables $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ are replaced by creation and annihilation operators, expression (43) becomes:

$$\hat{\boldsymbol{S}}_R = \int \mathrm{d}^3k \left\{ \left[\hat{a}_+^\dagger(\boldsymbol{k})\hat{a}_+(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} - \left[\hat{a}_-^\dagger(\boldsymbol{k})\hat{a}_-(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} \right\} \tag{44}$$

Operator $\hat{a}_+^\dagger(\boldsymbol{k})$ creates a photon with momentum $\hbar\boldsymbol{k}$ and spin angular momentum $+\hbar$ along the direction $\boldsymbol{\kappa}$ of $\boldsymbol{k}$; operator $\hat{a}_+(\boldsymbol{k})$ annihilates that photon, and operator $\hat{a}_+^\dagger(\boldsymbol{k})\hat{a}_+(\boldsymbol{k})$ corresponds to the number of photons in that mode. An analogous definition applies to the second term of (44) with a change of sign for the angular momentum.

Helicity

These results, which arise from the transverse character of free radiation, lead us to introduce the so-called "*helicity*". It is the projection of the photon spin angular momentum onto the direction $\boldsymbol{\kappa}$ of the wave vector $\boldsymbol{k}$, equal to $+1$ for photons with a right circular polarization with respect to $\boldsymbol{\kappa}$, and -1 for photons with a left circular polarization. The transverse character of the free radiation field forbids the photons to have zero helicity. Note also that helicity is a *pseudoscalar*: upon reflection in space, the polar vector $\boldsymbol{\kappa}$ changes sign whereas the spin vector $\boldsymbol{S}$, an axial vector, remains unchanged. Consequently, the scalar product of $\boldsymbol{\kappa}$ and $\boldsymbol{S}$ changes sign (as opposed to a scalar).

3-b. Experimental evidence of the radiation spin angular momentum

Consider a plane wave of wave vector $\boldsymbol{k}$ and polarization $\boldsymbol{\varepsilon}$. Using the expressions for the electric field $\boldsymbol{E}(\boldsymbol{r})$ and the vector potential $\boldsymbol{A}(\boldsymbol{r})$ given in Chapter XVIII, one can easily compute the two terms of equation (28) yielding, in position space, the radiation spin angular momentum and orbital angular momentum[2]. The result is that the radiation orbital angular momentum – second term of (28) – is always zero, whatever the

[2]These are the two terms that, transformed into the reciprocal space, and after the introduction of the normal variables, yield the two terms on the right-hand side of equation (33); by comparison with the expression of the angular momentum of a spin 1 particle, these terms have been interpreted as the two components of the radiation angular momentum.

polarization $\boldsymbol{\varepsilon}$ is. As for the spin angular momentum – first term of (28) – it is zero for a linear polarization, but different from zero for a circular polarization, with opposite signs for the right and left circular polarizations. This validates, in the simple plane wave case, the general conclusions of the previous paragraph.

Such a result suggests sending a linearly polarized light beam through a quarter-wave plate. Assuming the plate transforms the incident linear polarization into a right (left) circular polarization, the incident photons have a zero spin angular momentum before they go through the plate, and equal to $+\hbar$ $(-\hbar)$ as they come out of the plate. The radiation spin angular momentum thus changes as beam goes through the plate, and this must be accompanied by a change, in the opposite direction, of the plate's angular momentum. Suspending the plate by a thin torsion fiber, one should observe a rotation of the plate induced by the incident radiation, in a direction opposite to that of the circular polarization of the outgoing beam. This experiment, suggested by A. Kastler [30] was performed by R. Beth [31], confirming the existence of angular momentum transfer.

Comment

A paradox arises when computing, again for a plane wave, the angular momentum of the radiation, given by equation (20). In a plane wave, the Poynting vector $\boldsymbol{\Pi}(\boldsymbol{r}) = \boldsymbol{E}(\boldsymbol{r}) \times \boldsymbol{B}(\boldsymbol{r})$ is always parallel to the wave vector $\boldsymbol{k}$ at any point $\boldsymbol{r}$, and for any polarization. The integral over the entire space of $\boldsymbol{r} \times \boldsymbol{\Pi}(\boldsymbol{r})$ must then be zero. As the orbital angular momentum is also zero, it seems that the spin angular momentum should also be zero, whatever the polarization is. This paradox arises because infinite plane waves do not exist in the physical world: any real light beam has a finite spatial extension. The authors of [32] (see also [33]) show that the circular polarization at the center of the beam changes when the field amplitude changes around the edge of the beam. Taking this effect into account quantitatively confirms the result obtained above, namely that the beam spin angular momentum is equal to the sum of the angular momenta $\pm\hbar$ of that beam's photons.

3-c. Orbital angular momentum of radiation

An important difference between the radiation orbital and spin angular momenta is clearly seen in expression (28): the definition of orbital angular momentum involves a reference point O, since the vector $\boldsymbol{r}$, defined with respect to that point, explicitly appears in the expression for the orbital angular momentum. This is not the case for the spin angular momentum which, for this reason, is sometimes called "intrinsic" angular momentum. There are actually at least two cases where the choice of the point O is obvious, cases that we now analyze.

α. *Multipolar waves*

When studying the radiation emitted or absorbed by an atom or a nucleus between two discrete states, a natural choice for analyzing the exchanges of angular momentum between the system internal variables and the photons is the center of mass of that atom or that nucleus. In the next complement C_{XIX}, we study for example the exchange of angular momentum between photons and the internal variables of an atom in a particular case: an electric dipole transition, in the long wavelength approximation. Consequently, the expressions describing the photon absorption only involve the radiation polarization variables, and hence only the photon spin angular momentum; the radiation orbital

angular momentum does not actually play any role[3].

There are other transitions, especially for atomic nuclei, where the variation of the internal angular momentum between the two transition states is larger than or equal to 2; the photon spin angular momentum, equal to 1, can then no longer ensure the conservation of the angular momentum. Radiation states having a total angular momentum J larger than 1 must come into play, which implies a contribution from the radiation orbital angular momentum. Waves corresponding to such states are called "multipolar waves".

The simplest way to build multipolar waves having a total angular momentum characterized by the quantum numbers J and M is to associate a spherical harmonic $Y_\ell^{m_\ell}(\boldsymbol{\kappa})$ with a spin $S = 1$; the spherical harmonic is an eigenfunction of $\boldsymbol{L}^2$ and L_z with eigenvalues $\ell(\ell+1)\hbar^2$ and $m_\ell\hbar$; the spin $S = 1$ has three eigenstates $|S, m_S\rangle$, with $m_S = +1, 0, -1$, isomorphic to the three polarization states $\boldsymbol{e}_+ = -(\boldsymbol{e}_x + i\boldsymbol{e}_y)/\sqrt{2}$, $\boldsymbol{e}_z$, $\boldsymbol{e}_- = (\boldsymbol{e}_x - i\boldsymbol{e}_y)/\sqrt{2}$. We therefore obtain a *vector spherical harmonic*:

$$\boldsymbol{Y}_{J,\ell,1}^M(\boldsymbol{\kappa}) = \sum_{m_\ell}\sum_{m_S} \langle JM|\ell 1 m_\ell m_S\rangle \; Y_\ell^{m_\ell}(\boldsymbol{\kappa})\boldsymbol{e}_{m_S} \tag{45}$$

which is an eigenfunction of $\boldsymbol{J}^2$, $\boldsymbol{L}^2$, J_z, with eigenvalues $J(J+1)\hbar^2$, $\ell(\ell+1)\hbar^2$, $M\hbar$. In this equation, the first term on the right-hand side is a Clebsch-Gordan coefficient (Chapter X, § C-4-c), J can take one of the three values $J = \ell - 1, J = \ell, J = \ell + 1$ and $M = m_\ell + m_S$.

A difficulty is that the vector spherical harmonics are not all transverse functions and hence cannot be used as a basis of normal transverse functions for expanding the radiation field. For a given value of J, one can nevertheless build linear superpositions of vector spherical harmonics $\boldsymbol{Y}_{J,\ell,1}^M(\boldsymbol{\kappa})$ with $\ell = J, \ell = J \pm 1$ that are transverse and have, in addition, a well defined parity $\pi = \pm 1$. Each vector spherical harmonic, which only depends on the direction $\boldsymbol{\kappa}$ of $\boldsymbol{k}$, can also be multiplied by $\delta(k - \omega_0/c)$, hence yielding a function that is also an eigenfunction of the energy, with eigenvalue $\hbar\omega_0$. These functions form a possible basis of normal transverse variables for expanding the field; they are characterized by four quantum numbers: the energy $\hbar c k_0$, the total angular momentum J, M, and the parity π. They are called electric (for $\pi = -1$) or magnetic (for $\pi = +1$) multipolar waves. As this book is limited to the study of electric or magnetic dipole transitions, we do not give here the general expressions for multipolar waves. More details can be found in complement B_I of [16] and in [34].

β. Beams with cylindrical symmetry around one axis

It often happens that the beams under study have a cylindrical symmetry. This is the case, for example, for Gaussian beams propagating along a z axis, and whose transverse sections are circular. If the reference point O is taken on the axis, the beam symmetry causes the orbital angular momentum to be necessarily along the z axis and to have the same value regardless of the position of O along this axis. If the reference point O is taken outside this axis, the orbital angular momentum will change, but not its component L_z, which exhibits an intrinsic character.

[3] This orbital angular momentum may, however, come into play during the angular momentum exchanges with the atom's external variables for certain types of light beams, the Laguerre-Gaussian light beams (§ 3-b of complement C_{XIX}).

A particularly interesting case concerns the Laguerre-Gaussian beam (LG) whose field has an $\exp(im\varphi)$ dependence with respect to the azimuth angle φ that defines the direction of a point in the plane perpendicular to the beam axis. The cylindrical symmetry is preserved since a rotation of the beam of an angle φ_0 around the z axis yields the same field to within a global phase factor $\exp(-im\varphi_0)$. Relation (34) then shows that the Oz component of the orbital angular momentum of each photon of the LG beam is $m\hbar$.

Consider an LG beam propagating along the z axis, with a k wave vector along that axis. The phase at a point with cylindrical coordinates z, ρ, φ is that of $\exp i(kz + m\varphi)$. For an ordinary Gaussian beam (for which $m = 0$) the surfaces of constant phase are, in the vicinity of the focal point, planes perpendicular to the z axis. When $m \neq 0$, z and φ must both vary for the phase to remain constant, following the relation $k\mathrm{d}z + m\mathrm{d}\varphi = 0$. The surfaces of constant phase are therefore helicoidal surfaces spanned by a half-line perpendicular to the z axis, starting from this axis, and which rotates of an angle $-2\pi/m$ when z increases by λ. It is not surprising that under such conditions the field has a non-zero orbital angular momentum. Note also that the field must be zero on the z axis (otherwise its phase would vary discontinuously upon crossing that axis).

In conclusion, we showed in this complement that there are two types of radiation angular momenta, the spin angular momentum and the orbital angular momentum, and we studied their properties. The photon can be viewed as a spin 1 particle, except for the fact that it only has two (instead of three) internal states, with respective heliticity $+1$ and -1. We shall see in the next complement how the interactions between radiation and atoms permit transferring angular momentum from the first to the others.

Complement C_{XIX}

Angular momentum exchange between atoms and photons

1 Transferring spin angular momentum to internal atomic variables 2056

1-a Electric dipole transitions 2056

1-b Polarization selection rules 2056

1-c Conservation of total angular momentum 2058

2 Optical methods 2058

2-a Double resonance method 2059

2-b Optical pumping 2062

2-c Original features of these methods 2064

3 Transferring orbital angular momentum to external atomic variables 2065

3-a Laguerre-Gaussian beams 2065

3-b Field expansion on Laguerre-Gaussian modes 2065

Introduction

Angular momentum exchanges between atoms and photons are at the base of several experimental methods that played an important role in atomic physics and laser spectroscopy. The aim of this complement is to analyze the selection rules appearing in the photon absorption and emission processes by an atom; they express the conservation of the total angular momentum of the system atom + radiation during these processes.

We shall mainly focus on the transfer of angular momentum to the atomic internal (electronic) variables (§ 1). The polarization of the field, characterizing the spin angular momentum of that field (Complement B_{XIX}), then plays an essential role. For the absorption process, we will establish the selection rules relating the field polarization to the variation of the "magnetic quantum number" m characterizing the projection of the total internal angular momentum of the atom on a given axis. Two important applications of these selection rules will be described in § 2: the double resonance method, and optical pumping. We shall see that the proper choice of the polarization of the exciting light beam, and of both the direction and polarization of the detected light emitted by the excited atoms, allows controlling the atomic Zeeman sublevels that can be populated and detected by light. We shall emphasize in § 2 the importance of this selectivity. The transfer of the radiation orbital angular momentum to the atomic external variables will be briefly addressed in § 3.

1. Transferring spin angular momentum to internal atomic variables

1-a. Electric dipole transitions

We shall limit ourselves to the case where the transitions between the atomic internal states are electric dipole transitions. As seen in Chapter XIX (§ C-4), the interaction Hamiltonian between atom and radiation can then be written in the form:

$$\hat{H}_I = -\hat{\boldsymbol{D}} \cdot \hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) \tag{1}$$

where $\hat{\boldsymbol{D}}$ is the operator associated with the atomic electric dipole moment and $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ is the radiation transverse electric field operator at point $\hat{\boldsymbol{R}}$, the position of the atom's center of mass. Note that all the results established in this complement are still valid for magnetic dipole transitions; one must simply replace $\hat{\boldsymbol{D}}$ by the atomic magnetic dipole moment operator $\hat{\boldsymbol{M}}$ and $\hat{\boldsymbol{E}}_\perp$ by the radiation magnetic field $\hat{\boldsymbol{B}}$ operator. For the transitions where one photon is absorbed, the expressions (B-3) and (B-4) of Chapter XIX of the fields $\hat{\boldsymbol{E}}_\perp$ and $\hat{\boldsymbol{B}}$ can be replaced by the part containing only destruction operators, called "positive frequency component" (*cf.* § A-3 of Chapiter XX) and denoted $\hat{\boldsymbol{E}}_\perp^{(+)}$ and $\hat{\boldsymbol{B}}^{(+)}$. For the transitions where one photon is emitted, these fields can be replaced their components $\hat{\boldsymbol{E}}_\perp^{(-)}$ and $\hat{\boldsymbol{B}}^{(-)}$ containing only creation operators, called "negative frequency components".

Replacing $\hat{\boldsymbol{E}}_\perp^{(+)}$ and $\hat{\boldsymbol{B}}^{(+)}$ by their plane wave expansions, the internal variables only appear in the scalar products of $\hat{\boldsymbol{D}}$, or $\hat{\boldsymbol{M}}$, with the polarization vector $\boldsymbol{\varepsilon}$ of the plane wave $\boldsymbol{k}$. We shall assume in § 1 and § 2 that all the incident radiation states are plane waves, or linear superpositions of plane waves with wave vectors having directions very close to an optical axis (paraxial approximation). These waves are supposed to have the same polarization $\boldsymbol{\varepsilon}$, which means that the beam angular aperture must be sufficiently small. The transitions between the internal atomic states[1] g and e we shall consider are thus entirely characterized by the matrix elements $\langle e|\boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{D}}|g\rangle$.

1-b. Polarization selection rules

We start with the simple case of an atom with a single electron, and a transition between a ground state g with an orbital angular momentum $\ell = 0$ and an excited state e with an angular momentum $\ell = 1$. We assume the radiation has a right circular polarization, noted σ_+, with respect to an axis noted z: this means that the radiation electric field rotates around that axis *following the right-hand rule*, at the angular frequency ω. We have:

$$\boldsymbol{\varepsilon} = -\frac{1}{\sqrt{2}}(\boldsymbol{e}_x + i\boldsymbol{e}_y) \tag{2}$$

Since $\hat{\boldsymbol{D}} = q\hat{\boldsymbol{r}}$, where q is the electron charge and $\hat{\boldsymbol{r}}$ its position with respect to the nucleus, we have:

$$\boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{D}} = -\frac{q}{\sqrt{2}}(x + iy) = -\frac{q}{\sqrt{2}} r \sin\theta \exp(i\varphi) \tag{3}$$

[1] In this complement and as is usually done in the literature, we shall note g the ground state and e the excited state (instead of using our previous notation of a and b for the two atomic levels).

where r, θ, φ are the spherical coordinates of the electron with respect to the nucleus. Starting from a ground state with a magnetic quantum number $m_g = 0$ (defined with respect to the z axis) and with no dependence on φ, the excitation with σ_+ polarized light creates an excited state wave function that now varies with φ as $\exp(i\varphi)$. This wave function is an eigenfunction of the component $\hat{L}_z = (\hbar/i)\partial/\partial\varphi$ of the orbital angular momentum, with eigenvalue $+1$ (it corresponds to the state $m_e = +1$). In a similar way, an excitation with a σ_- polarization, for which $\boldsymbol{\varepsilon} = \frac{1}{\sqrt{2}}(\boldsymbol{e}_x - i\boldsymbol{e}_y)$, transfers the atom into the state $m_e = -1$. Consider finally an excitation with a π polarized light, for which $\boldsymbol{\varepsilon} = \boldsymbol{e}_z$: the electric field has a linear polarization parallel[2] to the Oz axis. We then have $\boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{D}} = qz = qr\cos\theta$, with no φ dependence; the excitation brings the atom to the state $m_e = 0$. To sum up, the polarization selection rules for a transition $\ell_g = 0 \longrightarrow \ell_e = 1$ are given by:

$$\sigma_+ \longrightarrow m_e = +1, \qquad \pi \longrightarrow m_e = 0 \qquad \sigma_- \longrightarrow m_e = -1. \tag{4}$$

The previous results can easily be generalized to any transition going from a ground state g with angular momentum J_g to an excited state e with angular momentum J_e, for an atom with any number of electrons. One simply has to use the Wigner-Eckart theorem (Complement $\mathrm{D_X}$ and exercise 8 in Complement $\mathrm{G_X}$). The dipole operator $\hat{\boldsymbol{D}}$ (sum of the dipole operators for each individual electron) is a vector operator whose three spherical components $\hat{\boldsymbol{D}}_s$ with $s = +1, 0, -1$ are equal to :

$$\hat{D}_{+1} = \frac{-1}{\sqrt{2}}(\hat{D}_x + i\hat{D}_y) \tag{5a}$$

$$\hat{D}_0 = \hat{D}_z \tag{5b}$$

$$\hat{D}_{-1} = \frac{1}{\sqrt{2}}(\hat{D}_x - i\hat{D}_y) \tag{5c}$$

The Wigner-Eckart theorem (Complement $\mathrm{D_X}$) states that

$$\langle J_e, m_e|\hat{\boldsymbol{D}}_s|J_g, m_g\rangle = \langle e||\hat{D}||g\rangle \ \langle J_e, m_e|J_g, 1; m_g, s\rangle \tag{6}$$

where the last term on the right-hand side is a Clebsch-Gordan coefficient (Chapter X, § C-4-c) and the first term is a "reduced matrix element" independent of m_g, s and m_e. The Clebsch-Gordan coefficient is different from zero only if:

- a triangle can be formed with J_e, 1 and J_g, which means that J_e is equal either to J_g, or to $J_g \pm 1$, the transition $J_e = 0 \leftrightarrow J_g = 0$ being forbidden;
- $m_e = m_g + s$.

As the three values of s ($s = +1, 0, -1$) correspond to the three polarizations σ_+, π, σ_- respectively, the selection rules (4) for any given transition $J_g \leftrightarrow J_e$ can be generalized to (see Figure 1):

$$\sigma_+ \longrightarrow m_e = m_g + 1, \qquad \pi \longrightarrow m_e = m_g \qquad \sigma_- \longrightarrow m_e = m_g - 1 \tag{7}$$

[2]Because of the transversality of the field, the light beam must then propagate in a direction perpendicular to the Oz axis.

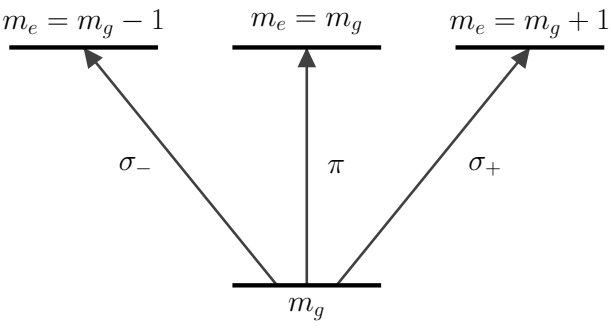


Figure 1: Selection rules for an electric or magnetic dipole transition. The magnetic quantum number m increases by one unit for an excitation with a σ_+ polarization, remains unchanged for a π polarization and decreases by one unit for a σ_- polarization.

1-c. Conservation of total angular momentum

We saw that, when an atom absorbs a photon having a polarization σ_+ with respect to a z axis, the component of the atom's angular momentum along that axis increases by one (in $\hbar$ units). The conservation of the total angular momentum means that the absorbed σ_+ photon must have an angular momentum $+\hbar$ along the z axis.

This result can also be obtained from the study of the radiation angular momentum presented in Complement B_{XIX}, as we now show. Although the interaction Hamiltonian pertaining to the atomic internal variables does not depend on the photon wave vector[3] $\boldsymbol{k}$, one can always choose a $\boldsymbol{k}$ wave vector parallel to the atomic quantization axis (since these two directions must be perpendicular to the polarization vector of the circular wave). Now we saw in Complement B_{XIX} that in a plane wave (or in a beam of plane waves with a small angular aperture) a photon with polarization σ_+ with respect to the wave vector has a total angular momentum (actually reduced to its spin angular momentum) equal to $+\hbar$ and parallel to its wave vector. The total angular momentum is indeed conserved.

2. Optical methods

The polarization selection rules show that it is possible to selectively excite a Zeeman sublevel of an atomic excited state. In a similar way, we shall see below that the observation of the emitted light in a given direction with a specific polarization allows determining from which excited sublevel the light was emitted. Such a selectivity in excitation and detection is the base of the optical methods for Hertzian spectroscopy, as will be illustrated below with two examples.

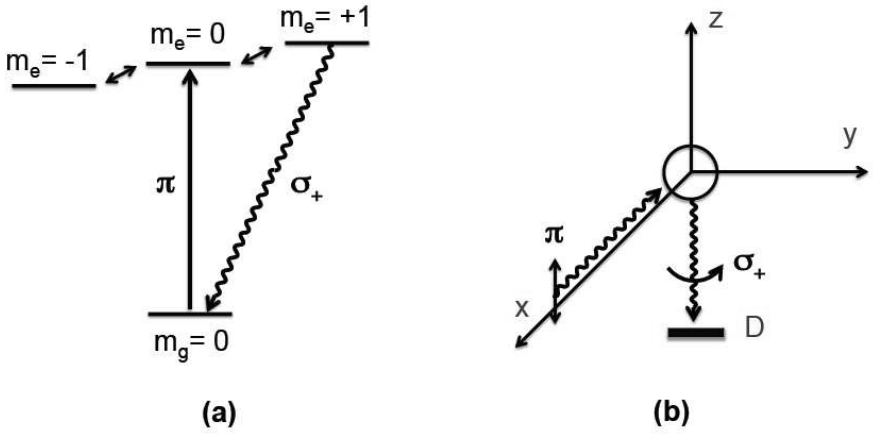


Figure 2: Energy diagram and Zeeman structure of the $6^1S_0 \longrightarrow 6^3P_1$ transition of the even isotopes of mercury at 253.7 nm (Fig.a). A static magnetic field, applied along the z axis, lifts the degeneracy of the excited state, which is split into three equidistant Zeeman sublevels: $m_e = -1, 0, +1$. A resonant light beam, propagating along the x axis, with a linear polarization π parallel to the z axis (Fig.b), selectively excites the atoms into the $m_e = 0$ sublevel (upward arrow on Fig.a). When they are in the excited state, the atoms are transferred from the $m_e = 0$ state to both states $m_e = \pm 1$ by a resonant radiofrequency field (oblique small double arrows on Fig.a). A detector D placed along the z axis, just behind an analyser that only transmits light with σ_+ polarization along the z axis, detects the light emitted by the atoms (Fig.b). Consequently, the detector is only sensitive to the light emitted from the $m_e = +1$ sublevel (wiggly arrow on Fig.a); it yields a signal proportional to the population of that sublevel.

2-a. Double resonance method

We now explain the principle of the method taking as an example the even isotopes of mercury, for which the first theoretical predictions were made by Brossel and Kastler [35]; the first experimental evidences were obtained by Brossel and Bitter [36]. Since for these isotopes the nuclear spin is zero, and as in a mercury atom all the electron shells are filled in the ground state, the energy diagram is particularly simple. The ground state g has a zero angular momentum ($J_g = 0$), and the first accessible excited states, an angular momentum $J_e = 1$. In the presence of a static magnetic field B applied along the z axis, the three Zeeman sublevels $m_e = -1, 0, +1$ of the excited state undergo Zeeman shifts proportional to m_e and to the applied field, so that the energy of the m_e state is written:

$$E_{m_e} = E_0 + g_L m_e \mu_B B \tag{8}$$

[3] In the long wavelength approximation, the radiation wave vector $\boldsymbol{k}$ no longer appears in the interaction Hamiltonian for the atomic internal variables. This wave vector only appears in the part of the interaction Hamiltonian, $\exp(i\boldsymbol{k} \cdot \hat{\boldsymbol{R}})$, pertaining to the external variables.

where E_0 is the excited state energy in the absence of the field, g_L is the Landé g-factor of that state, and μ_B the Bohr magneton. The ground state $J_g = 0, m_g = 0$, whose energy is chosen to be zero, is not affected by the B field (Figure 2a).

In the double resonance method, the atoms are selectively excited by a resonant light beam with polarization π into the sublevel $m_e = 0$. For example, the exciting beam propagates along the x axis, with a polarization $\boldsymbol{\varepsilon}_z$ perpendicular to the x axis, and parallel to the z axis, hence with a π polarization (Figure 2b).

If the atoms were left alone, with no perturbation while in the excited state during its radiative lifetime $\tau_R = 1/\Gamma$ (of the order of 1.5 10^{-7} sec), they would remain in that sublevel for the entire time they stay in the excited state. On the other hand, if they are subjected to a resonant radiofrequency field[4] that is strong enough to make them go from $m_e = 0$ to $m_e = \pm 1$ during the excited state lifetime τ_R, the two states $m_e = \pm 1$ will be equally populated.

Is it possible, observing the light emitted by the atoms as they go back to the ground state by spontaneous emission of a photon, to determine the sublevel m_e from which the light was emitted, and hence obtain a signal proportional to this sublevel population? This problem must be analyzed with more care than the absorption process for the following reason. Once the atom has reached the excited sublevel m_e, it can emit in any direction with all possible polarizations, which are not necessarily the three basic polarizations σ_+, π or σ_-. We are going to show that one can place the detector in a specific direction, far from the atom, and put in front of it a polarization analyzer suitably chosen so as to be able to determine from which sublevel the detected photon was emitted.

To demonstrate this result, it is important to first study the oscillations of the atomic dipole associated with a transition $m_g \longleftrightarrow m_e$. Let us assume the detector is placed on the z axis where the atom is located.

(i) For the transition $m_e = 0 \longleftrightarrow m_g = 0$, as the dipole oscillates along the z axis, it does not emit along that axis. The detector does not receive any fraction of the light emitted by an atom in the $m_e = 0$ state.

(ii) On the other hand, the dipole associated with the transition $m_g = 0 \longleftrightarrow m_e = +1$, which rotates around the z axis following the right-hand rule, in a plane perpendicular to that axis, emits along that axis light with a σ_+ polarization; this light yields a signal on the detector equipped with an analyzer selecting right circular polarization. This analyzer will, however, block the light emitted by an atom in the $m_e = -1$ sublevel, which has a left circular polarization.

To sum up, placing a detector in a well chosen direction, preceded by an analyzer selecting a suitable polarization, one can block out all light except the one coming from a specific unique m_e sublevel, and get a signal proportional to that sublevel population.

The principle of this double resonance experiment is to selectively excite the atoms in the $m_e = 0$ sublevel, and to detect, by observing the σ_+ light emitted along the z axis, the variations of the number of atoms transferred into the $m_e = +1$ sublevel by a radiofrequency field with angular frequency ω_{RF} close to the Zeeman frequency $\omega_Z = g_L \mu_B B/\hbar$. Observing the variation of the emitted light as one scans ω_{RF} for a fixed value of the magnetic field, or as one scans the magnetic field for a fixed ω_{RF}, one

[4] The atoms are thus submitted to two resonant excitations: an optical resonant excitation that brings them from $m_g = 0$ to $m_e = 0$; a radiofrequency resonant excitation that brings them from $m_e = 0$ to $m_e = \pm 1$. This is why this method is called "double resonance method".

can optically detect the magnetic resonance in the excited state.

Comment

What happens if the light emitted by the atom is not observed along the z axis, with a right circular analyzer, but with a detector placed along the y axis, behind an analyzer selecting a linear polarization parallel to the x axis, and hence perpendicular to the z axis, a polarization called "sigma"[5]? The light emitted along the y axis by the atom in the $m_e = 0$ state has a linear polarization parallel to the oscillating dipole, and hence parallel to the z axis. It is blocked by the analyzer that only lets through the orthogonal polarization. On the other hand, whether the atom is in the $m_e = +1$ or $m_e = -1$ sublevel, the rotating dipole in the xOy plane (following the right-hand or left-hand rule) emits in that plane light with a linear polarization perpendicular to the z axis; that light can be detected by the sigma polarization detector and it yields a signal proportional to the sum of the populations of both sublevels $m_e = +1$ and $m_e = -1$. Actually, the resonant radiofrequency field excites a linear superposition of these two Zeeman sublevels. It was later discovered that the waves emitted from these two $m_e = +1$ and $m_e = -1$ states gave rise to interference, hence modulating at the frequency $2\omega_{RF}$ the detected sigma light intensity[6]. The detector signal therefore contains a component modulated at the frequency $2\omega_{RF}$, on top of a continuous component (in steady state), proportional to the sum of the two populations of the Zeeman sublevels. This continuous component was the signal used in the first double resonance experiment.

The shape of the magnetic resonance line can be exactly computed; it leads to analytical expressions in excellent agreement with experimental observations. The center of the resonance line yields the Landé g-factor of the excited state, i.e. the magnetic moment of that state. The resonance width, extrapolated to zero radiofrequency intensity to eliminate radiative broadening, yields the natural Γ width of the excited state.

Calculation of the line shape

Broadband excitation with a π polarization prepares, in a quasi-instantaneous way, the atom in the $m_e = 0$ excited state. In the steady state, n_0 atoms per unit time are excited into that state. Each atom then evolves, because of its interaction with the radiofrequency field $\boldsymbol{B}_1$, and its state becomes a linear superposition of the three sublevels $m_e = -1, 0, +1$. Let us assume the radiofrequency field is a rotating field, which allows introducing the associated rotating reference frame (Complement $\mathrm{F_{IV}}$) where the atom's evolution actually becomes a simple rotation around $\boldsymbol{B}_1$. Using the rotation matrix for a spin 1, one can find the expression for $P(m_e = 0 \longrightarrow m_e = +1, t)$, the probability for the atom, initially in the $m_e = 0$ state, to be found after a time t in the $m_e = +1$ state [38]. Because of the radiative lifetime $\tau_R = 1/\Gamma$ of the excited sublevels, this probability is reduced by a factor $e^{-\Gamma t}$. Consequently, in steady state, the number of atoms transferred to the $m_e = +1$ state is equal to[7]:

$$N_{+1} = n_0 \int_0^\infty P(m_e = 0 \longrightarrow m_e = +1, t) e^{-\Gamma t} \mathrm{d}t \tag{9}$$

Using the expression for $P(m_e = 0 \longrightarrow m_e = +1, t)$, one finally obtains:

$$N_{+1} = \frac{n_0 \Omega^2}{2\Gamma} \frac{4\delta^2 + \Gamma^2 + \Omega^2}{(\delta^2 + \Gamma^2 + \Omega^2)(4\delta^2 + \Gamma^2 + 4\Omega^2)} \tag{10}$$

[5]This set-up was the one used in the first double resonance experiment [36].
[6]Such modulations, called "light beats", were observed in 1959 [37].
[7]A similar computation can be carried out for N_{-1}.

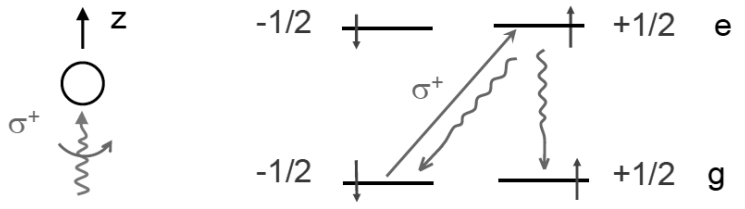


Figure 3: Principle of optical pumping for a $J_g = 1/2 \rightarrow J_e = 1/2$ transition. The resonant absorption of a photon with σ^+ polarization selectively excites the $m_g = -1/2 \longrightarrow m_e = 1/2$ transition. Once it has reached its excited state, the atom falls back, through spontaneous emission, into the $m_g = \pm 1/2$ states. If it falls in the $m_g = +1/2$ state, it can no longer absorb an incident photon, and remains in that state (since there is no transition σ^+ originating from the $m_g = +1/2$ state where the atoms accumulate). In addition, any change in the population difference between the $m_g = \pm 1/2$ sublevels can be detected by a change in the incident beam absorption, since this absorption is only possible starting from the $m_g = -1/2$ sublevel.

In this expression, Ω is the Rabi frequency associated with the radiofrequency field $\boldsymbol{B}_1$, proportional to that field's amplitude; $\delta = \omega_{RF} - \omega_Z$ is the difference between the frequency ω_{RF} of the RF field and the Zeeman frequency ω_Z associated with the gap between the Zeeman sublevels, which is proportional to the static field B_0. The resonance is plotted by scanning either the frequency of the RF field, or the static field, which amounts to scanning ω_Z.

2-b. Optical pumping

Optical pumping, proposed by A. Kastler [39], extends to atomic ground states the essential ingredients of the double resonance method. It also opens the possibility of achieving, in a steady state, large population differences between Zeeman sublevels in the ground state.

We shall explain this method's principle in the simple case where the ground state g (as well as the excited state e) has only two Zeeman sublevels $m_g = \pm 1/2$ (or $m_e = \pm 1/2$ for the excited state). The ideas introduced for that example remain valid for more complex transitions where the e and g states have a higher degeneracy.

The principle of optical pumping is illustrated in Figure 3. Atoms are excited by a resonant beam with σ^+ polarization, propagating along the z axis (left-hand side of the figure). A static magnetic field is also applied along that same axis. The absorption of a resonant photon with σ^+ polarization is selective, meaning it can only excite the $m_g = -1/2 \longrightarrow m_e = 1/2$ transition, during which the magnetic quantum number m increases by one unit. Once in the excited state, and after an average time $\tau_R = 1/\Gamma$, the atom falls back, through spontaneous emission, to the $m_g = \pm 1/2$ states. The probabilities of the various possible transitions are proportional to the square of the Clebsch-Gordan coefficients. If the atom falls into the $m_g = -1/2$ state, it can reabsorb a σ_+ photon. After a certain number of cycles, it will eventually fall in the $m_g = +1/2$ state, where it can no longer absorb an incident photon. It will remain in that state (since no σ^+ transition originates from the $m_g = +1/2$ state) and the atoms will accumulate in

that sublevel. The cycles of absorption of a σ_+ photon starting from the $m_g = -1/2$ state followed by a spontaneous emission a photon transferring the atom in the $m_g = +1/2$ level can be considered as an "optical pump" that empties the $m_g = -1/2$ sublevel to fill the $m_g = +1/2$ sublevel, hence the name "optical pumping".

Balance of angular momentum exchanges

During an optical pumping cycle, the atom gains an angular momentum $+\hbar$ since it goes from the state $m_g = -1/2$ to the state $m_g = +1/2$. The radiation looses one unit $\hbar$ of angular momentum as a σ_+ photon is absorbed. Since the total angular momentum of the system atom + radiation is conserved, the angular momentum of the radiation emitted by the atom going by spontaneous emission from the state $m_e = +1/2$ to the state $m_g = +1/2$ must be zero.

In order to directly demonstrate that the radiation does not lose, on average, any angular momentum during this transition, one must take into account that, as it goes from the $m_e = +1/2$ state to the $m_g = +1/2$ state, the atom actually emits a spherical wave in all directions, with all possible transverse polarizations; it is therefore not correct to say that the atom emits a single photon with a π polarization, which would imply the atom only emits photons with a wave vector perpendicular to the z axis. The correct way is to calculate the total (orbital and spin) angular momentum of the spherical wave emitted by the atom going from $m_e = +1/2$ to $m_g = +1/2$. We give a brief outline of the computation[8] as this involves expanding the field no longer onto plane waves but onto multipolar waves (Complement B_{XIX}, § 3-c-α). Those waves are field modes characterized by four quantum numbers: wave number k (or energy), parity equal to $+1$ or -1, total angular momentum J (which must be an integer), and finally the M_J component of the angular momentum J on the quantization axis z (which varies by unit steps between $-J$ and $+J$). For an electric dipole transition such as the one studied here, the parity equals -1 and the total angular momentum J equals 1. It can be shown that M_J is equal to 0 when the atom goes from m_e to $m_g = m_e$, and is equal to ± 1 when the atom goes from m_e to $m_g = m_e \mp 1$.

The correct language to describe the spontaneous emission into the ground state is as follows: for a $J_e = 1/2 \longleftrightarrow J_g = 1/2$ transition, an atom in the $m_e = +1/2$ state can fall back either into the state $m_g = -1/2$ by spontaneous emission of an electric dipole photon $M_J = +1$, or into the state $m_g = +1/2$ by spontaneous emission of an electric dipole photon $M_J = +0$. This correlation between m_g and M_J is due to conservation of total angular momentum, arising from the rotational invariance of the interaction Hamiltonian. The final state of the system after the spontaneous emission is thus an entangled state, a superposition of the $|m_g = -1/2\rangle \otimes |M_J = +1\rangle$ state and the $|m_g = +1/2\rangle \otimes |M_J = 0\rangle$ state, with coefficients weighted by the Clebsch-Gordan coefficients of the two atomic transitions, coming from the dipole matrix elements. This computation also yields the speed with which the two ground state sublevels repopulate, the branching ratio being given by the square of the Clebsch-Gordan coefficients.

Light plays a double role in these experiments. As we just saw, it polarizes the atoms by accumulating them in a ground state sublevel; it also permits the optical detection of the atoms' polarization. As the atoms can only absorb the incident σ_+ light if they are in the $m_g = -1/2$ sublevel, the absorption of that light yields a signal proportional to the population of that sublevel. Any change in the population differences

[8] The interested reader will find more details on multipolar waves properties in Complement B_I of [16] and in [34].

between the $m_g = -1/2$ and $m_g = +1/2$ sublevels, induced by a resonant radiofrequency field or by collisions, can therefore be detected by a change in the absorbed light.

2-c. Original features of these methods

We now review some of the original features of these optical methods, to understand their prominent role in the development of atomic physics.

- At the time it was suggested, the double resonance was among the first methods to extend the magnetic resonance techniques to atomic excited states. These techniques had been developed for ground states or metastables states with very long lifetimes, using essentially atomic or molecular beams: Stern-Gerlach type experimental set-ups were used to select atoms in given internal states; the flipping of the spins by a RF field would change the trajectories of the atoms or molecules in a detectable way, hence allowing, in most cases, the monitoring of the magnetic resonance (see for example reference [40]). These techniques could not be extended to the excited states because of their very short lifetimes.

- An interesting feature of these optical methods is their selectivity, both for the excitation and the detection. This selectivity comes from the light polarization, and not from the light frequency. The width of the spectral sources used at the time, and the Doppler width of the spectral lines of the atoms contained in a glass cell were considerably larger than the frequency differences between optical lines going from the ground state Zeeman sublevels to the excited states Zeeman sublevels; it was thus out of the question to try to excite or detect a single Zeeman component of the optical line.

- The measurements of the Zeeman or hyperfine structures in the excited states by the double resonance method are high resolution measurements. The structures under study are not determined by the measurement of the difference between two optical line frequencies, but by a direct measurement of the structure. In the radiofrequency or microwave domain, the Doppler width is negligible and the measurement resolution is only limited by the natural width.

- The optical methods are highly sensitive methods. A radiofrequency transition between two sublevels of the excited or ground state is not detected by the loss or gain in energy of the radiofrequency field, but via an absorbed or reemitted optical photon, which has a much higher energy than a RF photon, and whose polarization depends on which sublevel the atom is in. It is therefore possible to detect magnetic resonances in a very dilute medium, such as a vapor.

- Very high polarization ratios in the ground state, up to 90%, can be achieved by optical pumping. Such ratios are considerably larger than those expected at thermodynamic equilibrium: because of the very weak Zeeman shifts between the ground state sublevels, and the high temperature at which the experiments are conducted, the Boltzmann factors $\exp(-\hbar\omega_Z/k_BT)$ are all very close to 1. Note in addition that optical pumping may easily result, by a suitable choice of the polarization, in a larger population for the ground state Zeeman sublevel having the higher energy. This is one of the first examples of a method for achieving a population inversion, an essential condition for obtaining a maser or a laser effect.

More details on the optical methods and their applications can be found in the well documented work [24] and the references suggested therein.

3. Transferring orbital angular momentum to external atomic variables

3-a. Laguerre-Gaussian beams

In optics or atomic physics experiments, one often uses Gaussian beams, linear superpositions of plane waves with wave vectors nearly parallel to an optical axis (paraxial approximation). If all the plane waves forming the Gaussian beam have the same polarization $\boldsymbol{\varepsilon}$, the field phase in planes perpendicular to the beam axis does not depend on the azimuth angle φ that determines the direction around the beam axis. New types of Gaussian beams have recently been realized[9], called *Laguerre-Gaussian* (LG) beams, for which the field has an azimuthal dependence in $\exp im\varphi$, where $m = \pm 1, \pm 2, ...$, in planes perpendicular to the beam axis. We already mentioned the existence of such beams in § 3-c-β of Complement $\mathrm{B_{XIX}}$. We now show how the absorption of photons from such beams can transfer to the atomic center of mass a non-zero orbital angular momentum with respect to the beam axis.

3-b. Field expansion on Laguerre-Gaussian modes

The LG modes form a possible basis for expanding any field. They are characterized by three quantum numbers: the wave number k, the number of nodes p in the radial direction, and the integer number m characterizing the phase dependence on the azimuth angle φ. We assume the polarization $\boldsymbol{\varepsilon}$ to be uniform in the beam. We now place an atom in that beam, and use the LG modes basis $\boldsymbol{\varepsilon} F^m_{k,p}(\boldsymbol{r})$. Instead of being written $\langle \psi^{\mathrm{ext}}_{\mathrm{fin}} | \exp(i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}})\ |\psi^{\mathrm{ext}}_{\mathrm{in}}\rangle$, the matrix element pertaining to the external variables of the interaction Hamiltonian in this basis is now written :

$$\left\langle \psi^{\mathrm{ext}}_{\mathrm{fin}} \right|\, F^m_{k,p}(\hat{\boldsymbol{R}})\ \left|\psi^{\mathrm{ext}}_{\mathrm{in}}\right\rangle \tag{11}$$

This relation shows that the initial wave function $\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ of the atomic center of mass is now multiplied by the function $F^m_{k,p}(\boldsymbol{R})$ characterizing the mode. The phase factor $\exp(im\varphi)$ is, in a manner of speaking, "imprinted" on the initial wave function. Equation (11) means that the transition amplitude, concerning the external variables and induced by the interaction Hamiltonian, is equal to the scalar product of $F^m_{k,p}(\boldsymbol{R})\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ and the final center of mass wave function $\psi^{\mathrm{ext}}_{\mathrm{fin}}(\boldsymbol{R})$.

Imagine that the initial external state of the atom has a zero angular momentum with respect to the z axis, i.e. that $\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ does not depend on the angle φ. The absorption of a photon from such an LG beam, with quantum numbers k, p, m, gives to the product $F^m_{k,p}(\boldsymbol{R})\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ a φ dependence given by $\exp(im\varphi)$. This means that in its final state, the center of mass must have an orbital angular momentum $m\hbar$ with respect to the z axis, since $L_z = (\hbar/i)\partial/\partial\varphi$. The LG beam has transferred to the atom's center of mass an orbital angular momentum $m\hbar$. It is important to note that the transfer's efficiency, described by the matrix element (11), will only be significant if the

[9] The main method used to achieve such beams is to numerically design and fabricate holograms, then use them to diffract a Gaussian beam. For a review of the properties and applications of these new types of beams, see reference [41].

spatial extent of the initial and final wave functions are well adapted to the geometrical characteristics of the LG beam. In the vicinity of the focal point, the width of the beam is of the order of its "waist" w_0, i.e. of the order of a few microns, much larger than the atomic wave packets, of the order of nanometers at normal temperatures ($T \simeq 300K$). This explains why the orbital angular momentum transfer to the atoms' centers of mass became operational only when atoms could be cooled down to very low temperatures, in the microkelvin, or even nanokelvin range. The matter waves thus obtained, in Bose-Einstein condensates for example, can have spatial extensions of the order of a few microns. The transfer of orbital angular momentum by "phase imprint" can then be used to generate quantum vortices (see Complement D_{XV}, § 3-b-β) in matter waves, where the atoms rotate in phase around an axis. This method was actually used to create such vortices in a Bose-Einstein condensate of trapped atoms [42].

Chapter XX

Absorption, emission and scattering of photons by atoms

A A basic tool: the evolution operator 2068
A-1 General properties 2069
A-2 Interaction picture 2070
A-3 Positive and negative frequency components of the field . . . 2072
B Photon absorption between two discrete atomic levels . . . 2073
B-1 Monochromatic radiation 2073
B-2 Non-monochromatic radiation 2075
C Stimulated and spontaneous emissions 2080
C-1 Emission rate 2080
C-2 Stimulated emission 2081
C-3 Spontaneous emission 2081
C-4 Einstein coefficients and Planck's law 2083
D Role of correlation functions in one-photon processes . . . 2084
D-1 Absorption process 2084
D-2 Emission process 2085
E Photon scattering by an atom 2085
E-1 Elastic scattering 2086
E-2 Resonant scattering 2089
E-3 Inelastic scattering - Raman scattering 2091

Introduction

In this chapter, we will use the results established in the previous chapter to study some elementary processes concerning the absorption or emission of photons by atoms. Knowing the Hamiltonians describing the atomic energy levels and the radiation, as well

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

as their interactions, we can now focus on solving Schrödinger's equation that governs the evolution of the system atom plus field. Our objective is to compute the probability amplitude for that system to go from a given initial state $|\varphi_i\rangle$ at time t_i to a certain final state $|\varphi_f\rangle$ at a later time t_f.

In quantum mechanics, the evolution of the system's state vector between the instants t_i and t_f is controlled by the evolution operator $U(t_f, t_i)$, which is the basic tool for computing the amplitudes of the various processes studied in this chapter. This is why we start in § A by reviewing a number of equations satisfied by $U(t_f, t_i)$, which will be useful for the forthcoming computations.

The first processes we shall study in § B concern photon absorption or emission by an atom undergoing a transition between two discrete states. We shall first consider monochromatic incident radiation, and then a broadband excitation. We will then show in § C that two types of emission can occur during these interaction processes: stimulated emission, which is also predicted in a semiclassical treatment, and spontaneous emission, which requires a quantum treatment of the radiation. We will make a connection with the method used by Einstein to reestablish Planck's law (giving the spectral distribution of the black body radiation), and deduce the absorption and emission coefficients. The role of correlation functions (pertaining to the atomic dipole and to the incident field) in the computation of transition probabilities is discussed in § D.

An important example that involves not one, but two photons, is the scattering of a photon by an atom: during that process, an incident photon is absorbed and a new one is created either by spontaneous or induced emission. This process is studied in § E. When the frequency of the incident photon is close to the atomic transition frequency, the scattering is said to be "quasi-resonant". Its description requires a non-perturbative treatment that will be developed, based on the results of §§ 4 and 5 of Complement $\mathrm{D_{XIII}}$. In this entire chapter, we shall only consider cases where the atomic levels are discrete; a case where those levels include a continuum will be treated in Complement $\mathrm{B_{XX}}$.

> ***Notation:*** In Chapter XVIII, it was important to distinguish between the classical quantities and the corresponding quantum operators, so that the latter were denoted with "hats". In the present chapter, this distinction becomes less important, and we will come back to a more standard and simpler notation, without the hats; for instance, the annihilation and creation operators will be denoted a_i and $a_i^\dagger$, instead of $\hat{a}_i$ and $\hat{a}_i^\dagger$.

A. A basic tool: the evolution operator

The unitary evolution operator $U(t, t_0)$ has been defined in Complement $\mathrm{F_{III}}$; it yields the state of a quantum system at instant t knowing the state of that system at a previous time t_0:

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle \tag{A-1}$$

It is a unitary operator:

$$U^\dagger(t, t_0)U(t, t_0) = \mathbb{1} \tag{A-2}$$

If $H(t)$ is the system Hamiltonian, $U(t,t_0)$ obeys the differential equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{dt}}U(t,t_0) = H(t)U(t,t_0) \tag{A-3}$$

with the initial condition $U(t_0,t_0) = \mathbb{1}$. The integral equation:

$$U(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')U(t',t_0) \tag{A-4}$$

is equivalent to the differential equation together with its initial condition. If the Hamiltonian is time-independent, the evolution operator is simply:

$$U(t,t_0) = e^{-iH(t-t_0)/\hbar} \tag{A-5}$$

A-1. General properties

In this entire chapter, we use the evolution operator to express the probability amplitude $A_{fi}(t_f,t_i)$ for the system, starting from the initial state $|\varphi_i\rangle$ at instant t_i, to be found in the state $|\varphi_f\rangle$ at time t_f:

$$A_{fi}(t_f,t_i) = \langle\varphi_f|U(t_f,t_i)|\varphi_i\rangle \tag{A-6}$$

Consider the total Hamiltonian H:

$$H = H_A + H_R + H_I = H_0 + H_I \tag{A-7}$$

where $H_0 = H_A + H_R$ is the non-perturbed Hamiltonian (sum of the isolated atom Hamiltonian H_A and the free radiation Hamiltonian H_R) and H_I is the interaction Hamiltonian (between the atom and the field). The evolution operators U_0 and U associated respectively with H_0 and H read:

$$U_0(t,t_0) = e^{-iH_0(t-t_0)/\hbar} \tag{A-8a}$$

$$U(t,t_0) = e^{-iH(t-t_0)/\hbar} \tag{A-8b}$$

These operators are related through the integral relation:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U(t',t_0) \tag{A-9}$$

To demonstrate this relation, we take the derivative of each side (taking into account the derivative with respect to the integral's upper bound, which appears in addition to that of the function to be integrated):

$$i\hbar\frac{d}{dt}U(t,t_0) = H_0\; U_0(t,t_0) + U_0(t,t)H_I U(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; H_0\; U_0(t,t')\, H_I\, U(t',t_0) \tag{A-10}$$

that is, taking into account the relation $U_0(t,t) = \mathbb{1}$ and relation (A-9):

$$\begin{aligned} i\hbar\frac{d}{dt}U(t,t_0) &= H_I U(t,t_0) + H_0\left[U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U(t',t_0)\right] \\ &= [H_I + H_0]\, U(t,t_0) = HU(t,t_0) \end{aligned} \tag{A-11}$$

The operator defined in (A-9) therefore obeys relation (A-3) with the total Hamiltonian H; since, in addition, $U(t_0, t_0) = \mathbb{1}$, this operator is an evolution operator.

Inserting expression (A-9) for $U(t, t_0)$ into the integral on the right-hand side of that same expression (A-9), we obtain an expression of that operator containing a double integral in time. Reiterating this process several times, we get the series expansion of the evolution operator in powers of H_I:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U_0(t',t_0) + \\ + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\; U_0(t,t')\, H_I\, U_0(t',t'') H_I\, U_0(t'',t_0) + ... \qquad \text{(A-12)}$$

The n-order term of this series is a succession of $n+1$ non-perturbed evolutions, each described by U_0, separated by n interactions H_I.

The evolution operator also obeys another integral equation, symmetric to (A-9), which will be useful for what follows:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U(t,t')\, H_I\, U_0(t',t_0) \qquad \text{(A-13)}$$

Its demonstration is similar to that of (A-9).

Finally, we can also insert expression (A-13) for U in the integral of (A-9), replacing t by t' and t' by t''. We then obtain:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U_0(t',t_0) + \\ + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\; U_0(t,t')\, H_I\, U(t',t'') H_I\, U_0(t'',t_0) \qquad \text{(A-14)}$$

Contrary to (A-12), the right-hand side of (A-14) only has three terms, and not an infinity. It is, however, the perturbed evolution operator U, and not U_0, that appears in the last term on the right-hand side of (A-14), in between the two interaction Hamiltonians H_I. This form of the evolution operator U will be used in § E-2.

A-2. Interaction picture

For the following computations, it will often be useful to write Schrödinger's equation in the interaction picture. Let $|\psi(t)\rangle$ be the Schrödinger state vector. Setting:

$$|\bar{\psi}(t)\rangle = U_0^{\dagger}(t,t_0)|\psi(t)\rangle = \exp\left[i\,(t-t_0)\,H_0/\hbar\right]|\psi(t)\rangle \qquad \text{(A-15)}$$

the new state vector $|\bar{\psi}(t)\rangle$ obeys the time evolution equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\bar{\psi}(t)\rangle = \bar{H}_I(t)|\bar{\psi}(t)\rangle \qquad \text{(A-16)}$$

where $\bar{H}_I(t)$ is defined as:

$$\bar{H}_I(t) = U_0^{\dagger}(t,t_0) H_I U_0(t,t_0) \qquad \text{(A-17)}$$

The evolution operator $\bar{U}(t,t_0)$ in the interaction representation yields the evolution of $|\bar{\psi}(t)\rangle$:

$$|\bar{\psi}(t)\rangle = \bar{U}(t,t_0)|\bar{\psi}(t_0)\rangle \tag{A-18}$$

From equations (A-1) and (A-15), we can deduce the relation between evolution operators in the two points of view:

$$\bar{U}(t,t_0) = U_0^\dagger(t,t_0)U(t,t_0) = \exp\left[i\,(t-t_0)\,H_0/\hbar\right]U(t,t_0) \tag{A-19}$$

In addition, insertion of (A-18) into (A-16) shows that:

$$i\hbar\frac{d}{dt}\bar{U}(t,t_0) = \bar{H}_I(t)\bar{U}(t,t_0) \tag{A-20}$$

which leads to the following series expansion for $\bar{U}(t,t_0)$ (perturbation expansion):

$$\bar{U}(t,t_0) = \mathbb{1} + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\ \bar{H}_I(t') + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^{t}\mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\ \bar{H}_I(t')\bar{H}_I(t'') + ... \tag{A-21}$$

The great advantage of the interaction picture is that the state vector only evolves under the effect of the interaction – since $\bar{H}_I(t)$ is the only operator appearing on the right-hand side of (A-16). We shall see that this point of view also allows expressing the transition probabilities in terms of time correlation functions of dipole and field operators, i.e. as average values of products of physical quantities taken at two different instants, and evolving freely (under the effect of only H_0). Finally, when trying to calculate the transition probability between two eigenstates $|\varphi_i\rangle$ and $|\varphi_f\rangle$ of H_0, with respective energies E_i and E_f, it is often convenient to use the interaction picture since, according to (A-19), the transition amplitude is of the form:

$$\langle\varphi_f|\,U(t,t_0)\,|\varphi_i\rangle = \langle\varphi_f|\exp\left[-i\,(t-t_0)\,H_0/\hbar\right]\bar{U}(t,t_0)\,|\varphi_i\rangle \tag{A-22}$$

$$= e^{-iE_f(t-t_0)/\hbar}\,\langle\varphi_f|\,\bar{U}(t,t_0)\,|\varphi_i\rangle \tag{A-23}$$

As the phase factor $e^{-iE_f(t-t_0)/\hbar}$ disappears from the probability (modulus squared of the amplitude), it can be ignored; this allows simply replacing the evolution operator U by $\bar{U}$ and directly using the more compact expansion (A-21).

In this entire chapter and its complements, we describe the interaction between atom and field by the electric dipole Hamiltonian $H_I = -\boldsymbol{D}\cdot\boldsymbol{E}_\perp(\boldsymbol{R})$ introduced in § C-4 of Chapter XIX. For the sake of simplicity, we assume that $t_0 = 0$. In the interaction picture[1], this operator becomes:

$$\bar{H}_I(t) = -\bar{\boldsymbol{D}}(t)\cdot\bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t) \tag{A-24}$$

where:

$$\bar{\boldsymbol{D}}(t) = \exp\left(iH_0t/\hbar\right)\boldsymbol{D}\exp\left(-iH_0t/\hbar\right) \tag{A-25a}$$

$$\bar{\boldsymbol{E}}_\perp(t) = \exp\left(iH_0t/\hbar\right)\boldsymbol{E}_\perp\exp\left(-iH_0t/\hbar\right) \tag{A-25b}$$

[1]Here, the atom's external degrees of freedom are treated classically. The atom is supposed to be at rest at point $\boldsymbol{R}$, meaning $\boldsymbol{R}$ is not modified when going to the interaction picture.

A-3. Positive and negative frequency components of the field

In the present chapter, we assume the system is contained in a cubic box of volume L^3. The transverse electric field is given by relation (B-3) of Chapter XIX. It is a linear combination of annihilation a_i and creation $a_i^\dagger$ operators, and can be expressed as the sum of two terms:

$$\boldsymbol{E}_\perp(\boldsymbol{R}) = \boldsymbol{E}_\perp^{(+)}(\boldsymbol{R}) + \boldsymbol{E}_\perp^{(-)}(\boldsymbol{R}) \tag{A-26}$$

where:

$$\begin{aligned}
\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R}) &= i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i \boldsymbol{\varepsilon}_i e^{i\boldsymbol{k}_i\cdot\boldsymbol{R}} \\
\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R}) &= -i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i\boldsymbol{k}_i\cdot\boldsymbol{R}} = \left[\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R})\right]^\dagger
\end{aligned} \tag{A-27}$$

Operator $\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R})$, obtained by keeping only the annihilation operators a_i in the expansion of $\bar{\boldsymbol{E}}_\perp(\boldsymbol{R})$, is called [2] the electric field "*positive frequency component*". As for the operator $\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R})$, it is the "*negative frequency component*". These two operators are not Hermitian, and do not commute. In a product of field operators, the order is said to be *normal* if the creation operators are to the left of the annihilation operators, as in $\boldsymbol{E}_\perp^{(-)}\boldsymbol{E}_\perp^{(+)}$; the order is said to be *antinormal* for a product in the inverse order, as in $\boldsymbol{E}_\perp^{(+)}\boldsymbol{E}_\perp^{(-)}$.

In the interaction picture, the annihilation and creation operators become:

$$\begin{aligned}
\bar{a}_i(t) &= \exp\left(iH_0 t/\hbar\right) a_i \exp\left(-iH_0 t/\hbar\right) = a_i e^{-i\omega_i t} \\
\bar{a}_i^\dagger(t) &= \exp\left(iH_0 t/\hbar\right) a_i^\dagger \exp\left(-iH_0 t/\hbar\right) = a_i^\dagger e^{+i\omega_i t}
\end{aligned} \tag{A-28}$$

(these equalities can be verified by computing the matrix elements in the Fock state basis, eigenvectors of H_0, and using the fact that the only action of operator a_i is to annihilate a photon in the mode i). The positive and negative frequency components of the field are thus:

$$\begin{aligned}
\bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t) &= i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{R}-\omega_i t)} \\
\bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R},t) &= -i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{R}-\omega_i t)} = \left[\bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t)\right]^\dagger
\end{aligned} \tag{A-29}$$

Suppose now that we wish to study the lowest order process of photon absorption by atoms. To compute the action of $\bar{H}_I(t)$ on the system's initial state, we can keep

[2] The positive frequency component annihilates a photon, the negative frequency component creates one. Furthermore, in the Heisenberg picture, we shall see that the free evolution of the positive component goes as $e^{-i\omega t}$, and that of the negative one, as $e^{+i\omega t}$. This labeling as "positive frequency" may seem somewhat counter intuitive, but is widely accepted.

only the terms in $\bar{H}_I(t)$ that annihilate one photon, i.e. the Hamiltonian terms containing $\bar{a}_i(t)$. This amounts to keeping only the positive frequency component of the field operator, and using the simplified interaction Hamiltonian:

$$\bar{H}_I(t) \simeq -\bar{\boldsymbol{D}}(t) \cdot \boldsymbol{E}_\perp^{(+)}(\boldsymbol{R}, t) \qquad \text{(absorption)} \tag{A-30}$$

In a similar way, to study the lowest order photon emission process, we can use the expression:

$$\bar{H}_I(t) \simeq -\bar{\boldsymbol{D}}(t) \cdot \boldsymbol{E}_\perp^{(-)}(\boldsymbol{R}, t) \qquad \text{(emission)} \tag{A-31}$$

B. Photon absorption between two discrete atomic levels

We start with monochromatic radiation (§ B-1), and will study later the broadband radiation case (§ B-2). The base of the computation is the study of the transition rate between stationary states of the non-perturbed Hamiltonian H_0; Complement $\mathrm{D_{XX}}$ will present a more detailed study in terms of wave packets propagating in free space, built from coherent superpositions of stationary states.

B-1. Monochromatic radiation

B-1-a. Probability amplitude (absorption)

We call a and b two discrete levels with respective energies E_a and E_b, and set:

$$E_b - E_a = \hbar\omega_0 \tag{B-1}$$

where $\omega_0/2\pi$ is the atomic eigenfrequency, assumed to be positive (the b level has an energy higher than the a level). For the sake of simplicity, we shall ignore the external variables, which amounts to considering the atom as infinitely heavy and at rest[3].

We assume the radiation is at the initial time $t_i = 0$ in a state $|\,\phi_{\text{in}}^R\rangle = |n_i\rangle$ containing n_i photons i with wave vector $\boldsymbol{k}_i$, polarization $\boldsymbol{\varepsilon}_i$ and frequency $\omega_i/2\pi$; it is a monochromatic radiation. The initial state of the system atom + radiation is written:

$$|\psi_{\text{in}}\rangle = |a; n_i\rangle \qquad \text{with energy } E_{\text{in}} = E_a + n_i\hbar\omega_i \tag{B-2}$$

We are trying to compute the probability amplitude for the atom to absorb a photon and be in the excited state b at instant $t_f = \Delta t$. The final state of the system must then be:

$$|\psi_{\text{fin}}\rangle = |b; n_i - 1\rangle \qquad \text{with energy } E_{\text{fin}} = E_b + (n_i - 1)\hbar\omega_i. \tag{B-3}$$

As mentioned above, when $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{fin}}\rangle$ are eigenstates of H_0, it is easier to carry out the calculation in the interaction representation. In the expansion (A-21) of $\bar{U}$, the lowest order term that can link those two states is the first order term in H_I. Calling

[3] Complement $\mathrm{A_{XIX}}$ shows how taking into account the external variables and momentum conservation allows introducing the Doppler effect and the recoil effect in the absorption and emission of a photon. It also shows how the confinement of the atom in a region of space by a trapping potential allows controlling those effects.

Δt the duration of the interaction, the probability amplitude for the system, initially in the state $|\psi_{\rm in}\rangle$ at time $t_0 = 0$, to be at time $t = \Delta t$ in the final state $|\psi_{\rm fin}\rangle$ is:

$$\begin{aligned}\langle\psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle &= \frac{1}{i\hbar}\int_0^{\Delta t} {\rm d}t'\ \langle\psi_{\rm fin}|e^{iH_0t'/\hbar}H_I e^{-iH_0t'/\hbar}|\psi_{\rm in}\rangle \\ &= \frac{1}{i\hbar}\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\int_0^{\Delta t}{\rm d}t'\ e^{i(E_{\rm fin}-E_{\rm in})t'/\hbar}\end{aligned} \tag{B-4}$$

or else, taking into account the relation $E_{\rm fin} - E_{\rm in} = \hbar(\omega_0 - \omega_i)$:

$$\begin{aligned}\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle &= \frac{1}{i\hbar}\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\int_0^{\Delta t}{\rm d}t'\ e^{i(\omega_0-\omega_i)t'} \\ &= -\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\frac{1}{\hbar\,(\omega_0-\omega_i)}\left[e^{i(\omega_0-\omega_i)\Delta t}-1\right]\end{aligned} \tag{B-5}$$

which leads to:

$$\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle = -\frac{i}{\hbar}\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\ e^{i(\omega_0-\omega_i)\Delta t/2}\left[\frac{2\sin\left(\omega_0-\omega_i\right)\Delta t/2}{\left(\omega_0-\omega_i\right)}\right] \tag{B-6}$$

The absorption probability is the squared modulus of that expression:

$$\left|\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle\right|^2 = \frac{1}{\hbar^2}\left|\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\right|^2\frac{4\sin^2\left(\omega_0-\omega_i\right)\Delta t/2}{\left(\omega_0-\omega_i\right)^2} \tag{B-7}$$

Taking relations (A-24) and (A-27) into account, the matrix element of H_I, appearing in the amplitude (B-5), can be written:

$$\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle = -i\sqrt{\frac{\hbar\omega_i}{2\epsilon_0 L^3}}\sqrt{n_i}\ \langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle e^{i\boldsymbol{k}_i\cdot\boldsymbol{R}} \tag{B-8}$$

so that the probability becomes:

$$\left|\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle\right|^2 = \frac{\omega_i}{2\hbar\epsilon_0 L^3}n_i\left|\langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle\right|^2\frac{4\sin^2\left(\omega_0-\omega_i\right)\Delta t/2}{\left(\omega_0-\omega_i\right)^2} \tag{B-9}$$

It is proportional to the number of incident photons n_i, i.e. to the incident intensity in the state $|\psi_{\rm in}\rangle$, as well as to the squared modulus of the atomic dipole matrix element between the states a and b (since $\boldsymbol{D}$ is an odd operator, the absorption of the photon can only occur between two states of different parity). This probability is an oscillating function of the time Δt.

B-1-b. Energy conservation

The presence in the denominator of (B-9) of the factor $(\omega_i - \omega_0)^2$ means that, the closer ω_i gets to ω_0, the larger the absorption probability will be. A photon absorption is said to be resonant when the absorbed photon energy is exactly equal to the energy the atom must gain to go from a to b (energy conservation). The width of the resonance described by (B-9) is of the order of $\Delta\omega_i = 1/\Delta t$ or, in energy, of the order of $\Delta E = \hbar/\Delta t$. This is consistent with the time-energy uncertainty principle: in a process extending over a length time Δt, the energy is only conserved to within $\hbar/\Delta t$.

Actually, when the variable is the angular frequency ω_i, one may consider the function within brackets in (B-6) to be an approximate delta function $\delta^{(\Delta t)}(\omega_0 - \omega_i)$, having a non-zero width of the order of $1/\Delta t$. As relation (10) of Appendix II in Volume II shows that the integral over ω_i of $\sin[(\omega_0 - \omega_i)\Delta t/2]/(\omega_0 - \omega_i)$ is equal to π, and since $(\omega_0 - \omega_i)$ is proportional to the difference in total energy between the final and initial states, equation (B-6) can be written (ignoring the phase factor):

$$\langle \psi_{\text{fin}} | \bar{U}(\Delta t, 0) | \psi_{\text{fin}} \rangle \simeq -2i\pi \ \langle \psi_{\text{fin}} | H_I | \psi_{\text{in}} \rangle \ \delta^{(\Delta t)}(E_{\text{fin}} - E_{\text{in}}) \tag{B-10}$$

B-1-c. Limits of the perturbative treatment

The lowest order perturbative treatment in H_I that we have used cannot remain valid for arbitrarily long times. To understand why, imagine for example that the resonance condition $\omega_i = \omega_0$ is satisfied. Since $\sin(x\Delta t)/x$ tends toward Δt when x goes to 0, the absorption probability predicted by (B-9) becomes proportional to Δt^2, which gets very large as the time interval Δt increases. Now a probability can never be larger than one; it is therefore obvious that expression (B-9) is no longer valid for long times. The same is true if ω_i gets close to ω_0 without being strictly equal to it: it is then quite possible for the oscillation amplitude of the right-hand side of (B-9) to be larger than one. The previous results can thus only be used for short enough times, ensuring the validity of the perturbative treatment.

We shall use in Complement C_{XX} the "dressed atom method" to get a more precise treatment of the coupling effects between a two-level atom and a single mode of the field. The coupling intensity will be characterized by a constant Ω_1 called the "Rabi frequency". At resonance, one shows that the probability amplitude presents a "Rabi oscillation"[4] in $\sin \Omega_1 \Delta t$. The quadratic behavior in Δt^2 found above for the absorption probability at resonance is simply the first term of the series expansion of $\sin^2(\Omega_1 \Delta t)$ in powers of $\Omega_1 \Delta t$. We shall also discuss the extent to which relation (B-9) can be used far from resonance.

B-2. Non-monochromatic radiation

We now study the absorption and emission processes when the radiation is no longer monochromatic. The transitions are still between two discrete atomic levels a and b; the case of a continuum of atomic levels is discussed in Complement B_{XX}.

B-2-a. Absorption of a broadband radiation

We now assume the initial state of the system atom + radiation to be of the form:

$$|\psi_{\text{in}}\rangle = |a; \phi^R_{\text{in}}\rangle \tag{B-11}$$

The atom is still in the internal state of lower energy a, but the radiation is now in a state $|\phi^R_{\text{in}}\rangle$ where photons occupy several modes with different frequencies. The radiation frequency distribution is characterized by a certain spectral band $\Delta\omega$. We shall see below (very last part of § B-2-a-γ) the condition $\Delta\omega$ must satisfy for the results obtained in

[4] We shall also show how this Rabi oscillation is modified when taking into account the width of the excited level due to spontaneous emission.

this section to be valid. The calculations will be carried out in the same way as in § B-1, while focusing more on the time correlation properties of the incident field.

We introduce the notation:

$$\langle b\,|\boldsymbol{D}|a\rangle = \boldsymbol{\mathcal{D}} \qquad \text{with:} \qquad \boldsymbol{\mathcal{D}} = \mathcal{D}\mathbf{e}_d \tag{B-12}$$

where $\mathbf{e}_d$ is the unit vector[5] (a priori complex) parallel to $\boldsymbol{\mathcal{D}}$, and $\mathcal{D}$ the (real) modulus of that vector. In the interaction picture (with respect to the free atom Hamiltonian), this equality becomes:

$$\langle b\,|\bar{\boldsymbol{D}}(t)|a\rangle = \langle b\,|e^{iH_0t}\boldsymbol{D}e^{-iH_0t}|a\rangle = \boldsymbol{\mathcal{D}}e^{i\omega_0 t} \tag{B-13}$$

α. Transition probability

Let us insert expression (A-24) for $\bar{H}_I(t)$ into (A-21). The first order term in $\bar{H}_I(t)$ yields the probability amplitude for the system, starting at $t_i = 0$ from the state $|\,a;\phi^R_{\text{in}}\rangle$, to be found at time $t_f = \Delta t$ in the state $|b;\phi^R_{\text{fin}}\rangle$. Taking (B-13) into account, we obtain:

$$\begin{aligned}\langle\psi_{\text{fin}}|\bar{U}(t_f,t_i)|\psi_{\text{in}}\rangle &= -\frac{1}{i\hbar}\int_0^{\Delta t} \mathrm{d}t'\ \langle b|\ \bar{\boldsymbol{D}}(t')\,|a\rangle\cdot\left\langle\phi^R_{\text{fin}}\right|\bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t')\left|\phi^R_{\text{in}}\right\rangle \\ &= -\frac{\mathcal{D}}{i\hbar}\int_0^{\Delta t} \mathrm{d}t'\ e^{i\omega_0 t'}\left\langle\phi^R_{\text{fin}}\right| E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t')\left|\phi^R_{\text{in}}\right\rangle\end{aligned} \tag{B-14}$$

where we have included only the positive frequency component of the field, which is the only one involved since $\left|\phi^R_{\text{fin}}\right\rangle$ contains less photons than $\left|\phi^R_{\text{in}}\right\rangle$ (we are dealing with an atomic absorption process); in this equality and the following, we shall use the convenient notation:

$$E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t) = \mathbf{e}_d\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t) \quad ; \quad E^{(-)}_{\mathbf{d}^*}(\boldsymbol{R},t) = \mathbf{e}^*_d\cdot\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t) \tag{B-15}$$

where $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t)$ and $\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t)$ are the field operators in the interaction picture defined in (A-29); the atom is supposed to be fixed at point $\boldsymbol{R}$. The corresponding transition probability $P^{\text{abs}}_{a\to b}(\Delta t)$ is obtained by squaring the modulus of amplitude (B-14), and then summing over all possible radiation final states. Replacing in (B-14) the integral variable t' by t'', we obtain:

$$P^{\text{abs}}_{a\to b}(\Delta t) = \frac{\mathcal{D}^2}{\hbar^2}\int_0^{\Delta t}\mathrm{d}t'\int_0^{\Delta t}\mathrm{d}t''\ e^{i\omega_0(t''-t')}G^N_R(t',t'') \tag{B-16}$$

where $G^N_R(t',t'')$ is defined as:

$$\begin{aligned}G^N_R(t',t'') &= \sum_{\left|\phi^R_{\text{fin}}\right\rangle}\langle\phi^R_{\text{in}}|E^{(-)}_{\mathbf{d}^*}(\boldsymbol{R},t')\left|\phi^R_{\text{fin}}\right\rangle\left\langle\phi^R_{\text{fin}}\right| E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t'')|\phi^R_{\text{in}}\rangle \\ &= \langle\phi^R_{\text{in}}|E^{(-)}_{\mathbf{d}^*}(\boldsymbol{R},t')E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t'')|\phi^R_{\text{in}}\rangle\end{aligned} \tag{B-17}$$

[5]The matrix elements $\mathcal{D}_{x,y,z}$ of the three components of $\boldsymbol{\mathcal{D}}$ are three a priori complex numbers. We set $\mathcal{D} = \sqrt{|\mathcal{D}_x|^2+|\mathcal{D}_y|^2+|\mathcal{D}_z|^2}$ and introduce the vector $\mathbf{e}_d = \boldsymbol{\mathcal{D}}/\mathcal{D}$. It is a unit vector since $\mathbf{e}^*_d\cdot\mathbf{e}_d = 1$; we have $\mathcal{D} = \boldsymbol{e}^*_d\cdot\boldsymbol{\mathcal{D}} = \langle b|\boldsymbol{e}^*_d\cdot\boldsymbol{D}|a\rangle$.

This function characterizes the role of the incident beam in the absorption process under study. It is the average value in the initial state of the product of two field operators, arranged in normal order (§ A-3), and taken at two different times.

Changing the integral variable t'' for $\tau = t'' - t'$ (the jacobian of this change of variables is equal to unity), equation (B-16) becomes:

$$P_{a\to b}^{\text{abs}}(\Delta t) = \frac{\mathcal{D}^2}{\hbar^2} \int_0^{\Delta t} \mathrm{d}t' \int_{-t'}^{\Delta t - t'} \mathrm{d}\tau \; e^{i\omega_0 \tau} G_R^N(t', t'+\tau) \tag{B-18}$$

The transition probability is therefore proportional to the time integral of the Fourier transform[6] with respect to τ of the function $G_R^N(t', t'+\tau)$, limited to the time interval $[-t', \Delta t - t']$.

β. Excitation spectrum

If the radiation initial state $\left|\phi_{\text{in}}^R\right\rangle$ is an eigenstate of the free radiation, the function $G_R^N(t', t'')$ depends only on the difference $t' - t''$. For example, if $\left|\phi_{\text{in}}^R\right\rangle$ is a Fock state[7]:

$$\left|\phi_{\text{in}}^R\right\rangle = |n_1, ..., n_i, ...\rangle \tag{B-19}$$

where the populated modes have the polarization $\boldsymbol{\varepsilon}_i$, inserting expansions (A-29) for the field operators leads to:

$$\begin{aligned} G_R^N(t'-t'') &= \sum_i \frac{\hbar\omega_i}{2\varepsilon_0 L^3} \underbrace{\left\langle \phi_{\text{in}}^R \left| a_i^\dagger a_i \right| \phi_{\text{in}}^R \right\rangle}_{=\langle n_i \rangle} \left| \mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i \right|^2 \; e^{i\omega_i (t'-t'')} \\ &= \int \mathrm{d}\omega \; I_R^N(\omega) e^{i\omega(t'-t'')} \end{aligned} \tag{B-20}$$

where:

$$I_R^N(\omega) = \sum_i \frac{\langle n_i \rangle \hbar\omega_i}{2\varepsilon_0 L^3} \; \left| \mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i \right|^2 \delta(\omega_i - \omega) \tag{B-21}$$

If the radiation is initially in the Fock state (B-19), the value of $\langle n_i \rangle$ is simply n_i, the number of photons in that state. On the other hand, if the radiation is in a statistical mixture of such states (see note 7), $\langle n_i \rangle$ represents the corresponding statistical average.

Expression (B-20) thus appears as the value at $t' - t''$ of the Fourier transform of a function $I_R^N(\omega)$ of ω, which depends on the initial photon populations $\langle n_i \rangle$. As $\langle n_i \rangle \hbar\omega_i$ is the average energy of mode i with frequency $\omega_i/2\pi$, the function $I_R^N(\omega)$ actually gives the variation of the energy density of the radiation as a function of frequency; this function is also referred to as the spectral distribution of the incident radiation (excitation spectrum). This distribution can have, a priori, any shape, but the energy density often presents a single peak of width $\Delta\omega$, with no other particular structure. Its Fourier transform $G_R^N(t' - t'')$ is then a function of width $\simeq 1/\Delta\omega$; as a result, when $|t' - t''|$

[6] The components of the field $E_{\mathbf{d}}^{(+)}$ vary in $\mathrm{e}^{-i\omega t}$. This is why the Fourier component of G_R^N at $-\omega_0$ appears in (B-18).

[7] Instead of a Fock state as the initial state, one could choose a statistical mixture of such states with arbitrary weights; this would not significantly change the following calculations and conclusions.

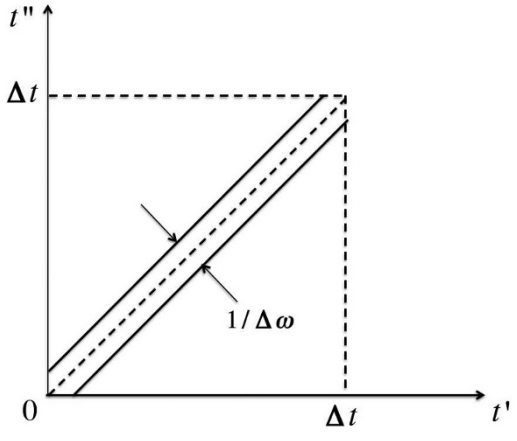


Figure 1: The function to be integrated in expression (B-16) yielding the absorption probability is significant only in a band along the first bisector, of width $1/\Delta\omega$, where $\Delta\omega$ is the width of the incident radiation spectrum. If $\Delta t \gg 1/\Delta\omega$, the area of the domain actually contributing to that integral increases linearly with the diagonal of the square, and hence with Δt (and not with its square).

becomes large compared to $1/\Delta\omega$, the correlation function $G_R^N(t'-t'')$ goes to zero and can be neglected. We shall see that in such a case the probability becomes proportional to Δt, which naturally leads to introducing a probability per unit time. Our reasoning will be similar to that of Complement E_{XIII} for a classical perturbation, but here the radiation is treated quantum mechanically.

γ. Probability per unit time

The integration domain of the double integral in (B-16) is plotted in Figure 1. As $G_R^N(t'-t'')$ goes to zero as soon as $|t'-t''|$ is large compared to $1/\Delta\omega$, the portion of that domain where the function to be integrated is not negligible is a band of width $1/\Delta\omega$, along the first bisector; the width of this band is very small compared to the domain extension if $\Delta t \gg 1/\Delta\omega$. To make use of that property, we again replace the integral variable t'' by $\tau = t'' - t'$ (the associated Jacobian for this change of variables in the double integral is equal to 1):

$$\int_0^{\Delta t} \mathrm{d}t' \int_0^{\Delta t} \mathrm{d}t'' = \int_0^{\Delta t} \mathrm{d}t' \int_{-t'}^{\Delta t - t'} \mathrm{d}\tau \tag{B-22}$$

In the second integral, the values of τ that actually yield a non negligible contribution are of the order of the correlation time $1/\Delta\omega$ of $G_R^N(t'-t'')$. If we assume $\Delta t \gg 1/\Delta\omega$, we can replace the limits of that second integral by $-\infty$ and $+\infty$. Inserting then (B-20) into (B-18) yields the integral:

$$\int_0^{\Delta t} \mathrm{d}t' \int_{-\infty}^{+\infty} \mathrm{d}\tau \int_{-\infty}^{+\infty} \mathrm{d}\omega \; e^{i(\omega_0-\omega)\tau} I_R^N(\omega) = 2\pi \, \Delta t \, I_R^N(\omega_0) \tag{B-23}$$

since the summation over $\mathrm{d}\tau$ yields the function $2\pi\delta(\omega_0-\omega)$, which allows integration over $\mathrm{d}\omega$; we get a function independent of t' whose integration is proportional to Δt.

We finally obtain:

$$P_{a\to b}^{\mathrm{abs}}(\Delta t)=\frac{2\pi}{\hbar^2}\ \Delta t\ \mathcal{D}^2\ I_R^N(\omega_0)=\frac{2\pi}{\hbar^2}\ \Delta t\ \mathcal{D}^2\sum_i\frac{\langle n_i\rangle\hbar\omega_i}{2\varepsilon_0 L^3}\ \left|\mathbf{e}_d\cdot\boldsymbol{\varepsilon}_i\right|^2\delta(\omega_i-\omega_0) \tag{B-24}$$

This means that $P_{a\to b}^{\mathrm{abs}}(\Delta t)$ increases **linearly** with Δt, which leads us to define an absorption probability per unit time (absorption rate):

$$W_{a\to b}^{\mathrm{abs}}=\frac{P_{a\to b}^{\mathrm{abs}}(\Delta t)}{\Delta t}=2\pi\frac{\mathcal{D}^2}{\hbar^2}I_R^N(\omega_0) \tag{B-25}$$

(remember that, for monochromatic radiation, we found an absorption probability increasing not linearly but as the square of Δt). This absorption rate is proportional to the radiation energy density at the atomic transition frequency ω_0. Formulas (B-21) and (B-25) give the dependence of the absorption rate on the various parameters of the incident radiation (populations of the modes, polarization $\boldsymbol{\varepsilon}_i$).

Our calculation is perturbative since it was carried out to the lowest order in H_I. It is therefore only valid for times Δt such that $P_{a\to b}^{\mathrm{abs}}(\Delta t)=W_{a\to b}^{\mathrm{abs}}\Delta t\ll 1$, i.e. such that $\Delta t\ll 1/W_{a\to b}^{\mathrm{abs}}$. In addition, we saw above that the linear variation of $P_{a\to b}^{\mathrm{abs}}(\Delta t)$ with Δt is obtained only if $\Delta t\gg 1/\Delta\omega$. These two inequalities are compatible if $\Delta\omega\gg W_{a\to b}^{\mathrm{abs}}$. The approximation to lowest order is thus valid only if the broad band radiation has a spectral width large compared to the absorption rate.

B-2-b. A specific case: isotropic radiation

The previous calculations can be pushed a step further when the radiation is isotropic, meaning when $\langle n_i\rangle$, the average number of photons in mode i, depends only on ω_i, and neither on the direction of the wave vector $\boldsymbol{k}_i$ nor on the polarization $\boldsymbol{\varepsilon}_i$. The results we shall obtain for this specific case will be useful for later computations (§ C-4) on the spontaneous emission rate, as well as for the isotropic radiation at thermal equilibrium.

Consider the limit of (B-25) when the volume L^3 containing the system goes to infinity. The summation over the index i can be replaced by an integral:

$$\sum_i\Rightarrow\frac{L^3}{(2\pi)^3}\int k^2\ \mathrm{d}k\ \mathrm{d}\Omega\sum_{\boldsymbol{\varepsilon}\perp\boldsymbol{k}} \tag{B-26}$$

where, for isotropic radiation, a sum is taken over two linear polarizations perpendicular to à $\boldsymbol{k}$. We assume that the vector $\mathbf{e}_d$ is real as well, and choose an Oz axis that is parallel to it[8]. We first calculate, for a given direction of $\boldsymbol{k}$, the sum of the quantities ε_{iz}^2 for the two polarizations. We take two polarization vectors, $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$, both perpendicular to $\boldsymbol{k}$, and perpendicular to each other. Imagine the first one is in the plane containing $\boldsymbol{k}$ and the z axis, so that $\varepsilon_{1z}=\sin\theta$, where θ is the angle between $\boldsymbol{k}$ and the z axis; the

[8] If $\mathbf{e}_d$ is complex, it is easy to see that the contributions of its real and imaginary part simply add in (B-24). The sum of these two contributions then gives (B-27) again.

second is necessarily perpendicular to that plane, which means $\varepsilon_{2z} = 0$. We then have $\sum_{\boldsymbol{\varepsilon}\perp\boldsymbol{k}} \varepsilon_{iz}^2 = \sin^2\theta$. The integral over all the directions of $\boldsymbol{k}$ is then trivial:

$$\int \mathrm{d}\Omega\ \sin^2\theta = 2\pi \int_0^\pi \mathrm{d}\theta\ \sin^3\theta = \frac{8\pi}{3} \tag{B-27}$$

We are left with the integral of the modulus k of $\boldsymbol{k}$. Using (B-25), (B-26), (B-27) and changing the variable k into the variable $\omega = ck$, we finally obtain:

$$W_{a\to b}^{\mathrm{abs}} = \frac{\mathcal{D}^2\omega_0^3}{3\varepsilon_0\hbar\pi c^3}\langle n(\omega_0)\rangle \tag{B-28}$$

C. Stimulated and spontaneous emissions

C-1. Emission rate

We now assume that the atom is initially in the upper state b, and that the radiation is in the initial Fock state (B-19); we study the emission processes where the atom falls back into the state a while emitting a photon. We still assume that the spectrum of the incident radiation is broad, so that its correlation function tends to zero in a time that is much shorter than Δt. The computations are then similar to the one we just did for the absorption process, but with a certain number of changes. First of all, we must now use expression (A-31) for the interaction Hamiltonian, the one that contains the negative frequency component of the field operator (with solely creation operators). Secondly, concerning the electric dipole operator, we must only keep the term connecting b to a, which amounts to replacing (B-13) by the complex conjugate expression:

$$\langle a\,|\tilde{\boldsymbol{D}}(t)|b\rangle = \boldsymbol{\mathcal{D}}^* e^{-i\omega_0 t} = \mathcal{D}\ \mathbf{e}_d^*\ e^{-i\omega_0 t} \tag{C-1}$$

The correlation function (B-17) of the field in normal order is now replaced by the correlation function in antinormal order $G_R^A(t', t'')$:

$$G_R^A(t', t'') = \left\langle \phi_{\mathrm{in}}^R \right|\ E_{\mathbf{d}}^{(+)}(\boldsymbol{R}, t')E_{\mathbf{d}^*}^{(-)}(\boldsymbol{R}, t'')\ \left|\phi_{\mathrm{in}}^R\right\rangle \tag{C-2}$$

For the radiation state (B-19), this function is given by:

$$\begin{aligned} G_R^A(t'-t'') &= \sum_i \frac{\hbar\omega_i}{2\varepsilon_0 L^3}\langle n_i\,|a_i a_i^\dagger|\,n_i\rangle\ |\mathbf{e}_d\cdot\boldsymbol{\varepsilon}_i|^2\ e^{-i\omega_i(t'-t'')} \\ &= \sum_i \frac{(\langle n_i\rangle+1)\hbar\omega_i}{2\varepsilon_0 L^3}\ |\mathbf{e}_d\cdot\boldsymbol{\varepsilon}_i|^2\ e^{-i\omega_i(t'-t'')} \end{aligned} \tag{C-3}$$

In the present case, it is now $a_i a_i^\dagger$ that is present in (C-3), and not $a_i^\dagger a_i$ as in (B-20). Since a_i and $a_i^\dagger$ do not commute, this leads to an important difference compared with expression (B-20) for $G_R^N(t'-t'')$: the number of photons $\langle n_i\rangle$ initially populating the mode i is replaced by $\langle n_i\rangle + 1$. It is the quantum character of the field (non-commuting operators) that is responsible for these essential differences between the absorption and emission processes.

The emission rate is obtained by computations similar to those that led to the absorption rate (B-25) (with a change of sign for ω_0 and for the ω_i). We obtain:

$$W_{b\to a}^{\mathrm{em}} = \frac{P_{b\to a}^{\mathrm{em}}(\Delta t)}{\Delta t} = 2\pi \frac{\mathcal{D}^2}{\hbar^2} \sum_i \frac{[\langle n_i\rangle + 1]\,\hbar\omega_i}{2\varepsilon_0 L^3} \left|\mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i\right|^2 \delta(\omega_i - \omega_0) \tag{C-4}$$

This result differs from the absorption rate only by the replacement of $\langle n_i\rangle$ by $\langle n_i\rangle + 1$. This formula gives the general expression of the emission rate as a function of the population $\langle n_i\rangle$ of the modes and of their polarization $\boldsymbol{\varepsilon}_i$. We shall now discuss its two components, one proportional to $\langle n_i\rangle$, and one that does not depend on $\langle n_i\rangle$.

C-2. Stimulated emission

We first consider the terms in (C-4) that contain $\langle n_i\rangle$, i.e. the contribution of the modes containing initially at least one photon. These terms correspond to an emission induced by the incident radiation; their rate is proportional to the incident light intensity. For this reason it is called *stimulated* emission. Its rate is the same as the absorption rate, since the terms depending on $\langle n_i\rangle$ are identical in (B-24) and (C-4):

$$W_{b\to a}^{\mathrm{stim.em}} = W_{a\to b}^{\mathrm{abs}} \tag{C-5}$$

In particular, for isotropic radiation, we get a result identical to (B-28):

$$W_{b\to a}^{\mathrm{stim.em}} = \frac{\mathcal{D}^2\omega_0^3}{3\varepsilon_0\hbar\pi c^3}\langle n(\omega_0)\rangle \tag{C-6}$$

A photon resulting from stimulated emission is created in the same mode i as the photons inducing that emission: the number n_i of photons in that mode goes from n_i to $n_i + 1$. The added photon has the same energy, same direction and same polarization as the initial n_i photons. If the incident radiation is coherent, one can show that radiation emitted by stimulated emission has the same phase as the incident one. This results in a constructive interference effect (in the direction of the incident radiation) between the radiation emitted by the induced dipole and the incident radiation, hence leading to an amplification effect. For this phenomenon to occur, the atomic populations must be inverted, meaning that the probability of occupation of the upper level b must be larger than that of the lower level a. However, if this is not the case, the interference is destructive, which explains the attenuation of an incident beam by the absorption process. The coherent amplification by stimulated emission of an incident beam propagating through atoms with an inverted population plays an essential role in laser systems. The word laser is an acronym of "Light Amplification by Stimulated Emission of Radiation".

C-3. Spontaneous emission

If all the $\langle n_i\rangle$ are equal to zero, the radiation is initially in the vacuum state; the absorption rate (B-25) is then equal to zero. On the other hand, the emission rate (C-4) is not, because of the term 1 in $\langle n_i\,|a_i a_i^\dagger|\,n_i\rangle = \langle n_i\rangle + 1$. It follows that an atom, initially in the upper state b and placed in a vacuum of photons, has a non-zero probability per unit time of emitting a photon and falling back into the lower state a. This is called the *spontaneous emission* process.

The term 1 appearing in (C-4) is the same for all the modes, as opposed to the term $\langle n_i \rangle$ that only exists for the modes already containing photons. As this term 1 does not favor any particular direction or polarization, the situation is similar to that where the radiation is isotropic; the calculations leading to (B-28) remain valid provided we replace $\langle n(\omega_0) \rangle$ by 1. This leads to the rate of spontaneous emission:

$$W_{b\to a}^{\text{spont.em.}} = \frac{\mathcal{D}^2 \omega_0^3}{3\varepsilon_0 \hbar \pi c^3} \tag{C-7}$$

This result could also be obtained by Fermi's golden rule (see Chapter XIII, § C-3-b) in the following way. The system atom + radiation starts from the initial state $|b;0\rangle$ (atom in state b, photon vacuum). This discrete state is coupled to an infinity of final states $|a;1_i\rangle$, atom in state a with one photon in mode i. The golden rule enables one to compute the probability per unit time of the transition from the discrete initial state towards the continuum of final states, which is simply the spontaneous emission rate:

$$W_{b\to a}^{\text{em.spont.}} = \frac{2\pi}{\hbar} \sum_i |\langle a;1_i|H_I|b;0\rangle|^2 \, \delta(\hbar\omega_i - \hbar\omega_0) \tag{C-8}$$

Using $\langle b;0|H_I|a;1_i\rangle = i\mathcal{D}(\mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i)(\hbar\omega_i/2\varepsilon_0 L^3)^{1/2}$, for a real polarization, as well as equations (B-26) and (B-27), we get the same result as (C-7).

This equation shows that the spontaneous emission rate increases as the cube of the atomic transition frequency; this explains why spontaneous emission is negligible in the radiofrequency domain, becomes important in the optical domain, and even more so in the ultraviolet or X ray domain. This ω_0^3 factor has two origins: on one hand, the square of the $\omega^{1/2}$ factor appearing in the electric field expression, on the other, the ω^2 factor present in the density of final states.

The spontaneous emission rate $W_{b\to a}^{\text{spont.em.}}$ is also called the "*natural width*" of the excited state b, and noted Γ. The inverse of Γ is the "*radiative lifetime*" of the excited state, which is the average time necessary for the atom to undergo radiative decay:

$$\Gamma = W_{b\to a}^{\text{spont.em.}} = \frac{1}{\tau_R} \tag{C-9}$$

It is instructive to compare the rate $W_{b\to a}^{\text{spont.em.}}$ to ω_0. It follows from (C-7) that:

$$\frac{\Gamma}{\omega_0} = \frac{\mathcal{D}^2 \omega_0^2}{3\varepsilon_0 \hbar \pi c^3} \tag{C-10}$$

The dipole $\mathcal{D}$ is of the order of qa_0 where q is the electron charge and a_0 the Bohr radius. This yields the quantity $q^2/(3\pi\varepsilon_0\hbar c)$ which, within a factor 4/3, is the fine-structure constant $\alpha \simeq 1/137$ multiplied by $a_0^2\omega_0^2/c^2$. Since $a_0\omega_0$ is of the order of the electron velocity in the first Bohr orbit, which is α times smaller than the speed of light c, $a_0^2\omega_0^2/c^2$ is of the order of α^2, so we finally have:

$$\frac{\Gamma}{\omega_0} \simeq \alpha^3 \tag{C-11}$$

The natural width of the excited level is therefore much smaller than the atomic transition frequency: the atomic dipole may oscillate a great number of times before these oscillations are damped. Typically, in the optical domain, Γ is of the order of 10^7 to 10^9 s^{-1} whereas $\omega_0/2\pi$ is much larger, of the order of 10^{14} to 10^{15} s^{-1}.

C-4. Einstein coefficients and Planck's law

Let us now assume the radiation is at thermal equilibrium at temperature T (black body radiation). In this case, it is current practice to use another notation for the various absorption and spontaneous or stimulated emission rates: the Einstein A and B coefficients from his 1917 article. In that article, he introduces for the first time the concept of stimulated emission:

$$W_{a\to b}^{\text{abs}} = B_{a\to b} \qquad W_{b\to a}^{\text{stim.em.}} = B_{b\to a} \qquad W_{b\to a}^{\text{spont.em.}} = A_{b\to a} \tag{C-12}$$

One can then write the change per unit time of the populations σ_{bb} and σ_{aa} of the b and a levels due to the various absorption and emission processes:

$$\begin{aligned} \dot{\sigma}_{bb} &= B_{a\to b}\sigma_{aa} - B_{b\to a}\sigma_{bb} - A_{b\to a}\sigma_{bb} \\ \dot{\sigma}_{aa} &= -B_{a\to b}\sigma_{aa} + B_{b\to a}\sigma_{bb} + A_{b\to a}\sigma_{bb} \end{aligned} \tag{C-13}$$

As an example, on the right-hand side of the first line of (C-13), the first term describes how level b fills up when a absorbs a photon, the second term how it is emptied by stimulated emission towards a, the third one how it is emptied by spontaneous emission. Similar explanations can be given for the second line.

In a steady state, there is a balance between the various processes, and we have:

$$\dot{\sigma}_{aa} = \dot{\sigma}_{bb} = 0 \tag{C-14}$$

We then get from the first equation:

$$\sigma_{bb} = \frac{B_{a\to b}}{B_{b\to a} + A_{b\to a}}\sigma_{aa} \tag{C-15}$$

Now, according to the Boltzmann distribution law, σ_{bb}/σ_{aa} must be equal to $e^{-(\hbar\omega_0/k_BT)}$. Relations (B-28), (C-6) and (C-7) then show that, at equilibrium, the populations σ_{bb} and σ_{aa} obey the relation:

$$\frac{\sigma_{bb}}{\sigma_{aa}} = e^{-(\hbar\omega_0/k_BT)} = \frac{B_{a\to b}}{B_{b\to a} + A_{b\to a}} = \frac{\langle n(\omega_0)\rangle}{\langle n(\omega_0)\rangle + 1} \tag{C-16}$$

which means that:

$$\langle n(\omega_0)\rangle = \frac{e^{-(\hbar\omega_0/k_BT)}}{1 - e^{-(\hbar\omega_0/k_BT)}} = \frac{1}{e^{(\hbar\omega_0/k_BT)} - 1} \tag{C-17}$$

Multiplying the average energy per mode $\hbar\omega_0\langle n(\omega_0)\rangle$ by the mode density $8\pi\nu_0^2/c^3$ in the vicinity of ω_0, yields Planck's law for the energy density per unit volume of the black body radiation, as a function of the frequency $\nu_0 = \omega_0/2\pi$:

$$u(\nu_0) = \frac{8\pi h\nu_0^3}{c^3}\frac{1}{e^{(h\nu_0/k_BT)} - 1} \tag{C-18}$$

In other words, when an ensemble of two-level atoms reaches Maxwell-Boltzmann equilibrium through absorption and spontaneous or stimulated emission of radiation, that radiation must necessarily obey Planck's law. This is the essence of the argument used by Einstein to establish this law[9].

[9]Einstein could not reason in 1917 in terms of the quantum theory of radiation, which was not available. The heuristic introduction of the A and B coefficients illustrates his remarkable intuition.

D. Role of correlation functions in one-photon processes

The probabilities associated with one-photon processes can be expressed in terms of atom and field correlation functions.

D-1. Absorption process

Let us write again the probability amplitude for the system, starting at $t_i = 0$ from the state $|\, a; \phi^R_{\rm in}\rangle$, to end up at time $t_f = \Delta t$ in the state $|b; \phi^R_{\rm fin}\rangle$ after undergoing a one-photon transition – see (B-14):

$$\langle \psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle = -\frac{1}{i\hbar}\int_0^{\Delta t} {\rm d}t'\, \langle b|\bar{\boldsymbol{D}}(t')|a\rangle \cdot \langle \phi^R_{\rm fin}|\bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t')|\phi^R_{\rm in}\rangle \tag{D-1}$$

When dealing with an absorption of radiation by the atom, the number of photons in the final state $\left|\phi^R_{\rm fin}\right\rangle$ is lower than in the initial state $\left|\phi^R_{\rm in}\right\rangle$; only the positive frequency component of the field $\bar{\boldsymbol{E}}^{(+)}_\perp$, which can destroy photons, yields a contribution to the matrix element $\left\langle \phi^R_{\rm fin}\right| \bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t') \left|\phi^R_{\rm in}\right\rangle$. Let us start with a radiation initial state containing photons with a single given polarization $\boldsymbol{\varepsilon}_{\rm in}$; only modes with this particular polarization are involved in the matrix element. If the polarization is not linear (circular for instance), $\boldsymbol{\varepsilon}_{\rm in}$ is complex, and we must replace it by its complex conjugate $\boldsymbol{\varepsilon}^*_{\rm in}$ is all negative frequency components of the field. Moreover, since $\boldsymbol{\varepsilon}^*_{\rm in} \cdot \boldsymbol{\varepsilon}_{\rm in} = 1$, the $\boldsymbol{\varepsilon}_{\rm in}$ polarization components are obtained by a scalar product of $\boldsymbol{\varepsilon}^*_{\rm in}$ by the field. We then have:

$$\langle \psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle = -\frac{1}{i\hbar}\int_0^{\Delta t} {\rm d}t'\, \langle b|\, \boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{D}}(t')|a\rangle \langle \phi^R_{\rm fin}|\boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\phi^R_{\rm in}\rangle \tag{D-2}$$

The probability $P^{\rm abs}_{a\to b}(\Delta t)$ of the absorption process is then:

$$\begin{aligned} P^{\rm abs}_{a\to b}(\Delta t) = \frac{1}{\hbar^2}\int_0^{\Delta t} {\rm d}t' \langle a|\, \boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{D}}(t')|b\rangle \langle \phi^R_{\rm in}|\boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t')|\phi^R_{\rm fin}\rangle \\ \times \int_0^{\Delta t} {\rm d}t'' \langle b|\, \boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{D}}(t'')|a\rangle \langle \phi^R_{\rm fin}|\boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')|\phi^R_{\rm in}\rangle \end{aligned} \tag{D-3}$$

To obtain the probability of finding the atom in any final state b other than a, whatever final state $|\phi^R_{\rm fin}\rangle$ the radiation is in, we must sum this result over all possible $|b\rangle$ and $|\phi^R_{\rm fin}\rangle$ states. This yields two closure relations, one in the atom state space[10], the other in the radiation state space. This leads to the following result:

$$\sum_b P^{\rm abs}_{a\to b}(\Delta t) = \frac{1}{\hbar^2}\int_0^{\Delta t} {\rm d}t' \int_0^{\Delta t} {\rm d}t''\; G_{\rm a}(t',t'')\; G^N_R(t',t'') \tag{D-4}$$

with the definitions:

$$G_{\rm a}(t',t'') = \langle a|\left[\boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{D}}(t')\right]\left[\boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{D}}(t'')\right]|a\rangle \tag{D-5}$$

[10] In the summation over b, the initial atomic state a can be included since the matrix element of the atomic dipole in that state is zero (because of a parity argument).

and:

$$G_R^N(t',t'') = \left\langle \phi_{\rm in}^R \right| \left[\boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R},t') \right] \left[\boldsymbol{\varepsilon}_{\rm in}^* \cdot \bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t'') \right] \left| \phi_{\rm in}^R \right\rangle \tag{D-6}$$

The two functions we just defined correspond, respectively, to the correlation function of the atomic dipole and to that of the electric field expressed in normal order.

In the more general case where the field initial state includes several polarizations, (D-1) must now include the matrix elements:

$$\sum_{j=x,y,z} \langle b| \ \bar{D}_j(t') \left|a\right\rangle \left\langle \phi_{\rm fin}^R \right| \bar{E}_j(\boldsymbol{R},t') \left| \phi_{\rm in}^R \right\rangle \tag{D-7}$$

with:

$$\bar{D}_j(t) = \mathbf{e}_j \cdot \bar{\boldsymbol{D}}(t) \quad \text{and} \quad \bar{E}_j(\boldsymbol{R},t) = \mathbf{e}_j \cdot \bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t) \tag{D-8}$$

where $\mathbf{e}_j$ is the unit vector of each of the three $j = x, y, z$ axes. Probability (D-4) then becomes:

$$\sum_b P_{a\to b}^{\rm abs}(\Delta t) = \sum_{j,j'=x,y,z} \frac{1}{\hbar^2} \int_0^{\Delta t} {\rm d}t' \int_0^{\Delta t} {\rm d}t'' \ G_{\rm a}^{jj'}(t',t'') \ G_R^{jj'}(t',t'') \tag{D-9}$$

with the following definitions of the 9 dipole correlation functions, and the other 9 field correlation functions (the vectors $\mathbf{e}_j$ are real) :

$$\begin{aligned} G_{\rm a}^{jj'}(t',t'') &= \langle a| \left[\mathbf{e}_j \cdot \bar{\boldsymbol{D}}(t') \right] \left[\mathbf{e}_{j'} \cdot \bar{\boldsymbol{D}}(t'') \right] |a\rangle \\ G_R^{jj'}(t',t'') &= \left\langle \phi_{\rm in}^R \right| \left[\mathbf{e}_j \cdot \bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R},t') \right] \left[\mathbf{e}_{j'} \cdot \bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t'') \right] \left| \phi_{\rm in}^R \right\rangle \end{aligned} \tag{D-10}$$

We thus get a correlation tensor, which slightly complicates the equations, but does not change the essence of the results.

D-2. Emission process

For the photon emission processes, spontaneous or stimulated, we can make similar computations. The main difference is that the $\bar{\boldsymbol{E}}_\perp^{(+)}$ and $\bar{\boldsymbol{E}}_\perp^{(-)}$ operators must be interchanged, which yields antinormal instead of normal field correlation functions; furthermore, we must use the more general formula (D-9) instead of (D-4) since an emission process does not favor any specific polarization.

E. Photon scattering by an atom

We now consider a photon scattering process where the initial state includes an atom in state a and an incident photon, and where in the final state the atom is still in state a, but the incident photon has been replaced by another one. This is a two-photon process, since one photon disappears, and is replaced by another one.

E-1. Elastic scattering

As in § B, we shall treat classically the center of mass of the atom, supposed to be fixed at the origin of the reference frame (see however the comment on page 2094). The initial state $|\psi_{\rm in}\rangle$ of the system atom + radiation at time t_i corresponds to an atom in state a in the presence of a photon in mode i, with wave vector $\boldsymbol{k}_i$, polarization $\boldsymbol{\varepsilon}_i$ and frequency ω_i (we assume the radiation to be monochromatic):

$$|\psi_{\rm in}\rangle = |a; \boldsymbol{k}_i, \boldsymbol{\varepsilon}_i\rangle \qquad \text{with energy} \qquad E_{\rm in} = E_a + \hbar\omega_i \tag{E-1}$$

At the final time t_f, the scattering process has replaced the initial photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ by a new photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ with wave vector $\boldsymbol{k}_f$ and polarization $\boldsymbol{\varepsilon}_f$. The final state of the system is:

$$|\psi_{\rm fin}\rangle = |a; \boldsymbol{k}_f, \boldsymbol{\varepsilon}_f\rangle \qquad \text{with energy} \qquad E_{\rm fin} = E_a + \hbar\omega_f \tag{E-2}$$

Conservation of total energy requires $\omega_i = \omega_f$. This scattering thus occurs without frequency change and is called for that reason "*elastic scattering*". We shall study in § E-3 the case where the final atomic state is different from the initial atomic state. This different process is called "*Raman scattering*".

As the electric dipole interaction Hamiltonian (A-24) can only change the photon number by one unit, the system must go through an intermediate state often called a "relay state" where the atom is in state c and the radiation in a state different from its initial state. The lowest order term of the series expansion (A-12) that contributes to the scattering amplitude is of order two.

E-1-a. Two possible types of relay state

There are two possible types of relay states: those corresponding to processes we shall label (α), where the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ is absorbed before the photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ is emitted; and those corresponding to processes labeled (β), where the photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ is emitted before the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ is absorbed. In the first case, the relay state is the state $|\psi^{\alpha}_{\rm rel}\rangle = |c; 0\rangle$, where c is an atomic relay state and 0 is the radiation vacuum, since the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ present in the initial state has been absorbed; the energy of this relay state is $E^{\alpha}_{\rm rel} = E_c$. In the second case, the relay state is $|\psi^{\beta}_{\rm rel}\rangle = |c; \boldsymbol{k}_i, \boldsymbol{\varepsilon}_i; \boldsymbol{k}_f, \boldsymbol{\varepsilon}_f\rangle$, since the photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ has been emitted before the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ was absorbed: the energy of this relay state is $E^{\beta}_{\rm rel} = E_c + \hbar\omega_i + \hbar\omega_f$. Figures 2 and 3 show two different diagrams representing these same two processes.

In Figure 2, the horizontal lines represent the atomic levels; an upwards arrow represents an absorption, whereas a downwards arrow represents an emission. The advantage of this representation is to directly show the energy difference between the initial state and the relay state, equal to $E_{\rm in} - E^{\alpha}_{\rm rel} = E_a + \hbar\omega_i - E_c$ for the (α) processes, and to $E_{\rm in} - E^{\beta}_{\rm rel} = E_a - \hbar\omega_f - E_c$ for the (β) processes: this difference is simply the distance between the height of the dashed line and the height of the line representing the atomic relay state c. In particular, these two lines coincide for the (α) processes when $E_a + \hbar\omega_i = E_c$, i.e. when the absorption of the incident photon is resonant for the transition $a \leftrightarrow c$ (resonant scattering, which will be studied later on).

In Figure 3, an incoming arrow represents an absorption, whereas an outgoing one represents an emission. Reading the diagram from bottom to top, one clearly sees which atomic state and photons are present in the initial state, the relay state and the final

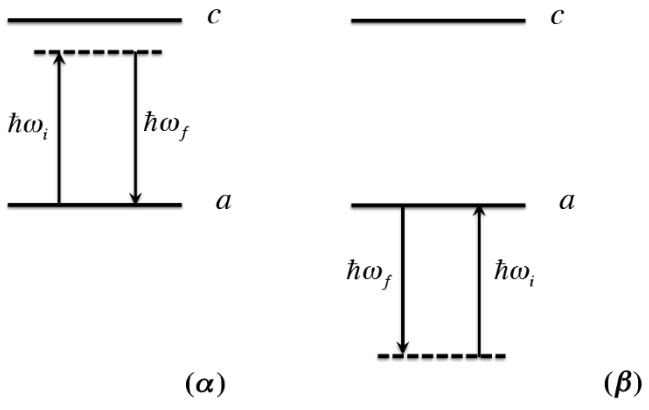


Figure 2: First diagram representation of the scattering processes labeled (α) *and* (β) *in the text; these processes have a different chronological order for the absorption of the incident photon and the emission of the scattered photon. The full horizontal lines represent the atomic levels, and the upwards arrows represent absorption processes and downwards arrows, emission processes; the horizontal dashed lines clearly show the energy differences that will appear in the denominator of the transition amplitude expression.*

state. For the (α) processes, no photons are present in the relay state, whereas both incident and scattered photons are present in that state in the (β) processes.

E-1-b. Computation of the scattering amplitude

The computation of the scattering amplitude is of the same type as the calculations already presented above. In addition, it is almost identical to the computation of the two-photon absorption amplitude explained in detail in § 1 of Complement $\mathrm{A_{XX}}$. Consequently, it will not be explicitly carried out here, but we shall merely highlight the differences with the computation of that complement. The reader interested in more details may want to read that complement before continuing with this paragraph. Relations (13) and (14) of that complement are written here:

$$\begin{aligned}\langle\psi_{\mathrm{fin}}|\bar{U}(t_f,t_i)|\psi_{\mathrm{in}}\rangle &= -2i\pi\left[A^{\alpha}_{fi}(E_{\mathrm{in}})+A^{\beta}_{fi}(E_{\mathrm{in}})\right]\delta^{(\Delta t)}(E_{\mathrm{fin}}-E_{\mathrm{in}})\\ &= -2i\pi\left[A^{\alpha}_{fi}(E_{\mathrm{in}})+A^{\beta}_{fi}(E_{\mathrm{in}})\right]\delta^{(\Delta t)}(\hbar\omega_f-\hbar\omega_i)\end{aligned} \tag{E-3}$$

where we have introduced the probability amplitudes:

$$\begin{aligned}A^{\alpha}_{fi}(E_{\mathrm{in}}) &= \sum_{|\psi^{\alpha}_{\mathrm{rel}}\rangle}\frac{\langle\psi_{\mathrm{fin}}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi^{\alpha}_{\mathrm{rel}}\rangle\langle\psi^{\alpha}_{\mathrm{rel}}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi_{\mathrm{in}}\rangle}{E_{\mathrm{in}}-E^{\alpha}_{\mathrm{rel}}}\\ &= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c\frac{\langle a|\boldsymbol{\varepsilon}^*_f\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c}\end{aligned} \tag{E-4}$$

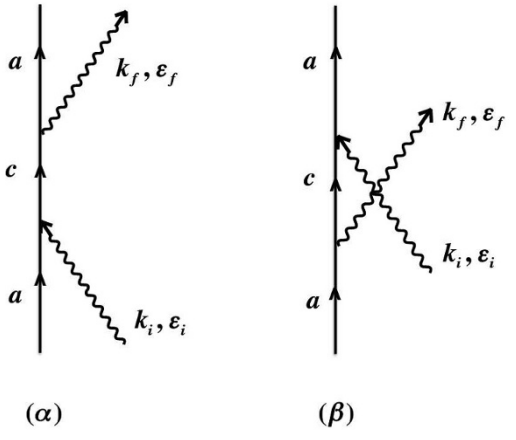


Figure 3: Another possible diagram representation of the scattering processes (α) *and* (β)*. The state of the atom is shown on the vertical full line. As for the photons, the incoming arrows (wiggly arrows pointing towards the vertical full line) represent absorption, whereas outgoing ones (wiggly arrows leaving the vertical full line) represent emissions. The bottom to top reading of each diagram shows the chronological succession of the states of the system atom + photons.*

$$A_{fi}^{\beta}(E_{\rm in}) = \sum_{|\psi_{\rm rel}^{\beta}\rangle} \frac{\langle\psi_{\rm fin}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi_{\rm rel}^{\beta}\rangle\langle\psi_{\rm rel}^{\beta}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi_{\rm in}\rangle}{E_{\rm in}-E_{\rm rel}^{\beta}}$$
$$= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c \frac{\langle a|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|a\rangle}{E_a-\hbar\omega_f-E_c} \qquad \text{(E-5)}$$

The two amplitudes $A_{fi}^{\alpha}(E_{\rm in})$ and $A_{fi}^{\beta}(E_{\rm in})$ correspond to the two types of relay states considered above. In (E-3), the delta function expressing the energy conservation is proportional to $\delta^{(\Delta t)}(\omega_f-\omega_i)$ since $E_{\rm fin}-E_{\rm in}=\hbar(\omega_f-\omega_i)$. To write (E-4) and (E-5), we have replaced in relation (13) of Complement $\mathrm{A_{XX}}$ the interaction Hamiltonian H_I by $-\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}$ or $-\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}$, depending on whether it pertains to an absorption or an emission process. In the numerator of the fractions on the right-hand side of (E-4), operator $\boldsymbol{E}^{(+)}$ (which absorbs the incident photon) acts before operator $\boldsymbol{E}^{(-)}$ (which creates the scattered photon); this is to be expected for an (α) type process. The order of the two operators $\boldsymbol{E}^{(+)}$ and $\boldsymbol{E}^{(-)}$ is reversed in (E-5), as expected for a (β) type process. The $\sqrt{\omega_i\omega_f}$ coefficients on the second lines of these equalities come from the plane wave expansion (A-27) of the electric field; L is the edge of the cubic cavity used to quantize the field. We could have added a factor $\sqrt{n_i n_f}$, where n_i and n_f are the photon numbers in the initial and final states; for the sake of simplicity, we have assumed that $n_i=n_f=1$.

E-1-c. Semiclassical interpretation

Elastic scattering can also be explained by a semiclassical treatment, where the quantum treatment only applies to the atom; the incident wave is described as a classical field of frequency ω_i. This wave induces, in the atom, an oscillating dipole at the same frequency. This dipole radiates into the entire space a field oscillating at that frequency.

The semiclassical approach also enables a simple interpretation of the absorption of the incident beam as resulting from a destructive interference, in the direction of the incident field, between that field and the field scattered by the induced dipole. One can also use such a description to account for the amplification of an incident beam by an ensemble of atoms whose population has been "inverted", meaning atoms for which the population of an excited level is larger than the population of a lower energy level. The scattered field then has the opposite phase of that it would have without population inversion, so that the interference becomes constructive.

E-1-d. Rayleigh scattering

Assume that the frequency ω_i of the radiation is much smaller than all the atomic frequencies $|E_c - E_a|/\hbar$. One can then ignore ω_i and ω_f in the denominators on the second lines of (E-4) and (E-5). The only ω_i dependence of the scattering amplitudes comes from the prefactor $\sqrt{\omega_i \omega_f}$, equal to ω_i since $\omega_f = \omega_i$. The scattering cross section involves the product of the squared modulus of that amplitude, proportional to ω_i^2, by the density of the radiation's final states at frequency $\omega_f = \omega_i$, also proportional to ω_i^2. The scattering cross section therefore varies as ω_i^4, much higher for blue light than for red light.

One usually calls "Rayleigh scattering" the elastic scattering when $\omega_i \ll |E_c - E_a|/\hbar$. It explains the scattering of the visible solar light by the atmospheric oxygen and nitrogen molecules, which have much higher resonant frequencies, in the ultraviolet domain. This rapid variation with frequency of the Rayleigh scattering cross section is a reason for the sky being blue.

E-2. Resonant scattering

Assume now that the frequency ω_i of the incident photon is very close to the frequency:

$$\omega_{ca} = (E_c - E_a)/\hbar \tag{E-6}$$

of a transition between a state a and a state c having a higher energy. The absorption of the incident photon is then resonant for the transition $a \leftrightarrow c$ and the amplitude (E-4) becomes very large when the state c becomes the atomic relay state – it even diverges if the resonant condition is exactly satisfied. In that case, one can neglect all the (β) type processes; in addition, even if there are other possible atomic relay states d, e,.., one can keep only the term involving c.

To avoid the difficulties related to the divergence of (E-4) when $\omega_i = \omega_{ca}$, it is convenient to use the exact expression (A-14), which only involves three terms, instead of an infinity as in expansion (A-12) for the evolution operator. Only the last of those three terms plays a role, since it can destroy a photon and create another one. A computation similar to that leading to relation (6) of Complement $\mathrm{A_{XX}}$, but using (A-14) instead of

(A-12), then yields:

$$\langle\psi_{\rm fin}|U(t_f,t_i)|\psi_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \int_{t_i}^{t_f} {\rm d}t \int_{t_i}^{t} {\rm d}t' \; \langle\psi_{\rm fin}|U_0(t_f,t)\, H_I\, U(t,t')H_I\, U_0(t',t_i)|\psi_{\rm in}\rangle \tag{E-7}$$

Note that it is now $U(t,t')$ that appears in the middle of the matrix element in (E-7), and not $U_0(t,t')$. Starting from $|\psi_{\rm in}\rangle = |a;\boldsymbol{k}_i,\boldsymbol{\varepsilon}_i\rangle$, the first interaction Hamiltonian H_I of (E-7) brings the system to the state $|c;0\rangle$. In a similar way, if the system starts from $\langle\psi_{\rm fin}|$ the second interaction Hamiltonian of (E-7) brings it to the state $\langle c;0|$. Expression (E-7) can then be written:

$$\begin{aligned}\langle\psi_{\rm fin}|U(t_f,t_i)|\psi_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \int_{t_i}^{t_f} {\rm d}t \int_{t_i}^{t} {\rm d}t' \langle a;\boldsymbol{k}_f,\boldsymbol{\varepsilon}_f|U_0(t_f,t)\, H_I\,|c;\,0\rangle \\ \times \langle c;\,0|U(t,t')|c;\,0\rangle\langle c;\,0|H_I\, U_0(t',t_i)|a;\boldsymbol{k}_i,\boldsymbol{\varepsilon}_i\rangle\end{aligned} \tag{E-8}$$

Had we used (A-12) instead of (A-14), the central matrix element of relation (E-8) would be $\langle c;\,0|U_0(t,t')|c;\,0\rangle = \exp(-iE_c(t-t')$. In our case, using (A-14) leads to $\langle c;\,0|U(t,t')|c;\,0\rangle$ which is the probability amplitude for the system, starting from the state $|c;\,0\rangle$ at time t', to still be in that same state at time t. The calculation of that amplitude appears in the study of the radiative decay of the excited state through spontaneous emission of a photon, hence the decay of a discrete state $|c;\,0\rangle$ coupled to a continuum of final states $|a;\boldsymbol{k},\boldsymbol{\varepsilon}\rangle$. Those states represent the atom in state a in the presence of a photon with any wave vector $\boldsymbol{k}$ and polarization $\boldsymbol{\varepsilon}$. Now we showed in Complement $\rm D_{XIII}$ (§ 4) that it was possible to obtain a solution of Schrödinger's equation yielding the amplitude $\langle c;\,0|U(t,t')|c;\,0\rangle$ at long times (and not only at short times, as for the perturbative solution). This solution is written:

$$\langle c;\,0|U(t,t')|c;\,0\rangle = \exp\left[-i(E_c+\delta E_c)(t-t')/\hbar\right]\exp\left[-\Gamma_c(t-t')/2\right] \tag{E-9}$$

where δE_c is the energy shift of the state $|c;\,0\rangle$ due to its coupling with the continuum of final states[11], and Γ_c the natural width of the excited state c (the inverse of the radiative lifetime of that state). We shall assume from now on that the shift δE_c is included in the definition of the energy E_c of the state c. Starting from the more precise expression (A-14) instead of (A-12) thus leads to a very simple result: we just have to replace, in all the computations of the scattering amplitude, E_c by $E_c - i\hbar\Gamma_c/2$.

$$E_c \to E_c - i\hbar\Gamma_c/2 \tag{E-10}$$

Once this replacement has been made, we get, keeping only the amplitude (E-4) and a single relay state c, the following scattering amplitude:

$$\bar{A}_{fi}(E_{\rm in}) = \frac{\hbar\omega_i}{2\varepsilon_0 L^3}\, \frac{\langle a|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{\hbar(\omega_i-\omega_{ca}+i\Gamma_c/2)} \tag{E-11}$$

This resonant scattering amplitude no longer diverges when $\omega_i = \omega_{ca}$; as ω_i is scanned around ω_{ca}, it exhibits a resonant behavior over a range equal to Γ_c.

[11] This shift is related to the "Lamb shift" of the excited atomic states.

E-3. Inelastic scattering - Raman scattering

We now consider a scattering process where, as before, an incident photon is absorbed and another emitted, but the final atomic state a' is now supposed to be different from the initial atomic state a.

E-3-a. Differences with elastic scattering

Figure 4 shows an example of such a process, called "*Raman scattering*", where the energy of the scattered photon is different from that of the incident photon[12]. The initial state of the scattering process is, as before, the state $|\psi_{\rm in}\rangle = |a; \boldsymbol{k}_i, \boldsymbol{\varepsilon}_i\rangle$, with energy $E_{\rm in} = E_a + \hbar\omega_i$; the final state, however, is now $|\psi_{\rm fin}\rangle = |a'; \boldsymbol{k}_f, \boldsymbol{\varepsilon}_f\rangle$ where $a' \neq a$, with energy $E_{\rm fin} = E_{a'} + \hbar\omega_f$.

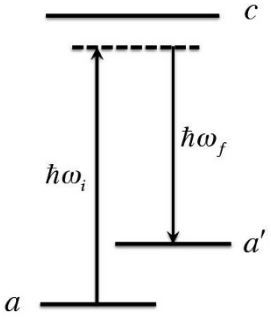


Figure 4: Raman scattering: an atom in state a absorbs an incident photon, with energy $\hbar\omega_i$; a photon $\hbar\omega_f$ is then spontaneously emitted by the atom, which ends up into a final state a' different from the initial state a.

Conservation of total energy requires:

$$E_a + \hbar\omega_i = E_{a'} + \hbar\omega_f \tag{E-12}$$

If $E_{a'} > E_a$, the Raman scattering is called "*Raman Stokes scattering*"; the energy of the scattered photon is lower than that of the incident photon. If $E_{a'} < E_a$, the Raman scattering is called "*Raman anti-Stokes scattering*"; the energy of the scattered photon is higher than that of the incident photon. As we assumed here that the mode $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$ was initially empty, the scattered photon is emitted spontaneously. The process is then called "*spontaneous Raman scattering*". We shall study later the case where $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$ photons are initially present, a situation resulting in "*stimulated Raman scattering*".

Equation (E-12) shows that the angular frequency of the scattered light is different from that of the incident light by a quantity $\omega_{aa'} = (E_a - E_{a'})/\hbar$, equal to the frequency of the atomic transition $a \leftrightarrow a'$. This means that Raman light spectrum provides information about the eigenfrequencies of the scattering system; this is the base of *Raman*

[12] This figure only shows a type (α) process, where the incident photon is absorbed in the first place.

spectroscopy. The systems under study are often molecules and the states a, a' are vibrational or rotational sublevels of the ground state, which means that the $\omega_{aa'}$ frequencies belong to the microwave or infrared domain. Instead of measuring directly these frequencies using spectroscopic techniques in a frequency domain where detection might be a problem, it is sometimes more convenient to illuminate the medium with an optical or ultraviolet frequency beam, and to measure the $\omega_{aa'}$ frequency via the frequency shifts of the Raman scattered light. Raman spectroscopy developed considerably once laser sources became available, yielding much higher signal intensities. The detection conditions have also been greatly improved by focusing the incident laser light on very small volumes. The analysis of the scattered light spectrum is now a very powerful tool for analyzing the chemical composition of any scattering media, since each molecule can be identified by its specific vibration-rotation eigenfrequencies.

To keep things simple, we have limited our discussion to Raman scattering by atoms or molecules in a dilute medium, where each scattering entity acts individually. Raman scattering is also used in condensed media such as liquids, crystals, surfaces, etc. and provides valuable information on the dynamics of these structures.

E-3-b. Scattering amplitude

The computation of the Raman scattering amplitude is very similar to the one leading to (E-3), (E-4) and (E-5), and yields:

$$\begin{aligned}\langle\psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle &= -2\pi i\left[A^{\alpha}_{fi}(E_{\rm in})+A^{\beta}_{fi}(E_{\rm in})\right]\delta^{(\Delta t)}(E_{\rm fin}-E_{\rm in})\\ &= -2\pi i\left[A^{\alpha}_{fi}(E_{\rm in})+A^{\beta}_{fi}(E_{\rm in})\right]\delta^{(\Delta t)}(E_{a'}+\hbar\omega_f-E_a-\hbar\omega_i)\end{aligned}\quad\text{(E-13)}$$

where:

$$\begin{aligned}A^{\alpha}_{fi}(E_{\rm in}) &= \sum_{|\psi^{\alpha}_{\rm rel}\rangle}\frac{\langle\psi_{\rm fin}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi^{\alpha}_{\rm rel}\rangle\langle\psi^{\alpha}_{\rm rel}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi_{\rm in}\rangle}{E_{\rm in}-E^{\alpha}_{\rm rel}}\\ &= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c\frac{\langle a'|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c}\end{aligned}\quad\text{(E-14)}$$

$$\begin{aligned}A^{\beta}_{fi}(E_{\rm in}) &= \sum_{|\psi^{\beta}_{\rm rel}\rangle}\frac{\langle\psi_{\rm fin}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi^{\beta}_{\rm rel}\rangle\langle\psi^{\beta}_{\rm rel}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi_{\rm in}\rangle}{E_{\rm in}-E^{\beta}_{\rm rel}}\\ &= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c\frac{\langle a'|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|a\rangle}{E_a-\hbar\omega_f-E_c}\end{aligned}\quad\text{(E-15)}$$

When the photon frequency ω_i is close to the $a\to c$ transition frequency, Raman scattering becomes resonant and amplitude (E-14) can become very large. As we did for the resonant elastic scattering, to avoid the divergence of (E-14), we just have to replace E_c by $E_c-i\Gamma_c/2$, where Γ_c is the natural width of the c state.

E-3-c. Semiclassical interpretation

As in § E-1-c, let us consider the dipole induced by the incident field on the scattering object. When that object is a molecule which vibrates and rotates, its polarizability

changes with time, and is modulated by its rotation and vibration frequencies. The dipole's oscillations induced by the incident field have an amplitude modulated at the rotation and vibration frequencies of the molecule. The Fourier spectrum of the dipole's motion contains lateral bands at frequencies shifted from the incident field frequency; these frequency shifts are equal to the molecule's rotation and vibration frequencies. This semiclassical interpretation accounts for the essential properties of the Raman scattering spectrum.

E-3-d. Stimulated Raman scattering. Raman laser

We now assume the Raman photon appears in a mode that is not initially empty, as $n_f \neq 0$ photons already occupy the mode $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$. Similarly, we assume several photons (n_i for example), initially occupy the mode $(\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i)$. To compute the Raman scattering amplitude $|a; n_i, n_f\rangle \rightarrow |a'; n_i - 1, n_f + 1\rangle$, we must include the factor $\sqrt{n_i(n_f + 1)}$ in expressions (E-14) and (E-15). Stimulated emission now comes into play: the factor $n_f + 1$ appearing in the probability expresses the fact that the initial presence of n_f photons in mode $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$ stimulates the emission probability of a Raman photon in that mode.

Consider now (right-hand side of Fig.5) the inverse scattering process symbolized by $|a'; n_i - 1, n_f + 1\rangle \rightarrow |a; n_i, n_f\rangle$. The corresponding scattering amplitude is simply the complex conjugate of the previous one, meaning that the probability of these two processes are equal. If we start with the same number of atoms in state a and state a', the number of photons created in one of the processes is equal to the number of photons that

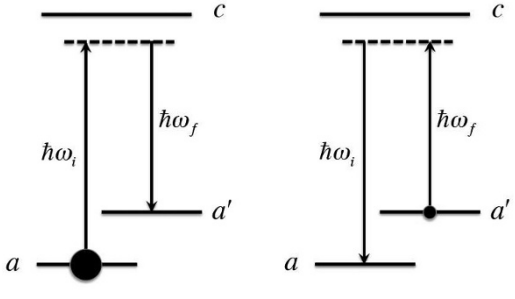


Figure 5: The left-hand side of the figure represents a stimulated Raman process where an atom in level a absorbs a photon ω_i, and ends up in state a' after the stimulated emission of an ω_f photon. The right-hand side of the figure shows the inverse process, where an atom starting from state a' absorbs an ω_f photon, and falls back in state a after the stimulated emission of an ω_i photon. These two processes have, a priori, the same probability. However, if the population of state a (shown as a large dot on the left-hand side of the figure) is higher than that of the state a' (shown as a small dot on the right-hand side of the figure), the number of processes resulting in the stimulated emission of an ω_f photon is bigger than the number of processes where that photon is absorbed. The radiation at frequency ω_f is thus amplified by stimulated emission, which allows creating a "Raman laser" at this frequency.

disappear in the inverse process. What will happen now if we start with an ensemble of atoms where the populations of states a and a' are not equal? If, for example, level a has a lower energy than level a', the relaxation mechanisms leading to thermal equilibrium tend to create a larger population in a than in a'. The number of scattering processes $|a; n_i, n_f\rangle \rightarrow |a'; n_i - 1, n_f + 1\rangle$ is then larger than the number of the inverse processes $|a'; n_i - 1, n_f + 1\rangle \rightarrow |a; n_i, n_f\rangle$, leading to an amplification of the number of ω_f photons. This amplification mechanism is the basis of *Raman laser* operation. This type of laser is different in two major ways from lasers involving a transition between an upper level, populated by a pumping process, and a lower level (with no relay level). First of all, they do not require a population inversion; the atomic media can be at thermal equilibrium, since the stimulated Raman scattering at the origin of the amplification starts from the atomic state with the largest population. They do require, however, a high intensity radiation at frequency ω_i, furnished by another laser called the "pump laser". Secondly, the ω_f frequency of the Raman laser oscillation can be scanned by changing the "pump frequency" ω_i, whereas lasers using a two-level system necessarily oscillate at a frequency very close to the atomic transition, and have thus a very small tuning range.

Conservation of total momentum

If the position $\boldsymbol{R}$ of the atomic center of mass is no longer treated classically, and placed at the origin as we have done until now, we must keep the exponential functions $\exp(i\boldsymbol{k}_i \cdot \boldsymbol{R})$ and $\exp(-i\boldsymbol{k}_f \cdot \boldsymbol{R})$ in the interaction Hamiltonian describing the absorption of an ω_i photon and the emission of an ω_f photon. The matrix element of the product of those two operators must be taken between an initial state of the center of mass, with momentum $\hbar\boldsymbol{K}_{\text{in}}$ and a final state with momentum $\hbar\boldsymbol{K}_{\text{fin}}$. This yields a $\delta(\boldsymbol{K}_{\text{fin}} - \boldsymbol{K}_{\text{in}} - \boldsymbol{k}_i + \boldsymbol{k}_f)$ function expressing the global momentum conservation in a Raman process: the momentum of the atom increases by the quantity $\hbar(\boldsymbol{k}_i - \boldsymbol{k}_f)$ during that process. It often happens that the two atomic states a and a' are two sublevels of the same electronic ground state, so that the frequency $\omega_{a'a} = (E_{a'} - E_a)/\hbar$ falls in the microwave domain; it is then much smaller than the frequencies ω_i and ω_f, which are optical frequencies. The energy conservation equation (E-12) then shows that the moduli of the two wave vectors $\boldsymbol{k}_i$ and $\boldsymbol{k}_f$ are practically the same. If the two wave vectors $\boldsymbol{k}_i$ and $\boldsymbol{k}_f$ have opposite directions, the momentum gained by the atom during a Raman process is equal to $\hbar(\boldsymbol{k}_i - \boldsymbol{k}_f) \simeq 2\hbar\boldsymbol{k}_i$. The interest of such a Raman process is to couple two states a and a', energetically very close to each other, by transferring to the atom in one of the two states a very large momentum $2\hbar\boldsymbol{k}_i$, equal to twice that of an optical photon. On the other hand, if the two states a and a' were to be coupled directly by absorption of a single photon in the microwave domain, the momentum transfer would be much smaller. This possibility of coupling two sublevels of the ground state (hence having long lifetimes) by transferring a large momentum to the atom in one of these two states, has interesting applications, in particular in atomic interferometry.

To remain concise, in this chapter we have not treated a certain number of interesting related problems. Among them are multophotonic processes, photoionization, the dressed atom method that facilites the study of light shifts, or the use of photon wave packets. All these subjects are treated in the complements of this chapter.

A_{XX} : A MULTIPHOTON PROCESS: TWO-PHOTON ABSORPTION

In a multiphoton absorption process, an atom simultaneously absorbs several photons. This complement focuses on the simplest case where two photons are absorbed, while presenting general ideas that also apply to processes involving a larger number of photons. Monochromatic and broad band excitations are successively considered. The very short time the system spends in the relay state violating energy conservation is proportional to the inverse of that energy mismatch.

B_{XX} : PHOTOIONIZATION

In a photoionization process, a photon can remove an electron from an atom, which then becomes an ion (photoelectric effect). This complement studies this process by using a quantum theory of radiation that no longer couples two discrete atomic states but rather a discrete (ground) state to a continuum of (excited) states. Two important cases are considered: a quasi-monochromatic incident radiation, and a broad band excitation. In that second case, this study provides a justification for Einstein's equation of the photoelectric effect. Lastly, we consider the case where the radiation field is so very intense that the atomic ionization no longer occurs through the absorption of one or several photons, but rather by the tunnel effect.

C_{XX} : TWO-LEVEL ATOM IN A MONOCHROMATIC FIELD. DRESSED ATOM APPROACH

The dressed atom approach is a powerful tool for describing and interpreting higher order effects that appear as a two-level atom interacts with a quasi-resonant radiation. It is valid both in the weak coupling domain (low field intensity) and the strong coupling domain (very high field intensity). An essential parameter is the ratio of the Rabi oscillation (characterizing the coupling with the field) and the natural width of the atomic levels. This general approach allows, in particular, a full understanding of the various properties of light shifts.

D_{XX} : LIGHT SHIFTS: A TOOL FOR MANIPULATING ATOMS AND FIELDS

Using light shifts has become a basic tool in atomic physics, as it allows manipulating atoms and photons. A number of applications of such methods are briefly described in this complement: laser trapping of atoms by dipole forces, mirrors for atoms, optical lattices, "Sisyphus" cooling, and one by one detection of photons in a cavity.

E_{xx} : DETECTION OF ONE- OR TWO-PHOTON WAVE PACKETS, INTERFERENCE

Just as for a massive particle, on can build wave packets for a photon by the coherent superposition of states, each having a different momentum; this leads to a description of radiation propagation in free space, and we have the possibility to model the arrival of a photon on an atom. We obtain a description of the photon absorption or scattering processes that is more realistic than that given in Chapter XX, where the incident radiation is described by a Fock state having a well defined number of photons (and hence without any spatial propagation). We introduce for the photon a function that is not its wave function, but rather yields the probability amplitude that it might be detected at a given point. Absorption and scattering of wave packets are studied, as well as the one- or two-photon detection signals; the case of two entangled photons (parametric down-conversion) is treated at the end of the complement.

Complement A_{XX}

A multiphoton process: two-photon absorption

1 Monochromatic radiation 2097
2 Non-monochromatic radiation 2101
2-a Probability amplitude, probability 2101
2-b Probability per unit time when the radiation is in a Fock state 2103
3 Discussion 2105
3-a Conservation laws 2105
3-b Case where the relay state becomes resonant for one-photon absorption 2106

In the process studied in this complement the atom absorbs, not one, but two photons of energy $\hbar\omega_i$, to go from a discrete level a to another discrete higher energy level b[1]; this process is schematized in Figure 1. For the moment we shall ignore the external degrees of freedom and suppose the atom to be infinitely heavy[2]. Conservation of total energy then requires:

$$E_b - E_a = \hbar\omega_0 = 2\hbar\omega_i \tag{1}$$

The calculation of the transition amplitude will explicit the role played by this conservation law.

1. Monochromatic radiation

Studying this process involves the computation of the transition amplitude:

$$\langle\psi_{\text{fin}}|U(t_f,t_i)|\psi_{\text{in}}\rangle \tag{2}$$

The initial state at time t_i of the system atom + radiation is:

$$|\psi_{\text{in}}\rangle = |a;n_i\rangle \qquad \text{with energy} \qquad E_{\text{in}} = E_a + n_i\hbar\omega_i \tag{3}$$

It describes an atom in state a in the presence of n_i photons in mode i, with wave vector $\boldsymbol{k}_i$, polarization $\boldsymbol{\varepsilon}_i$ and frequency ω_i (we assume the radiation is monochromatic). After the absorption process, at time t_f, the photon number is lowered by two units, going from n_i to $n_i - 2$. The final state of the system is:

$$|\psi_{\text{fin}}\rangle = |b;n_i-2\rangle \qquad \text{with energy} \qquad E_{\text{fin}} = E_b + (n_i-2)\hbar\omega_i \tag{4}$$

The atom-radiation interaction is described, as before, by the electric dipole Hamiltonian (A-30) of Chapter XX, which lowers the photon number by a single unit. This

[1] In Complement B_{XX}, we shall see how multiphoton processes can also make an atom go from a discrete state to a final state belonging to a continuum of states (photoionization).

[2] When the atom's mass is finite, there are interesting physical effects arising from the conservation of total momentum; these will be studied later on (cf. § 3; see also § 2 of Complement A_{XIX}).

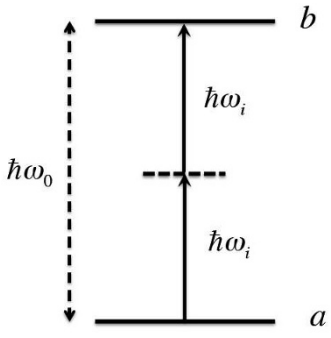


Figure 1: During a two-photon transition, the atom goes from state a to state b by absorbing two photons of energy $\hbar\omega_i$. The horizontal dashed line represents the energy half-way between the a and b levels. A third level, the relay level c, is also involved in the transition; its energy is not necessarily between those of levels a and b and it is not shown in the figure. However, we assume that the energy of that relay atomic level c is so far from the dashed line that no one-photon resonant transition can occur between a and c.

means that in the expansion (A-21) of Chapter XX for $\bar{U}(t_f, t_i)$, the lowest term that gives a non-zero contribution to the transition amplitude (2) is of order two, hence containing two operators H_I. The first brings the system atom + radiation from the initial state $|\psi_{\rm in}\rangle = |a, n_i\rangle$ to an intermediate "relay state"

$$|\psi_{\rm rel}\rangle = |c, n_i - 1\rangle \qquad \text{with energy} \qquad E_{\rm rel} = E_c + (n_i - 1)\hbar\omega_i \tag{5}$$

where c is any atomic state; the second operator H_I brings the system from this relay state $|\psi_{\rm rel}\rangle$ to the final state $|\psi_{\rm fin}\rangle = |b, n_i - 2\rangle$. One must of course sum over all accessible intermediate states c. Nevertheless, to keep the computation simple, we shall only take into account a single intermediate state (the summation of the probability amplitudes over several such states does not pose a serious problem). If we insert relation (A-21) of Chapter XX between the bra $\langle\psi_{\rm fin}|$ and the ket $|\psi_{\rm in}\rangle$, the first two terms on the right-hand side yield zero, and the third one becomes[3]:

$$\langle\psi_{\rm fin}|\bar{U}(\Delta t, 0)|\psi_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \sum_{\psi_{\rm rel}} \langle\psi_{\rm fin}|H_I|\psi_{\rm rel}\rangle\langle\psi_{\rm rel}|H_I|\psi_{\rm in}\rangle \times$$
$$\times \int_0^{\Delta t} {\rm d}t' \int_0^{t'} {\rm d}t''\ e^{iE_{\rm fin}t'/\hbar}\ e^{-iE_{\rm rel}(t'-t'')/\hbar}\ e^{-iE_{\rm in}t''/\hbar} \tag{6}$$

Let us write explicitly the argument of the exponents appearing in the integral on

[3]We assumed that the two absorbed photons were identical. If the two absorbed photons 1 and 2 were different, either by their energies, their wave vector directions, or their polarizations (albeit satisfying the conservation of energy $\hbar\omega_1 + \hbar\omega_2 = E_b - E_a = \hbar\omega_0$), this would lead to a situation similar to that of § E-1-a of Chapter XX. Two types of processes should then be considered, those where photon 1 is first absorbed, photon 2 next, and those where the photons are absorbed in the inverse order (*cf.* Figure 2).

the second line:

$$[E_b + (n_i - 2)\hbar\omega_i]\, t' - [E_c + (n_i - 1)\hbar\omega_i]\, (t' - t'') - [E_a + n_i\hbar\omega_i]\, t'' \tag{7}$$

The terms in $n_i t'$ and $n_i t''$ cancel out, leaving the expression:

$$[E_b - E_a - 2\hbar\omega_i]\, t' - [E_a - E_c + \hbar\omega_i]\, (t'' - t') \tag{8}$$

Choosing t' and $\tau = t'' - t'$ as the integral variables, the second line of (6) becomes the double integral:

$$\int_0^{\Delta t} \mathrm{d}t'\ e^{i[E_b - E_a - 2\hbar\omega_i]t'/\hbar} \int_{-t'}^{0} \mathrm{d}\tau\ e^{i[E_c - E_a - \hbar\omega_i]\tau/\hbar} \tag{9}$$

or else:

$$\int_0^{\Delta t} \mathrm{d}t' e^{i[E_b - E_a - 2\hbar\omega_i]t'/\hbar} \frac{i\hbar}{[E_a + \hbar\omega_i - E_c]} \left[1 - e^{i[E_a + \hbar\omega_i - E_c]t'/\hbar}\right] \tag{10}$$

We are dealing with situations where the frequency $\omega_i/2\pi$ of the photons is close to the two-photon resonance expressed by the conservation of energy (see relation (1)). In addition, we assume that the process is a real direct two-photon transition, and not successive one photon absorptions. In other words, we assume level a cannot absorb a photon in a resonant way and get to the intermediate level c; this means that the energy E_c of that intermediate level c must be very different from the half-way energy $(E_a + E_b)/2$ shown as the horizontal dashed line in Figure 1. It is easy to see that the first term 1 in the bracket of (10) does correspond to a two-photon resonant absorption, since its probability amplitude is written:

$$\frac{i\hbar}{[E_a + \hbar\omega_i - E_c]} \int_0^{\Delta t} \mathrm{d}t' e^{i[E_b - 2\hbar\omega_i - E_a]t'/\hbar} \sim \frac{2i\pi\hbar^2}{[E_a + \hbar\omega_i - E_c]} \delta^{(\Delta t)}\,(E_b - E_a - 2\hbar\omega_i) \tag{11}$$

(the sign $\sim$ simply means we have ignored an irrelevant phase factor), with:

$$\delta^{(\Delta t)}\,(E) = \frac{1}{\pi} \frac{\sin(E\ \Delta t/2\hbar)}{E} \tag{12}$$

The function $\delta^{(\Delta t)}$ tends towards a delta function when $\Delta t \to \infty$, as shown by relation (10) in Appendix II; it expresses an energy conservation satisfying condition (1), within $\hbar/\Delta t$. On the other hand, the second term in the bracket of (10) introduces, in the sum over t', an exponential $e^{i[E_b - E_c - \hbar\omega_i]t'/\hbar}$ that oscillates rapidly as a function of t' when condition (1) is (exactly or approximately) satisfied[4]. This term yields a non-resonant contribution, hence negligible. Its physical significance (sudden branching of the coupling between atom and field) will be discussed in comment (ii) below. We shall ignore it for the moment because of its non-resonant character. This leads us to:

$$\langle \psi_{\mathrm{fin}} | \bar{U}(\Delta t, 0) | \psi_{\mathrm{in}} \rangle = -2i\pi \frac{\langle \psi_{\mathrm{fin}} | H_I | \psi_{\mathrm{rel}} \rangle \langle \psi_{\mathrm{rel}} | H_I | \psi_{\mathrm{in}} \rangle}{E_a + \hbar\omega_i - E_c}\ \delta^{(\Delta t)}\,(E_b - E_a - 2\hbar\omega_i) \tag{13}$$

[4] In that case, $e^{i[E_b - E_c - \hbar\omega_i]t'/\hbar} \simeq e^{i[E_b + E_a - 2E_c]t'/2\hbar}$, whose exponent is necessarily large since we assumed that E_c is very different from the half-sum of E_a and E_b.

Comparison with relation (B-10) of Chapter XX shows a great similarity between the probability amplitude of a one-photon absorption process and that of a two-photon transition. We go from the first to the second by substituting the variable in the function $\delta^{(\Delta t)}$ by the one relevant for the two-photon energy conservation written in (1), and by replacing the matrix element $\langle\psi_{\text{fin}}|H_I|\psi_{\text{in}}\rangle$ by[5]:

$$\frac{\langle\psi_{\text{fin}}|H_I|\psi_{\text{rel}}\rangle\langle\psi_{\text{rel}}|H_I|\psi_{\text{in}}\rangle}{E_{\text{in}}-E_{\text{rel}}} = -\frac{\hbar\omega_i}{2\varepsilon_0 L^3}\sqrt{n_i(n_i-1)}\ \ \frac{\langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c} \tag{14}$$

This means that we just have to replace, in relation (B-6) of Chapter XX for the transition amplitude, the matrix element H_I by a product of matrix elements divided by an energy difference.

Comments

(i): Characteristic time of the intermediate transition

The transit of the physical system through the intermediate relay state occurs without energy conservation, since it involves a difference $\delta E = E_a + \hbar\omega_i - E_c$ with the initial energy. Mathematically, this results in the presence, in the second time integral of (9), of an oscillating term; the larger the energy difference, the more rapid the oscillation. Once that integral is performed, we obtain the bracket appearing in (10), multiplied by a prefactor. This bracket starts from zero at time $t' = 0$, then oscillates as a function of the intermediate time t'. After a time $t' = h/\delta E$, which corresponds to one oscillation period, its average value over time equals one, precisely the value we have used for the computation of the probability amplitude.

The transit through the relay state brings in a characteristic time $\delta\tau = h/\delta E$, after which the modulus of the integral over time τ no longer increases. The larger the departure from energy conservation, the shorter that time is (this short transit through such a relay state is sometimes referred to as a "virtual transition"). The integral over time τ thus behaves completely differently from the integral over t', which, at resonance, increases linearly with time as shown from (11) and (12); this latter integral over t' may accumulate contributions over much larger times. The limitation of the times τ that actually contribute to the probability amplitude has a natural interpretation in the context of the Heisenberg time-energy uncertainly relation $\delta E \times \delta\tau \geq \hbar$.

(ii): Physical meaning of the term left out of the transition amplitude

We have left out the second term in the bracket of relation (10). Its origin is nevertheless interesting, as it arises from the sudden branching of the coupling between atom and field at time $t = 0$, as assumed in the computation. To see this, we can use a model where the interaction Hamiltonian H_I is replaced by an operator $f(t)H_I$; the time dependence of the function $f(t)$ allows introducing an adiabatic turning on of the coupling. It can be shown that the term we had ignored does disappear when turning on the interaction very slowly.

A more rigorous description can be obtained by describing the field as a wave packet propagating in space (Complement E_{XX}), and overlapping the atom only during a limited time. In that case, the interaction Hamiltonian only acts during that overlap time, even

[5] We use for H_I expression (A-24) of Chapter XX, as well as expression (A-27) for the electric field.

though the operator itself is time independent. For a wave packet with a progressive wave front, this approach shows that the term we have ignored does not even come into play.

As for the transition probability, the computation is the same as for a one-photon transition. We get for the transition probability $P^{(2)}_{a\to b}(t)$:

$$P^{(2)}_{a\to b}(t) = \left(\frac{\hbar\omega_i}{2\hbar\varepsilon_0 L^3}\right)^2 [n_i(n_i-1)] \left|\frac{\langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c}\right|^2 \frac{4\sin^2(\omega_0-2\omega_i)\,\Delta t/2}{(\omega_0-2\omega_i)^2} \tag{15}$$

At short times, it is proportional to the square of $(\Delta t)^2$.

Finally, if several relay states are involved in the two-photon transition, one must sum expressions (13) and (14) over all the possible intermediate states $|\psi_{\rm rel}\rangle$ (taking into account the fact they each have different energies $E_{\rm rel}$); on the right-hand side of (14), this amounts to summing over all accessible intermediate atomic states c with energy E_c. Interference effects between amplitudes associated with different intermediate states c may then appear in the probability (15).

2. Non-monochromatic radiation

Consider now what happens when the initial state of the system $|\psi_{\rm in}\rangle$ contains photons of different frequencies. We are going to show that, just as for one-photon transitions, the two-photon transitions can lead to a transition probability per unit time; however, this probability involves a higher order correlation function (order 4 instead of 2).

2-a. Probability amplitude, probability

The computation of the probability amplitude is similar to that discussed in § 1; it is based, as before, on the expression of $\bar{U}$ to second order in H_I. We will carry out the calculation so as to highlight the properties of the time correlation functions of the incident electric field. The radiation initial state is $|\,\phi^R_{\rm in}\rangle$, its final state $|\,\phi^R_{\rm fin}\rangle$, and $|\,\phi^R_{\rm rel}\rangle$ is its intermediate state when the atom is in the relay state c. The two-photon transition is described by the sequence of the following states for the system atom + photon:

$$|\,a;\phi^R_{\rm in}\rangle \Rightarrow |\,c;\phi^R_{\rm rel}\rangle \Rightarrow |\,b;\phi^R_{\rm fin}\rangle \tag{16}$$

corresponding to the transition amplitude to the lowest order:

$$\langle b;\phi^R_{\rm fin}|\,\bar{U}(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \int_0^{\Delta t} {\rm d}t'\langle b|\bar{\boldsymbol{D}}(t')|c\rangle\cdot\langle\phi^R_{\rm fin}|\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm rel}\rangle \times \int_0^{t'} {\rm d}t''\langle c|\bar{\boldsymbol{D}}(t'')|a\rangle\cdot\langle\phi^R_{\rm rel}|\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')|\,\phi^R_{\rm in}\rangle \tag{17}$$

($\boldsymbol{R}$ is the atom's position). By analogy with relation (B-13) of Chapter XX, we set:

$$\begin{aligned}\langle c|\bar{\boldsymbol{D}}(t)|a\rangle &= \boldsymbol{\mathcal{D}}_{ca}\,e^{i(E_c-E_a)t/\hbar} \qquad \text{with:} \quad \boldsymbol{\mathcal{D}}_{ca}=\mathcal{D}_{ca}\,\mathbf{e}_{ca}\\ \langle b|\bar{\boldsymbol{D}}(t)|c\rangle &= \boldsymbol{\mathcal{D}}_{bc}\,e^{i(E_b-E_c)t/\hbar} \qquad \text{with:} \quad \boldsymbol{\mathcal{D}}_{bc}=\mathcal{D}_{bc}\,\mathbf{e}_{bc}\end{aligned} \tag{18}$$

We then get:

$$\langle b;\phi^R_{\rm fin}\,|\,\bar U(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle = \\ -\frac{\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar^2}\int_0^{\Delta t}{\rm d}t' e^{i(E_b-E_c)t'/\hbar}\langle\phi^R_{\rm fin}\,|\,\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm rel}\rangle \\ \times\int_0^{t'}{\rm d}t'' e^{i(E_c-E_a)t''/\hbar}\langle\phi^R_{\rm rel}\,|\mathbf{e}_{ca}\cdot\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')|\,\phi^R_{\rm in}\rangle \tag{19}$$

Expression (A-29) of Chapter XX for the electric field operator $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t)$ is a sum of modes' contributions, each including the exponential $e^{-i\omega_i t}$ associated with its eigenfrequency. Let us focus on the contribution of mode i in $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')$ and mode j in $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')$. It involves the time integrals:

$$\int_0^{\Delta t}{\rm d}t' e^{i(E_b-E_c)t'/\hbar}e^{-i\omega_i t'}\int_0^{t'}{\rm d}t'' e^{i(E_c-E_a)t''/\hbar}e^{-i\omega_j t''} \tag{20}$$

with an exponent containing:

$$[E_b-E_c-\hbar\omega_i]\,t'+[E_c-E_a-\hbar\omega_j]\,t'' \\ =[E_b-E_a-\hbar\omega_i-\hbar\omega_j]\,t'+[E_c-E_a-\hbar\omega_j]\,(t''-t') \tag{21}$$

reminiscent of result (8) obtained for monochromatic excitation. The computation is then very similar to that of § 1, assuming that the relay state is not half-way between levels a and b, and that the frequency distribution of the incident photons does not include any of the resonance frequencies for the one-photon transitions $a\leftrightarrow c$ and $c\leftrightarrow b$. We make the usual change of variable $\tau=t''-t'$, and, in the integral over τ, we only keep the upper boundary contribution, as we did going from (10) to (13):

$$\int_{-t'}^{0}{\rm d}\tau\; e^{i[E_c-E_a-\hbar\omega_i]\tau/\hbar}\Rightarrow\frac{i\hbar}{[E_a+\hbar\omega_j-E_c]} \tag{22}$$

(we discussed in comment (ii) at the end of § 1 the significance of the lower boundary contribution, and why it is justified to ignore it). In addition, we assume the width of the frequency spectrum of the incident photons to be small compared to the one-photon resonance detuning $E_a+\hbar\omega_i-E_c$; consequently, the denominator of (22) does not vary significantly in relative value, and can be replaced by the constant value $E_a+\hbar\omega_{\rm ex}-E_c$, where $\omega_{\rm ex}/2\pi$ is the central excitation frequency. We note Δ_{ac} the distance from the resonant absorption of a photon in the intermediate state:

$$\Delta_{ac}=\frac{E_a+\hbar\omega_{\rm ex}-E_c}{\hbar} \tag{23}$$

The replacement of the integral over τ by i/Δ_{ac} yields an exponential depending on the variable t', with argument $i\,[E_b-E_a-\hbar\omega_j-\hbar\omega_i]\,t'/\hbar$. Each summation over the modes i and j with the exponential factors $e^{-i\omega_{i,j}t'}$ reconstructs the electric field $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')$, which leads to:

$$\langle b;\phi^R_{\rm fin}\,|\,\bar U(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle=-\frac{i\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar^2\Delta_{ac}} \\ \times\int_0^{\Delta t}{\rm d}t' e^{i\omega_0 t'/\hbar}\langle\phi^R_{\rm fin}\,|\,\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm rel}\rangle\langle\phi^R_{\rm rel}\,|\mathbf{e}_{ca}\cdot\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm in}\rangle \tag{24}$$

where ω_0 was defined as $\omega_0 = (E_b - E_a)/\hbar$. Finally, the summation over all the radiation intermediate states $|\,\phi^R_{\rm rel}\rangle$ yields a closure relation and we obtain:

$$\sum_{|\,\phi^R_{\rm rel}\rangle} \langle b;\phi^R_{\rm fin}\,|\,\bar{U}(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle = -\frac{i\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar^2\Delta_{ac}} \times \int_0^{\Delta t} {\rm d}t' e^{i\omega_0 t'/\hbar}\langle\phi^R_{\rm fin}\,|\,\left[\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]\left[\mathbf{e}_{ca}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]|\,\phi^R_{\rm in}\rangle \tag{25}$$

This result is similar to the probability amplitude for a one-photon transition, written on the second line of (B-14) in Chapter XX, provided we make the substitution:

$$\begin{aligned} &-\mathcal{D}\langle\phi^R_{\rm fin}\,|\,\mathbf{e}_d\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm in}\rangle \\ &\Rightarrow \frac{\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar\Delta_{ac}}\langle\phi^R_{\rm fin}\,|\,\left[\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]\left[\mathbf{e}_{ca}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]|\,\phi^R_{\rm in}\rangle \end{aligned} \tag{26}$$

We then follow the same line of reasoning as for a one-photon transition. In equality (B-17) of Chapter XX, we must now substitute:

$$\begin{aligned} G^N_R(t',t'') \Rightarrow \langle\phi^R_{\rm in}|\left[\mathbf{e}^*_{ca}\cdot\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t')\right]\left[\mathbf{e}^*_{bc}\cdot\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t')\right] \\ \times\left[\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')\right]\left[\mathbf{e}_{ca}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')\right]|\,\phi^R_{\rm in}\rangle \end{aligned} \tag{27}$$

which is then inserted in (B-18) to yield the transition probability. This probability is given by the Fourier transform, at the angular frequency ω_0 of the atomic transition, of a correlation function of the field in the initial state, and which involves four field operators (4-point correlation function). This function is in general different from a product of correlation functions involving two field operators (those determining the absorption probability of a single photon). This means that measurements of two-photon transition probabilities yield access to characteristics of the quantum field that are different from those measured in single photon transitions.

2-b. Probability per unit time when the radiation is in a Fock state

Let us assume now that the radiation is initially in a Fock state such as that described by (B-19) in Chapter XX. In (25), we replace the positive frequency components of the electric fields by their expressions (A-27) of Chapter XX. Only the occupied modes in state $|\,\phi^R_{\rm in}\rangle$ now come into play, since each annihilation operator yields a factor equal to the square root of the mode's initial population; the other modes give a zero result. We then consider two modes, k_1 and k_2, initially occupied in the incident radiation. They yield two contributions (Figure 2) to relation (25): in one of them (term $i = k_1$ and $j = k_2$), the photon k_1 is absorbed first and brings the atom from the ground state to the relay state, then photon k_2 completes the two-photon transition and brings the atom to level b; in the other (term $i = k_2$ and $j = k_1$), the order of the two absorptions is inverted. These two contributions interfere in the probability: once the amplitude modulus is squared, four terms arise from the cross contributions of the two modes (to which we must add two non-crossed contributions $i = j = k_{1,2}$ where only one mode is involved).

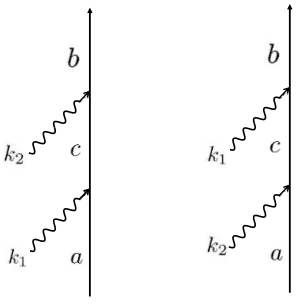


Figure 2: Two diagrams schematizing a two-photon transition with a multimode source where two modes k_1 and k_2 are initially occupied. In the left-hand side diagram, photon k_1 is absorbed first, bringing (in a non-resonant fashion) the atom from the initial sate a to the relay state c; photon k_2 then completes the (resonant) two-photon transition. In the right-hand side diagram, the order of absorption of photons k_1 and k_2 is inverted. These two diagrams describe probability amplitudes that interfere when computing the two-photon transition probability.

The same line of reasoning as in § B-2-a of Chapter XX, and summarized in Figure 1 of that chapter, can be followed here. We assume that the 4-point correlation of the field goes rapidly to zero when $t'' - t'$ is larger that a value $1/\Delta\omega$ that is small compared to Δt. One can then show that the transition probability becomes proportional to Δt, and that the two-photon transition probability per unit time can be written:

$$\begin{aligned} W^{(2)}_{a\to b} &= \frac{P^{(2)}_{a\to b}(\Delta t)}{\Delta t} \\ &= \frac{2\pi}{\hbar^2}\left(\frac{\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar\Delta_{ac}}\right)^2 \frac{1}{4\varepsilon_0^2 L^6}\sum_{i,j}\Big[2\,\langle n_i n_j\rangle - \delta_{ij}\langle n_i(n_i+1)\rangle\Big]\,(\hbar\omega_i)\,(\hbar\omega_j) \\ &\qquad\qquad \times\ \left|(\mathbf{e}_{bc}\cdot\boldsymbol{\varepsilon}_i)\,(\mathbf{e}_{ca}\cdot\boldsymbol{\varepsilon}_j)\right|^2\ \delta(\omega_i+\omega_j-\omega_0) \end{aligned} \tag{28}$$

The delta function at then end of this expression obviously expresses total energy conservation: for the atomic transition to occur, the sum of the energies of the absorbed photons must be equal to the energy of the transition. As expected, the probability includes the photon populations that satisfy this condition. A general property is that, for $i \neq j$ (photons absorbed from two different modes), it is the average value $\langle n_i n_j\rangle$ of the product of the mode populations that appears in the two-photon transition probability, and not the product of the average values which different in general (they are nevertheless equal in the special case of a Fock state of the radiation). For $i = j$ (two photons absorbed from the same mode, as in § 1), it is the average value $\langle n_i\,(n_i-1)\rangle$ that comes into play; this value equals zero if only one photon is present in the mode, as obviously one single photon cannot induce a two-photon transition.

Comment: There also exists 3- ,..., n-photon transitions, corresponding to an energy conservation relation $E_b - E_a = n\hbar\omega_i$. The corresponding transition rates are proportional to field correlation functions of order 6, .., $2n$.

3. Discussion

Even though the transition amplitudes for one- and two-photon absorption processes are similar, the second type of processes has a number of specific features we now discuss.

3-a. Conservation laws

α. Total energy conservation

As we just saw, the function $\delta^{(\Delta t)}(E_b - E_a - 2\hbar\omega_i)$ appearing in (13) expresses the conservation of total energy. When the atom absorbs two photons to go from state a to state b, its gain in energy $E_b - E_a$ equals the sum $2\hbar\omega_i$ of the energies of the two absorbed photons.

β. Total momentum conservation

In the computation leading to the two-photon transition amplitude, the external variables have been ignored. To take them into account, we must assume the atomic center of mass is at point $\boldsymbol{R}$ and keep the exponentials $\exp(i\boldsymbol{k}_i \cdot \boldsymbol{R})$ appearing in the operators $\boldsymbol{E}^{(+)}(\boldsymbol{R})$ in the two interaction Hamiltonians H_I; as these two exponentials are multiplied by each other, they yield the operator $\exp(2i\boldsymbol{k}_i \cdot \boldsymbol{R})$. We must also include in the initial and final states the quantum numbers $\boldsymbol{K}_{\text{in}}$ and $\boldsymbol{K}_{\text{fin}}$ characterizing the initial $\hbar\boldsymbol{K}_{\text{in}}$ and final $\hbar\boldsymbol{K}_{\text{fin}}$ momenta of the center of mass. We then get in the transition amplitude an additional term:

$$\langle \boldsymbol{K}_{\text{fin}} | \exp(2i\,\boldsymbol{k}_i \cdot \boldsymbol{R}) | \boldsymbol{K}_{\text{in}} \rangle = \delta(\boldsymbol{K}_{\text{fin}} - \boldsymbol{K}_{\text{in}} - 2\boldsymbol{k}_i) \tag{29}$$

which shows that the atom's momentum increases by $2\hbar\boldsymbol{k}_i$ when it absorbs two photons.

Comment:

Imagine the atom is excited by two light beams 1 and 2, having the same frequency $\omega_i/2\pi$ but propagating in opposite directions. The previous computations must then be generalized to the case where the two photons absorbed by the atom belong, one to beam 1, and the other to beam 2. The momenta $+\hbar\boldsymbol{k}_i$ and $-\hbar\boldsymbol{k}_i$ of these two photons are then opposite and the total momentum gained by the atom during the transition is zero. As the Doppler effect and the recoil effect are linked to the variation of the atomic momentum during the transition (see Complement A_{XIX}), it follows that the two-photon absorption line does not present any Doppler broadening nor any recoil shift. Such a situation presents many advantages for high resolution spectroscopy, and is used for example in the study of the two-photon transition between the $1s$ and $2s$ states of the hydrogen atom.

γ. Conservation of total angular momentum and parity

Expression (14) appearing in the two-photon absorption amplitude is a product of two matrix elements of a component of the atomic electric dipole – which is a vector operator – and an energy denominator. In a rotation of the atom, this expression will be transformed as the product of two vectors, since the energy denominator is rotation invariant. Vectors are irreducible tensor operators (Complement G_X, exercise 8) of order $K = 1$. Consequently, expression (14) may be expanded[6] as a sum of components with total angular momentum $F = 0, 1, 2$. Using the Wigner-Eckart theorem (Complement E_X), we can show that the two-photon absorption amplitude between two levels with quantum numbers F, M and F', M' (where M and M' are the components of $\boldsymbol{F}$ on the Oz axis) is different from zero only if:

$$\begin{aligned} F - F' &= \pm 2, \pm 1, 0 \\ M - M' &= \pm 2, \pm 1, 0 \end{aligned} \tag{30}$$

In addition, the electric dipole operator appearing in (14) is an odd operator, as it is proportional to the electron position operator. Consequently, the initial and final states of the two-photon transition must have the same parity, and a parity inverse to that of the relay state.

These selection rules can be applied to the $1s \leftrightarrow 2s$ transition of the hydrogen atom, that occurs between two states having the same parity and a total angular momentum difference equal to ± 1 at most (the $1/2$ spins of the electron and the proton are taken into account). For electric dipole transitions, a two-photon transition $1s \leftrightarrow 2s$ is allowed, whereas it is forbidden for a one-photon transition.

3-b. Case where the relay state becomes resonant for one-photon absorption

In the denominator of expression (14), we have the quantity:

$$\hbar\Delta_{ac} = E_a + \hbar\omega_i - E_c \tag{31}$$

which is the difference between the energy of the atomic state a, increased by $\hbar\omega_i$, and the energy of the relay sate c. If Δ_{ac} goes to zero, we get a divergence and the expressions we have obtained become meaningless. In the computation, we did explicitly assume that the intermediate level was not resonant for a one-photon absorption, so that this divergence should not occur. Let us examine, however, what would be involved if Δ_{ac} were to go to zero. As the resonance condition for the two-photon transition is written $E_b = E_a + 2\hbar\omega_i$, the condition $\Delta_{ac} = 0$ means that the atomic relay level[7] c is exactly

[6] The product of two irreducible tensor operators K and K' can be decomposed as the product of two kets with angular momentum K and K', hence involving Clebsch-Gordan coefficients. This means that, according to the general results of Chapter X on the addition of angular momenta, a product of irreducible tensor operators can be decomposed as the sum of other irreducible tensor operators of order K'', where K'' varies between $|K - K'|$ and $K + K'$. In the particular case where $K = K' = 1$, we get three possible values $F = 0, 1, 2$.

[7] We assume here that the relay state c is discrete. If it belongs to a continuum, the sum over this relay state c in (14) becomes an integral over E_c. An adiabatic branching calculation then introduces a fraction $1/(E_a + \hbar\omega_i + i\eta - E_c)$ with $\eta \to 0$, which can then be expressed in terms of $\delta(E_a + \hbar\omega_i - E_c)$ and $\mathcal{P}(1/(E_a + \hbar\omega_i - E_c))$, where $\mathcal{P}$ is the Cauchy principal value. This calculation yields, after integration over E_c, functions of $E_{\text{in}} = E_a + 2\hbar\omega_i$ which have no reason to diverge in the vicinity of $E_{\text{in}} \simeq E_{\text{fin}}$.

half-way in energy between a and b, the initial and final atomic levels. This means that, starting from level a, a resonance would occur for both a two-photon and a one-photon process.

Is it possible to study the case where Δ_{ac} is zero or very small, while avoiding the divergence of the two-photon absorption? A method to overcome this difficulty is to note that state c has a finite lifetime τ_c, because of spontaneous emission of photons from that state. As we did before in § E-2 of Chapter XX, when studying the resonant scattering of a photon by an atom and where similar divergences appeared, we can show that it is legitimate to replace the energy E_c of state c by $E_c - i\hbar(\Gamma_c/2)$, where $\Gamma_c = 1/\tau_c$ is the natural width of level c. The denominator of the transition amplitude no longer goes to zero and the divergence disappears. The transition amplitude still varies significantly over an interval of width Γ_c when E_c varies around $E_a + \hbar\omega_i$.

Replacing E_c by $E_c - i(\hbar\Gamma_c/2)$ leads to valid results only if the matrix elements $\langle c; n_i - 1|H_I|a; n_i\rangle$ and $\langle b; n_i - 2|H_I|c; n_i - 1\rangle$, characterizing the coupling of the field with the atom for the $c \leftrightarrow a$ and $b \leftrightarrow c$ transitions, are small compared to $\hbar\Gamma_c$. If this is not true, we cannot limit the computation to the lowest field order. We must then diagonalize the Hamiltonian of the global system atom + field within the subspace $\mathcal{E}$ of the states which, in the absence of coupling, are very close to each other[8]. When no relay state is resonant, the subspace $\mathcal{E}$ is two-dimensional; it is spanned by the two states $|a; n_i\rangle$ and $|b; n_i - 2\rangle$. When one relay state c becomes resonant, we must include the state $|c; n_i - 1\rangle$ in the subspace $\mathcal{E}$ – which then becomes three-dimensional. To study the dynamics of the system, we must diagonalize the matrix:

$$\begin{pmatrix} E_a + n_i\hbar\omega_i & \langle a; n_i|H_I|c; n_i - 1\rangle & 0 \\ \langle c; n_i - 1|H_I|a; n_i\rangle & E_c + (n_i - 1)\hbar\omega_i & \langle c; n_i - 1|H_I|b; n_i - 2\rangle \\ 0 & \langle b; n_i - 2|H_I|c; n_i - 1\rangle & E_b + (n_i - 2)\hbar\omega_i \end{pmatrix} \tag{32}$$

This general treatment allows taking into account simultaneously the one- and two-photon transitions.

Concluding this complement, let us emphasize that the two-photon transitions involve a physical process different from the mere succession of two one-photon absorptions. We stressed in the discussion of § 1, and in particular in its two comments, the difference between populating the final state, which is cumulative in time and conserves the energy, and a transit through an intermediate relay state, which can only last a very short time $\Delta\tau$, limited by the non-conservation of energy. It is also noteworthy that the two-photon transition amplitude can take a form very similar to that of a one-photon transition; the only major change is the replacement of the matrix element to first order in the interaction, by a second order matrix element, divided by an energy defect factor in the relay state. These concepts can be generalized to higher order processes: similar techniques can be used to evaluate three-, four-, etc.. photon transition amplitudes.

[8]Such a description of the atom + radiation interactions is called the "dressed atom method" (see for example Chapter VI of reference [21]). In Complement C_{XX}, this method is applied to the problem of a two-level atom interacting with a strong field. The eigenstates of the total Hamiltonian restricted to the subspace $\mathcal{E}$ are called the "dressed states".

Complement B_{XX}

Photoionization

1 Brief review of the photoelectric effect 2110
1-a Interpretation in terms of photons 2110
1-b Photoionization of an atom 2111
2 Computation of photoionization rates 2112
2-a A single atom in monochromatic radiation 2112
2-b Stationary non-monochromatic radiation 2113
2-c Non-stationary and non-monochromatic radiation 2116
2-d Correlations between photoionization rates of two detector atoms 2117
3 Is a quantum treatment of radiation necessary to describe photoionization? 2118
3-a Experiments with a single photodetector atom 2118
3-b Experiments with two photodetector atoms 2119
4 Two-photon photoionization 2123
4-a Differences with the one-photon photoionization 2123
4-b Photoionization rate 2124
4-c Importance of fluctuations in the radiation intensity 2125
5 Tunnel ionization by intense laser fields 2126

All the atomic processes of absorption, emission or scattering of photons studied in Chapter XX involved transitions between two discrete states of the atom. In addition to discrete levels, atoms also have continuums of energy levels. The most directly accessible one is the simple ionization continuum, which corresponds to the loss of a single electron by the atom (ionization). This continuum starts at an energy threshold E_I above the ground state energy, and extends over all energies larger than this threshold. This energy E_I is called the "ionization energy". The aim of this complement is to study the "photoionization" process where incident radiation takes the atom from its ground state to a state belonging to the ionization continuum.

Once the atom's electron has reached the ionization continuum, it can travel an arbitrary distance from the remaining ion; it has been ejected from the atom by the incident radiation. Such a process is reminiscent of the photoelectric effect where radiation ejects an electron from a metal. This is why we shall review in § 1 a few properties of the photoelectric effect to underline its analogies with photoionization.

We shall then use quantum theory to compute, in § 2, the probability per unit time for the incident radiation to photoionize an atom. We shall assume that the incident radiation spectrum is entirely above the ionization threshold, so that no resonant absorption can bring the atom to a discrete excited state. Since the emitted electron can be amplified in a photomultiplier, the atom can play the role of a photodetector. In the case where only one photodetector D is used (§ 2-a), the computations are very similar to those exposed in Chapter XX; the only differences arise from the continuous character

of the final atomic state. Another interesting situation occurs when two detectors D1 and D2 are placed in the radiation field at points $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ (§ 2-d) and when we focus on the correlations between their signals. For example, we shall compute the probability per unit time to observe a photoionization at $\boldsymbol{R}_1$ at time t_1 and another one at $\boldsymbol{R}_2$ at time t_2.

One may wonder if a quantized radiation theory is needed to quantitatively account for the photoionization processes. Could a semiclassical theory suffice to describe the ionization of one or several quantized atoms by a classical field? In other words, can the photoelectric effect be explained "without photons" [45]? This question will be discussed in § 3.

The atom can also be photoionized by the absorption of a number of photons n, larger than one. These processes are called "*multiphoton ionization*" and play an important role in experiments using high intensity laser sources. In § 4, we shall give an idea of how to compute the rates of those processes for $n = 2$. We shall then briefly mention in § 5 another mechanism for atoms' ionization, based on a tunnel effect, and occurring when the incident radiation electric field becomes of the order of the Coulomb field between an atom's electron and the nucleus.

1. Brief review of the photoelectric effect

In 1905, Albert Einstein [43] introduced for the first time in physics the concept of "light quanta", which we now call photons. Considering the great analogy between certain statistical properties of black body radiation and those of an ideal gas of particles, Einstein proposed the idea that radiation was in fact composed of discrete quanta, each having an energy $h\nu$. In view of the successes of the wave theory of light, this return to a particle description seemed totally unrealistic for most physicists at the time. Energy quantization had indeed been introduced a few years earlier by Max Planck to account for the spectral distribution of black body radiation, but it was the exchanges of energy between matter and radiation that were quantized, not the radiation itself.

1-a. Interpretation in terms of photons

In that same 1905 article, Einstein used the concept of light quanta to give a new description of the photoelectric effect. In this process, an electron is ejected from a metal irradiated by light. Einstein postulated that the energy $h\nu$ of a light quantum from the incident beam was absorbed by an electron in the metal, hence allowing it to escape from the metal. This escape requires an energy at least equal to that of the binding energy E_L of the electron in the metal. The frequency ν of the light beam must therefore be larger than a threshold value given by $h\nu = E_L$. If $h\nu < E_L$, no electron can be ejected. If $h\nu > E_L$, the energy surplus $h\nu - E_L$ provides the electron with a kinetic energy $E_{\text{kin}} = mv^2/2$. This interpretation leads to Einstein's equation:

$$\frac{1}{2}mv^2 = h\nu - E_L \quad \text{if} \quad h\nu > E_L \tag{1}$$

giving the kinetic energy of the ejected electron as a function of ν. It means, in particular, that the kinetic energy of each electron depends only on the frequency of the light beam,

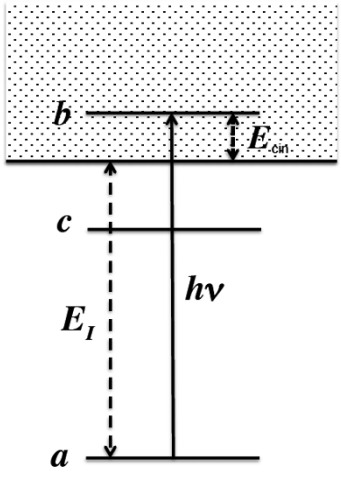


Figure 1: Photoionization of an atom. State a is the ground state, state c one of the discrete states. The continuum of states belonging to the continuous part of the spectrum (ionization continuum) starts at an energy E_I above the ground state. Energy E_I is called the "ionization energy". As the atom in state a absorbs a photon with frequency ν such that $h\nu > E_I$, it goes into a state b, which is part of the ionization continuum. The electron is ejected, and its kinetic energy $E_{\rm kin}$ (when it is far enough from the ion formed as it left the atom) is equal to the difference between $h\nu$ and E_I.

and not on its intensity[1](which, on the other hand, determines the number of ejected electrons per unit time). Equation (1) also tells us that if ν is varied and one plots the variation of $mv^2/2$ as a function of ν, one should get a straight half-line with slope h, starting from the abscissa axis at point E_L/h. All these predictions generated a certain skepticism and it was not until several years later (1913) that an experimental confirmation of the predictions of equation (1) was obtained by the work of R. Millikan and H. Fletcher on the photoelectric effect [44].

1-b. Photoionization of an atom

Figure 1 represents the ground state a of an atom, one of the states c from the discrete part of the spectrum, and the continuum part of the spectrum which starts at a distance E_I above the ground state. The origin of the energies is often chosen at the beginning of the continuum, and hence the discrete states have a negative energy. When it is in the positive energy states, belonging to the continuum and called "scattering states", the electron is no longer bound to the nucleus, although it is still attracted to it.

Consider an atom in the ground state a. A photon with energy $h\nu$ has different ways to bring energy to the atomic electron. If $h\nu < E_I$, this photon can be absorbed only if ν coincides with the frequency $\nu_{ca} = (E_c - E_a)/h$ of a transition between state a

[1]A classical theory would tend to predict that the higher the light intensity, the more energy could be furnished to the electron, thereby increasing its acceleration.

and state c belonging to the discrete spectrum. This is the absorption process between two discrete states already studied in Chapter XX. On the other hand, if $h\nu > E_I$, the atom can always absorb the photon and end up in a state within the continuum. The electron is no longer bound and can move away from the remaining ion formed once the electron has left. When the electron is far enough from the ion for their Coulomb interaction energy to be neglected, its kinetic energy is given by:

$$\frac{1}{2}mv^2 = h\nu - E_I \quad \text{if} \quad h\nu > E_I \tag{2}$$

This is the photoionization process of an atom. Equation (2) is the generalization for an atom of equation (1) introduced by Einstein for the photoelectric effect in a metal.

2. Computation of photoionization rates

Let us now see how to adapt the computations of Chapter XX to the calculation of a photoionization rate.

2-a. A single atom in monochromatic radiation

We start from expression (B-7) of Chapter XX for the probability that the system, leaving at $t = 0$ the initial state $|\psi_{\text{in}}\rangle = |\, a; n_i\rangle$ (atom in state a in the presence of n_i photons ω_i), ends up at time Δt in the final state $|\psi_{\text{fin}}\rangle = |\, b; n_i - 1\rangle$ (atom in state b with an energy $\hbar\omega_0$ above a, and one photon less in mode i) :

$$|\langle\psi_{\text{fin}}|\bar{U}(\Delta t, 0)|\psi_{\text{in}}\rangle|^2 = \frac{1}{\hbar^2}\,|\langle\psi_{\text{fin}}|H_I|\psi_{\text{in}}\rangle|^2\,\frac{4\sin^2\left[(\omega_0-\omega_i)\Delta t/2\right]}{\left[(\omega_0-\omega_i)/2\right]^2} \tag{3}$$

In Chapter XX, we used this expression for studying the case where states a and b are both discrete states. It is nevertheless still valid when b belongs to a continuum; its interpretation, however, is different. Whereas the probability of finding the atom in a discrete final state makes sense, from a physical point of viewn, when dealing with a continuum we must compute the probability of finding the atom within a non-zero energy interval. We must then sum probability (3) over states b.

As b varies, $\omega_0 = (E_b - E_a)/\hbar$ varies, and so does the matrix element of H_I. However, for large enough Δt, the variation of the H_I matrix element is much slower than that of the ratio in the right-hand side of (3). This ratio is the square of a diffraction function, whose maximum equals Δt^2 for $\omega_0 = \omega_i$ and whose width is of the order of $1/\Delta t$. The area under this function is thus of the order of $\Delta t^2 \times (1/\Delta t) = \Delta t$. Compared to functions of ω_0 with slow variations over an interval of the order of $1/\Delta t$, this function behaves, within a proportionality factor, as the product $\Delta t\,\delta(\omega_0 - \omega_i)$. It follows that the sum over b of (3) is proportional to Δt, meaning we can define a probability per unit time for the atom to reach the continuum, i.e. a photoionization rate.

The proportionality factor between a diffraction function and a delta function is given by relation (11) in Appendix II, which is written:

$$\lim_{\Delta t\to\infty}\frac{\sin^2\left[(\omega_0-\omega_i)\Delta t/2\right]}{\left[(\omega_0-\omega_i)/2\right]^2} = \pi\ \Delta t\ \delta(\frac{\omega_0-\omega_i}{2}) = 2\pi\hbar\ \Delta t\ \delta(E_b - E_a - \hbar\omega_i) \tag{4}$$

Inserting (4) into (3), summing over b, and dividing by Δt, finally yields the photoionization rate:

$$\frac{1}{\Delta t}\sum_b |\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle|^2 = \frac{2\pi}{\hbar}\sum_b |\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle|^2\delta(E_b - E_a - \hbar\omega_i)$$
$$= \frac{2\pi}{\hbar}\underbrace{|\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle|^2}_{E_b=E_a+\hbar\omega_i}\ \rho(E_a+\hbar\omega_i) \tag{5}$$

where $\rho(E_a+\hbar\omega_i)$ is the density of states in the continuum around the energy $E_a+\hbar\omega_i$. This expression is just a consequence of the Fermi golden rule (Chapter XIII, § C-3-b) applied to the coupling between the discrete level $|\,a;n_i\rangle$ and the continuum $|\,b;n_i-1\rangle$. It is also reminiscent of expression (C-37) of Chapter XIII which yields the transition probability per unit time between a discrete atomic state and a continuum, with an excitation induced by a classical wave described by a time-dependent sinusoidal function. This point will be discussed further in § 3.

2-b. Stationary non-monochromatic radiation

We now consider a single atom interacting with non-monochromatic radiation, described by a spectral distribution $I(\omega)$. We first assume that the field statistical properties are time-invariant. We shall consider the case of non-stationary radiation in § 2-c.

α. *Field and atomic dipole correlation functions*

In § D-1 of Chapter XX, we obtained the expression for the probability $\sum_b P^{\rm abs}_{a\to b}(\Delta t)$ for the atom to go, through the absorption of a photon, from state a to any state b different from a, after a time Δt. This probability is given by relation (D-4) of that chapter as a double integral of a sum of products of field and atomic dipole correlation functions.

We first examine the correlation functions of the atomic dipole. As in relation (B-13) of Chapter XX, we write the matrix elements of this dipole, in the Heisenberg picture (with respect to the Hamiltonian of the free atom), as:

$$\langle b|\bar{\boldsymbol{D}}(t)|a\rangle = \langle b|\boldsymbol{D}|a\rangle e^{i\omega_{ba}t} = \boldsymbol{\mathcal{D}}_{ba}e^{i\omega_{ba}t} \tag{6}$$

with:

$$\boldsymbol{\mathcal{D}}_{ba} = \mathcal{D}_{ba}\ \boldsymbol{e}_d \tag{7a}$$

where $\boldsymbol{e}_d$ is the unit vector parallel to the vector $\langle b|\boldsymbol{D}|a\rangle$. Since $\boldsymbol{e}_d^*\cdot\boldsymbol{e}_d = 1$, we have:

$$\mathcal{D}_{ba} = \boldsymbol{e}_d^*\cdot\boldsymbol{\mathcal{D}}_{ba} = \langle b|\boldsymbol{e}_d^*\cdot\boldsymbol{D}|a\rangle \tag{7b}$$

For the sake of simplicity, we shall assume the unit vector $\boldsymbol{e}_d$ to have the same direction for all the states b.

Comment:

This vector $\boldsymbol{e}_d$ is, a priori, not the same for all the states b related to a by matrix elements of $\boldsymbol{D}$. The direction of that vector actually depends on the rotation symmetry properties

of the two states[2]. One can sort the different b states by categories having the same symmetry (hence the same direction for $\boldsymbol{e}_d$) and for which the computations presented hereafter are valid. One must then add all the ionization probabilities calculated for each category.

Since the field always appears in a scalar product with $\boldsymbol{D}$, assuming that $\boldsymbol{e}_d$ has the same direction for all states b implies that only the scalar products $E_d^{(+)}$ (and $E_{d^*}^{(-)}$) of the fields with $\boldsymbol{e}_d$ (and with $\boldsymbol{e}_d^*$) appear in the correlation functions of the field.

The calculation is then very similar to that of § D-1 of Chapter XX, and we obtain:

$$\sum_b P_{a\to b}^{\text{abs}}(\Delta t) = \frac{1}{\hbar^2}\int_0^{\Delta t} \mathrm{d}t' \int_0^{\Delta t} \mathrm{d}t''\; G_A(t''-t')\; G_R^N(t''-t') \tag{8}$$

where:

$$G_A(t''-t') = \sum_b |\mathcal{D}_{ba}|^2 e^{i\omega_{ba}(t''-t')} \tag{9}$$

$$\begin{aligned} G_R^N(t''-t') &= \langle \phi_{\text{in}}^R | \left[\boldsymbol{e}_d^* \cdot \bar{\boldsymbol{E}}^{(-)}(\boldsymbol{R},t')\right] \left[\boldsymbol{e}_d \cdot \bar{\boldsymbol{E}}^{(+)}(\boldsymbol{R},t'')\right] |\phi_{\text{in}}^R\rangle \\ &= \langle \phi_{\text{in}}^R | \bar{E}_{d^*}^{(-)}(\boldsymbol{R},t') \bar{E}_d^{(+)}(\boldsymbol{R},t'') | \phi_{\text{in}}^R\rangle \end{aligned} \tag{10}$$

In this last equality, $|\phi_{\text{in}}^R\rangle$ is the radiation initial state. As this state is stationary, its properties are invariant under time translation; consequently, the correlation function depends only on the difference $t''-t'$.

The atomic correlation function (9) can also be rewritten as:

$$G_A(t''-t') = \int \mathrm{d}\omega\, \tilde{G}_A(\omega) e^{i\omega(t''-t')} \tag{11}$$

where:

$$\tilde{G}_A(\omega) = \sum_b |\mathcal{D}_{ba}|^2 \delta(\omega - \omega_{ba}) \tag{12}$$

The quantity $\tilde{G}_A(\omega)$ represents the spectral sensitivity of the "photodetector atom", that is the variation with ω of the transition intensities from the ground state a to a level b in the ionization continuum, at an energy $\hbar\omega$ above a. We shall assume here that the width $\Delta\omega_A$ of the function $\tilde{G}_A(\omega)$ is much larger than the bandwidth $\Delta\omega_R$ of the incident radiation.

$$\Delta\omega_A \gg \Delta\omega_R \tag{13}$$

[2]The ground state a is an eigenvector of the angular momentum component along the quantization axis Oz, with eigenvalue $m_a\hbar$; for the state b, the eigenvalue is $m_b\hbar$. The a to b transition thus corresponds to a variation $\Delta m = m_b - m_a$. If $\Delta m = 0$, symmetry arguments show that $\boldsymbol{e}_d$ is parallel to Oz; if $\Delta m = \pm 1$, $\boldsymbol{e}_d$ is in the plane perpendicular to Oz: $\boldsymbol{e}_d = (\boldsymbol{e}_x \mp i\boldsymbol{e}_y)/\sqrt{2}$; we note that the complex conjugate of this vector appears in the matrix element (7b).

Such a condition defines what we shall call a "broadband photodetector".

The field correlation function (9) has already been calculated in § B-2-a-β of Chapter XX assuming the radiation initial state is a Fock state $|n_1, ..., n_i, ...\rangle$, or a statistical mixture of such states with weights $p(n_1, ..., n_i, ...)$ – see equations (B-20) and (B-21) of Chapter XX. For the problem we are studying now, i.e. stationary non-monochromatic radiation, we can use the same assumption for the radiation initial state.

β. Photoionization rate

To transform equation (8), it is useful to study in more detail the $t''-t'$ dependence of $G_R^N(t''-t')$. We assume that the spectral distribution is centered around a non-zero value $\omega_{\rm ex}$, and that this distribution is entirely above the ionization threshold. Since the field $\bar{E}_d^{(+)}$ varies in $\mathrm{e}^{-i\omega t}$, and the field $\bar{E}_d^{(-)}$ in $\mathrm{e}^{i\omega t}$, we can then write:

$$G_R^N(t''-t') = e^{i\omega_{\rm ex}(t'-t'')} C_R^N(t''-t') \tag{14}$$

where $C_R^N(t''-t')$ is an "envelope" function whose Fourier transform is a function centered at $\omega = 0$ and of width $\Delta\omega_R$. This envelope function varies very slowly over time intervals short[3] compared to $1/\Delta\omega_R$. For $t'' = t'$, i.e. for $t''-t' = 0$, equation (14) leads to:

$$\begin{aligned} C_R^N(0) &= G_R^N(0) \\ &= \langle \phi_{\rm in}^R | \bar{E}_{d^*}^{(-)}(\boldsymbol{R}, t') \bar{E}_d^{(+)}(\boldsymbol{R}, t') | \phi_{\rm in}^R \rangle = I \end{aligned} \tag{15}$$

where I is the radiation intensity (which is time-independent since the radiation state is supposed to be stationary). We shall see in the next section how to generalize our results to non-stationary radiation.

Let us go back to the double integral of (8) and assume that the integration interval Δt satisfies the condition $\Delta t \gg \tau_{CA}$ where $\tau_{CA} = 1/\Delta\omega_A$ is the detector correlation time. In the plane t', t'', the function to be integrated in (8) is different from zero only in a band along the first bisector (Figure 1 of Chapter XX), of width τ_{CA} very narrow compared to Δt. If we change the integral variables t' and t'' to the variables t' and $\tau = t''-t'$, we can neglect the variation of $C_R^N(\tau)$ since, according to (13), $\tau_{CA} \ll 1/\Delta\omega_R$, and use (15) to rewrite (14) as:

$$\begin{aligned} G_R^N(t''-t') &= e^{i\omega_{\rm ex}(t'-t'')} C_R^N(t''-t') \\ &\simeq e^{i\omega_{\rm ex}(t'-t'')} C_R^N(0) = I e^{i\omega_{\rm ex}(t'-t'')} \end{aligned} \tag{16}$$

The double integral of (8) is easily performed with the new variables t' and τ. Using expression (3) for $G_R^N(t''-t')$, the integral over t' of a function that no longer depends on this variable introduces a simple factor Δt. We are then left with the integral over τ which leads to:

$$\sum_b P_{a\to b}^{\rm abs}(\Delta t) = \frac{1}{\hbar^2}\, I\, \Delta t \underbrace{\int_{-\infty}^{+\infty} \mathrm{d}\tau\; e^{-i\omega_{\rm ex}\tau} G_A(\tau)}_{=2\pi \bar{G}_A(\omega_{\rm ex})} \tag{17}$$

[3]This is not the case for $G_R^N(t''-t')$, because of the exponential $e^{i\omega_{\rm ex}(t'-t'')}$ that varies a lot over time intervals of the order of $1/\Delta\omega_R$, since $\omega_{\rm ex} \gg \Delta\omega_R$.

As this probability is proportional to Δt, we can define a photoionization rate:

$$W_{\text{phot.}} = \frac{1}{\Delta t} \sum_b P^{\text{abs}}_{a\to b}(\Delta t) = \frac{2\pi}{\hbar^2} I\, \bar{G}_A(\omega_{\text{ex}}) \tag{18}$$

This rate is proportional to the incident intensity I and to the spectral sensitivity $\tilde{G}_A(\omega_{\text{ex}})$ of the photodetector, evaluated at the radiation central frequency ω_{ex}.

2-c. Non-stationary and non-monochromatic radiation

For non-stationary radiation, the initial radiation state is no longer a Fock state or a statistical mixture of Fock states; it is rather a linear superposition of such states, creating wave packets such as those described in Complement D_{XX}. The radiation correlation function $G^N_R(t', t'')$ is no longer a function of the single variable $t'' - t'$, but depends on both t' and t''. One can still assume that the frequencies appearing in $G^N_R(t', t'')$ are centered around ω_{ex} in an interval of width $\Delta\omega_R$, which permits generalizing expressions (14) and (15) to:

$$G^N_R(t', t'') = e^{i\omega_{\text{ex}}(t'-t'')} C^N_R(t', t'') \tag{19}$$

and[4]:

$$G^N_R(t', t') = \langle \phi^R_{\text{in}} | E^{(-)}_{d^*}(\boldsymbol{R}, t') E^{(+)}_d(\boldsymbol{R}, t') | \phi^R_{\text{in}} \rangle = C^N_R(t', t') = I(\boldsymbol{R}, t') \tag{20}$$

where $I(\boldsymbol{R}, t')$ is the intensity at point $\boldsymbol{R}$ and at time t'.

Using the expansions in a_i and $a_i^\dagger$ for the field operators appearing in (10), we get an expression for $G^N_R(t', t'')$ that generalizes equation (B-20) of Chapter XX to non-stationary fields:

$$G^N_R(t', t'') = \sum_i \sum_j \frac{\hbar\sqrt{\omega_i \omega_j}}{2\varepsilon_0 L^3} (\boldsymbol{e}^*_d \cdot \varepsilon^*_j)(\boldsymbol{e}_d \cdot \varepsilon_i) \langle \phi^R_{\text{in}} | a^\dagger_j a_i | \phi^R_{\text{in}} \rangle\, e^{-i(\boldsymbol{k}_j \cdot \boldsymbol{R} - \omega_j t')} e^{i(\boldsymbol{k}_i \cdot \boldsymbol{R} - \omega_i t'')} \tag{21}$$

For fixed values of $\boldsymbol{k}_j$ and ω_j, the summation over i of (21) represents a wave packet of central frequency ω_{ex}, and whose envelope passes by a point $\boldsymbol{R}$ in a time interval of the order of $1/\Delta\omega_R$. If t' varies over a time interval $\Delta t \ll 1/\Delta\omega_R$, the variation of the envelope can be neglected. Similar conclusions are valid for a summation over j of relation (21), with fixed values for $\boldsymbol{k}_i$ and ω_i.

Let us now go back to the double integral in (8). As the phenomenon now depends on t, the integration interval will be taken between t and $t + \Delta t$ (instead of between 0 and Δt). We assume that Δt satisfies the condition :

$$\frac{1}{\Delta\omega_A} \ll \Delta t \ll \frac{1}{\Delta\omega_R} \tag{22}$$

which is possible when (13) is taken into account. Since $\Delta t \ll 1/\Delta\omega_R$, we can neglect in (19) the variation of $C^N_R(t', t'')$ when t' and t'' vary in the integration domain; we therefore replace $C^N_R(t', t'')$ by:

$$C^N_R(t, t) = \langle \phi^R_{\text{in}} | E^{(-)}_{d^*}(\boldsymbol{R}, t) E^{(+)}_d(\boldsymbol{R}, t) | \phi^R_{\text{in}} \rangle = I(\boldsymbol{R}, t) \tag{23}$$

[4]From now on, we simplify the notation by omitting the bar over operators in the Heisenberg picture, since the explicit time dependence is sufficient to indicate this point of viewn.

Using this equality in (19) we get, in the integration interval between t and $t + \Delta t$:

$$G_R^N(t', t'') \simeq e^{i\omega_{\rm ex}(t'-t'')} I(\boldsymbol{R}, t) \tag{24}$$

The computations are then quite similar to those carried out above for a stationary field: the integral over t' leads to a Δt term; the integral over $\tau = t'' - t'$ is equal to:

$$\int_{-\infty}^{+\infty} {\rm d}\tau \; e^{-i\omega_{\rm ex}\tau} G_A(\tau) = 2\pi \tilde{G}_A(\omega_{\rm ex}) \tag{25}$$

We call $W_I(\boldsymbol{R}, t)\Delta t$ the probability that a photodetector atom, placed at $\boldsymbol{R}$, will undergo a photoionization between times t and $t + \Delta t$. We get the result:

$$W_I(\boldsymbol{R}, t) = s\langle \phi_{\rm in}^R | E_{d^*}^{(-)}(\boldsymbol{R}, t) E_d^{(+)}(\boldsymbol{R}, t) | \phi_{\rm in}^R \rangle \tag{26}$$

where $s = 2\pi \bar{G}_A(\omega_{\rm ex})/\hbar^2$ is a factor characterizing the photodetector sensitivity at the radiation central frequency $\omega_{\rm ex}$. The atomic photoionization rate is thus a signal that constantly follows the time variations of the incident radiation intensity, written in (20).

2-d. Correlations between photoionization rates of two detector atoms

The previous computations can be generalized to analyze other experiments where two detector atoms are placed at $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$, and where we study correlations between the photoionizations observed on those two atoms at times t_1 and t_2. More precisely, let us call $W_{II}(\boldsymbol{R}_2, t_2; \boldsymbol{R}_1, t_1)\Delta t_1 \Delta t_2$ the probability to detect a photoionization at $\boldsymbol{R}_1$ between t_1 and $t_1 + \Delta t_1$ and another one at $\boldsymbol{R}_2$ between t_2 and $t_2 + \Delta t_2$. Computations very similar to those performed above, and which will not be explicited here (for more details, see Complement $\rm A_{II}$ of [21]), lead to:

$$W_{II}(\boldsymbol{R}_2, t_2; \boldsymbol{R}_1, t_1) = \\ s^2 \langle \phi_{\rm in}^R | E_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1, t_1) E_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_2, t_2) E_{\boldsymbol{d}}^{(+)}(\boldsymbol{R}_2, t_2) E_{\boldsymbol{d}}^{(+)}(\boldsymbol{R}_1, t_1) | \phi_{\rm in}^R \rangle \tag{27}$$

It is easy to understand why two operators $\boldsymbol{E}_\perp^{(+)}$ preceded by two operators $\boldsymbol{E}_\perp^{(-)}$ appear in (27). The double photoionization rate W_{II} is computed from a probability that is the modulus squared of a probability amplitude for a photon to be absorbed at $\boldsymbol{R}_1, t_1$ and another one at $\boldsymbol{R}_2, t_2$. This amplitude must contain a product of two operators $\boldsymbol{E}_\perp^{(+)}$. Its conjugate must contain two operators $\boldsymbol{E}_\perp^{(-)}$ arranged in inverse order. We therefore should find in (27) two operators $\boldsymbol{E}_\perp^{(-)}$ followed by two operators $\boldsymbol{E}_\perp^{(+)}$ with different orders of $\boldsymbol{R}_1, t_1$ and $\boldsymbol{R}_2, t_2$.

There is a great analogy between the simple and double photoionization rates W_I and W_{II} given by equations (25) and (27) and the correlation functions G_1 and G_2 studied in Chapter XVI. The functions G_1 and G_2 give the probability densities of finding a particle at $\boldsymbol{r}_1, t_1$ for G_1, or a particle at $\boldsymbol{r}_1, t_1$ and another one at $\boldsymbol{r}_2, t_2$ for G_2. Note, however, that G_1 and G_2 give the probability of finding one or two particles at specific points, whereas we are now dealing with the probability of photoionization of atoms placed at specific points. The field operators appearing in (26) and (27) are the positive or negative frequency components of the electric field, since these are the operators describing the emission or absorption of photons.

3. Is a quantum treatment of radiation necessary to describe photoionization?

In a semiclassical treatment of photoionization, the radiation field is described as a classical field, while the atom follows a quantum treatment. The atom-field coupling is then a time-dependent perturbation that can induce transitions between a discrete atomic state, such as the ground sate a, and a state b, part of the ionization continuum. Does such a treatment yield the same results as those obtained in a quantum treatment? We are going to show that, while this is often the case, this is not always true.

3-a. Experiments with a single photodetector atom

In the simple case of an oscillating monochromatic classical field, with frequency ω_i, the transition probability per unit time takes a form that is reminiscent of the Fermi golden rule, valid for a constant perturbation coupling a discrete state to a continuum[5].

In the more general case where the classical field is non-monochromatic but stationary, one can follow the same line of computations that led to equation (8). This shows that the transition probability from a discrete state a to any state of the continuum can still be expressed as the integral of the product of two correlation functions: one, $G_A(t''-t')$, for the atomic dipole, another, $G_R(t''-t')$, for the radiation. In both cases, the field appears in the transition probability only via a correlation function. For the classical case, the quantum average value (10) must be replaced by the product of the negative and positive frequency components of the classical field:

$$\mathcal{G}_R^N(t''-t') = \mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R},t')\ \mathcal{E}_{\boldsymbol{d}}^{(+)}(\boldsymbol{R},t'') \tag{28}$$

Note that in this relation, in order to distinguish the quantum fields from the classical fields, these latter fields are written with curly letters; the subscript d means that the field has been projected onto the polarization unit vector defined in (7a).

Expression (28) has been obtained for a perfectly known classical field. Another possibility is that the classical field is only known in a probabilistic sense, as is the case with a classical statistical mixture of fields, with given probabilities. The transition probability (8), where the correlation function $\bar{G}_R^N(t''-t')$ is replaced by $\mathcal{G}_R^N(t''-t')$, must then be averaged over all the states of the statistical mixture, which amounts to replacing (28) by:

$$\mathcal{G}_R^N(t''-t') = \overline{\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R},t')\ \mathcal{E}_{\boldsymbol{d}}^{(+)}(\boldsymbol{R},t'')} \tag{29}$$

where the bar above the product of the two fields symbolizes the statistical average. It seems that for all signals involving one single photodetector atom, the quantum predictions are identical to those of a semiclassical theory, using a classical field with the same correlation function as the quantum field. In particular, for a stationary field, the Fourier transform of the field correlation function is simply the spectral distribution $I_R(\omega)$ of that field. For a quantum field, this property was established in Chapter XX – see relation (B-20). For a classical field, this property is a consequence of the Wiener-Khintchine theorem. It follows that the photoionization probability of an atom is the same, whether it is computed with a stationary quantum or classical field, as long as they both have the same spectral distribution.

[5] See for example § C-3 of Chapter XIII, and relation (C-37) in particular.

Comment

The equivalence of the predictions of the two theories is also valid for a non-stationary field. The semiclassical theory predicts that the photoionization rate at time t' is proportional to the classical field intensity $\overline{\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R},t')\mathcal{E}_{\boldsymbol{d}}^{(+)}(\boldsymbol{R},t')}$, which now depends on t' since the field is no longer stationary. We shall see in Complement $\mathrm{D_{XX}}$ that a similar result is obtained in quantum theory: the photoionization probability at time t' of an atom receiving a one-photon wave packet is again given by the modulus squared of a function, which can be considered to be the photon wave function, evaluated at $t = t'$.

3-b. Experiments with two photodetector atoms

The same line of computations that led to equation (27) can be followed for a classical field. It leads to an expression similar to (27), where the field quantum correlation function is replaced by the statistical average of the product of two negative frequency components and two positive frequency components of the classical field:

$$\overline{\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1,t_1)\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_2,t_2)\mathcal{E}_{d}^{(+)}(\boldsymbol{R}_2,t_2)\mathcal{E}_{d}^{(+)}(\boldsymbol{R}_1,t_1)} \tag{30}$$

As classical fields commute with each other, expression (30) can be rewritten as:

$$\overline{\mathcal{I}(\boldsymbol{R}_1,t_1)\mathcal{I}(\boldsymbol{R}_2,t_2)} \tag{31}$$

where:

$$\mathcal{I}(\boldsymbol{R}_1,t_1) = \mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1,t_1)\mathcal{E}_{d}^{(+)}(\boldsymbol{R}_1,t_1) \tag{32}$$

is the classical field intensity at point $\boldsymbol{R}_1$ at time t_1 and with a similar equation for $\mathcal{I}(\boldsymbol{R}_2,t_2)$. Correlations between the photoionization rates of the two photodetector atoms thus involve, in the semiclassical theory, the product of the intensities arriving on both photodetectors. The correlation function of the field amplitude is replaced by the correlation function of the intensity.

α. *Situations where a semiclassical treatment is adequate*

A first situation where quantum and semiclassical predictions agree is the case where the field state is a coherent state $|\{\alpha_i\}\rangle$ described by the set of classical normal variables $\{\alpha_i\}$ (Chapitre XIX, § B-3-b). Each mode i is in a coherent state $|\alpha_i\rangle$, meaning the state $|\{\alpha_i\}\rangle$ is an eigenket of the operator $\boldsymbol{E}^{(+)}(\boldsymbol{R},t)$ with eigenvalue $\mathcal{E}^{(+)}(\{\alpha_i\},\boldsymbol{R},t)$ equal to the classical field corresponding to the set of classical normal variables $\{\alpha_i\}$. In a similar way, the bra $\langle\{\alpha_i\}|$ is an eigenbra of $\boldsymbol{E}^{(-)}(\boldsymbol{R},t)$ with eigenvalue $\mathcal{E}^{(-)}(\{\alpha_i\},\boldsymbol{R},t)$. For a coherent state of the field, the rate W_{II} written in (27) therefore becomes: égal à:

$$\begin{aligned}
&W_{II}(\boldsymbol{R}_2,t_2;\boldsymbol{R}_1,t_1)\\
&= s^2\langle\{\alpha_i\}|\ E_{\perp\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1,t_1)E_{\perp\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_2,t_2)E_{\perp\boldsymbol{d}}^{(+)}(\boldsymbol{R}_2,t_2)E_{\perp\boldsymbol{d}}^{(+)}(\boldsymbol{R}_1,t_1)\ |\{\alpha_i\}\rangle\\
&= s^2\mathcal{E}^{(-)}(\{\alpha_i\},\boldsymbol{R}_1,t_1)\ \mathcal{E}^{(-)}(\{\alpha_i\},\boldsymbol{R}_2,t_2)\ \mathcal{E}^{(+)}(\{\alpha_i\},\boldsymbol{R}_2,t_2)\ \mathcal{E}^{(+)}(\{\alpha_i\},\boldsymbol{R}_1,t_1)
\end{aligned} \tag{33}$$

The quantum result for W_{II} does coincide with the semiclassical prediction. The same conclusion holds when the state of the quantum field is a statistical mixture of coherent states $|\{\alpha_i\}\rangle$ with statistical weights $P(\{\alpha_i\})$.

Another situation where the quantum and semiclassical predictions agree is the case of a thermal field. In the quantum description, Wick's theorem (Complement C_{XVI}) allows expressing the four-point correlation function appearing in W_{II} as the sum of products of two-point correlation functions. In a similar way, in the semiclassical theory, the thermal field is a Gaussian random field, and here again, the classical four-point correlation function is the sum of products of two-point correlation functions. Provided we use the same two-point correlation functions in both theories, their predictions agree.

Comment

The interferometric analysis of the electric field of the light emitted by the stars (to measure their angular diameter) is confronted with the problem of atmospheric fluctuations, which introduce a random phase shift between the two arms of the interferometer. The analysis of intensity correlations is much less sensitive to these fluctuations. As different parts of the stars emit incoherent waves, the total field received from the star is Gaussian, and the result we just established shows that analyzing intensity correlations allows obtaining two-point correlation functions and hence the same information as the one contained in a field correlation measurement. The validity of such a method, based on intensity correlations, was experimentally demonstrated in 1956 by Robert Hanbury Brown and Richard Twiss [46].

β. Situations requiring a quantum radiation treatment

An example of a situation where a quantum treatment of the radiation becomes essential is shown in Figure (2). A one-photon wave packet is emitted by atom A; this wave packet is described by ϕ (cf. Complement D_{XX}). It then goes through a beamsplitter LS that divides it into two wave packets: a transmitted wave packet ϕ_T and a reflected wave packet ϕ_R, which then arrive on two detectors D_1 and D_2.

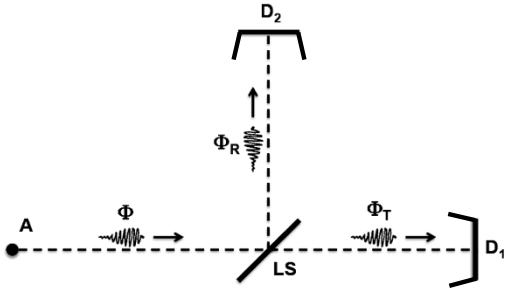


Figure 2: Atom A emits a photon described by a wave packet ϕ. This wave packet goes through a beamsplitter that divides it into a transmitted wave packet ϕ_T and a reflected wave packet ϕ_R, which then arrive on two detectors D_1 and D_2. The quantum and semiclassical predictions concerning the correlations between the signals detected on D_1 and D_2 are significantly different (see text).

In a quantum description of the radiation, the radiation state ψ after crossing the beamsplitter is still a one-photon state, described by the linear superposition of the two one-photon wave packets ϕ_T and ϕ_R, that is:

$$|\psi\rangle = c_T|\phi_T\rangle + c_R|\phi_R\rangle \tag{34}$$

In the expression (27) for the rate W_{II} giving the probability of a photoionization at time t_1 of detector D_1 at $\boldsymbol{R}_1$, and a photoionization at time t_2 of detector D_2 at $\boldsymbol{R}_2$, appears the squared norm of the ket:

$$E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_2, t_2)\ E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_1, t_1)\ |\psi\rangle \tag{35}$$

where ψ is given by (34). The first operator $E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_1, t_1)$, which destroys a photon, yields the vacuum $|0\rangle$ when it acts on state $|\psi\rangle$ that contains only one photon. The second destruction operator $E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_2, t_2)$ will then yield 0 when acting on the vacuum. The two detectors D_1 and D_2 cannot both undergo a photoionization. This result was, a priori, obvious: a single photon cannot produce two photoionizations.

If, on the other hand, the radiation emitted by the atom is described classically, the two wave packets ϕ_T and ϕ_R are classical wave packets which can ionize the detectors D_1 and D_2 they encounter.

Single photon sources are not easy to fabricate. An experiment close to the situation in Figure 2 is described in reference [47]. Instead of the atom A in Figure 2, it uses as a light source atoms emitting pairs of photons in a radiative cascade: the atom emits a photon of frequency ν_α going from a state α to a state β, then a photon ν_β going from state β to a state γ. If we call τ_β the radiative lifetime of state β, photon ν_β is emitted after photon ν_α in a time window having a width of the order of τ_β. Imagine we add to the experimental set-up of Figure 2 a third detector (not shown in the figure) that detects the ν_α photon and can trigger a departure time: after each detection of a ν_α photon, detectors D_1 and D_2 are activated, but for a short time interval of the order of τ_β. The probability of detecting a single photon ν_β during that time window is much higher than in a time window of the same length, but not triggered by the detection of a ν_α photon. This trigger method provides an equivalent of a single photon source, and was used to observe that a single ν_β photon could not simultaneously excite both detectors D_1 and D_2.

γ. *Resonance fluorescence of a single atom. Photon antibunching*

Another experiment clearly shows the need for a quantum description of radiation: the study of the second order correlation function W_{II} of the fluorescent light emitted by a single atom or ion, and excited by a resonant laser beam.

Imagine the emitting object A in Figure 2 is a single trapped ion[6]. Submitted to the resonant laser excitation, the ion emits a series of photons which enter the set-up of Figure 2. The distances between the beamsplitter and the detectors D_1 and D_2 are equal, so that the two wave packets associated with each photon arrive at the same time on D_1 and D_2.

[6]Ion trapping is now a well mastered technique. The results presented in Figure 3 have been obtained on a single $^{24}Mg^+$ ion [48]. The first experimental evidence for photon antibunching in the fluorescent light from a single atom were obtained on a sodium atomic beam at very low intensity, with an observation volume small enough for the probability of its containing more than one atom to be negligible [49].

For a continuous laser excitation with a constant intensity, the statistical properties of the fluorescent light are invariant under time translation; consequently, the quantum correlation function $W_{II}(\boldsymbol{R}_2, t_2; \boldsymbol{R}_1, t_1)$ characterizing the photoionizations detected on D_1 and D_2 depends only on $\tau = t_2 - t_1$. It shall be noted $g^{(2)}(\tau)$. Figure 3 plots the variations of $g^{(2)}(\tau)$ as a function of τ, for increasing values (from bottom to top) of the laser intensity. This figure shows that $g^{(2)}(\tau)$ is zero for $\tau = 0$; in other words, the detected photons are "antibunched" in time (one cannot detect simultaneously a photon at D_1 and another one at D_2).

The quantum interpretation of that result is as follows. Each photon emitted by the ion is detected either by D_1, or by D_2. Right after the emission of a photon, the ion is "projected" into the ground state a of the $a - b$ transition excited by the laser. This means it cannot immediately emit another photon as it must first be re-excited by the laser, and that takes a certain time. This is why $g^{(2)}(\tau)$ is zero for $\tau = 0$. Actually, after it emits a photon, the atom starting from g will oscillate between the state a and the excited state b at the Rabi frequency characterizing the atom-laser coupling, and proportional to the laser field amplitude. The Rabi oscillations explain the oscillations of $g^{(2)}(\tau)$ that appear in Figure 3 at frequencies higher and higher as the laser intensity increases.

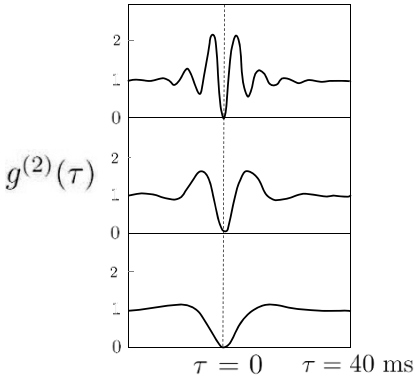


Figure 3: Intensity correlations in the resonance fluorescence of a single ion excited by a laser. The figure shows the time correlations $g^{(2)}(\tau)$ between the signals from the two detectors D_1 and D_2 as a function of the delay τ between two detections. The three curves correspond to increasing intensities (from bottom to top) of the laser beam exciting the resonant fluorescence of the trapped ion. It shows that $g^{(2)}(\tau)$ is zero for $\tau = 0$ and, for small positive values of τ, it increases with τ (figure adapted from [48]).

Let us examine now the predictions of a theory that classically treats the field emitted by the ion. We established above that the correlations between the photoionization rates of the two detectors are described by the correlation function $\overline{\mathcal{I}(t)\mathcal{I}(t+\tau)}$ of

the classical intensity $\mathcal{I}(t)$. For a stationary field, this classical correlation function $g_{\rm cl}^{(2)}$ depends only on τ:

$$\overline{\mathcal{I}(t)\mathcal{I}(t+\tau)} = g_{\rm cl}^{(2)}(\tau) \tag{36}$$

In addition, writing that $\overline{(\mathcal{I}(t)-\mathcal{I}(t+\tau))^2} \geqslant 0$, we get:

$$\overline{(\mathcal{I}(t))^2} + \overline{(\mathcal{I}(t+\tau))^2} \geqslant 2\ \overline{\mathcal{I}(t)\mathcal{I}(t+\tau)}$$

that is, taking into account the field stationarity and relation (36):

$$g_{\rm cl}^{(2)}(\tau=0) \geqslant g_{\rm cl}^{(2)}(\tau) \tag{37}$$

The semiclassical theory therefore predicts that $g_{\rm cl}^{(2)}(\tau)$ should not be an increasing function of τ in the vicinity of $\tau = 0$. This is contradicted by the experimental results shown in Figure 3, and hence proves that the fluorescent light emitted by a single ion excited by a resonant laser beam cannot be described as a classical field.

The radiation quantum theory is thus essential to account for all photoionization experimental results. This remains true even though the simple photoelectric effect observed on a single photodetector can be described by a semiclassical theory (without photons).

4. Two-photon photoionization

4-a. Differences with the one-photon photoionization

We now consider a two-photon absorption process similar to those studied in Complement A_{XX}, but where the final state b of the two-photon absorption process is now part of the atomic ionization continuum. This continuum starts at an energy E_I (ionization energy) above the energy of the ground state a (Fig. 4). This process is called two-photon photoionization.

The photoionization process transforms the atom into an ion and an electron, which moves away. When the distance between the electron and the ion is large enough, their Coulomb interaction energy becomes negligible and the electron energy is just its kinetic energy. Total energy conservation tells us that this kinetic energy is equal to:

$$E_{\rm kin} = 2\hbar\omega_i - E_I \tag{38}$$

If we plot the variations of $E_{\rm kin}$ as a function of ω_i, we get a straight half-line with slope $2\hbar$, which starts from the abscissa axis at point $E_I/2\hbar$. This result is a generalization of the photoelectric law established in 1905 by Einstein.

The previous result clearly shows that it is not necessary for the incident photon energy $\hbar\omega_i$ to be larger than the ionization energy for the atom to undergo photoionization. Figure 4 shows that $\hbar\omega_i$ is lower than E_I whereas $2\hbar\omega_i$ is larger than E_I. This result can be generalized: if $p\hbar\omega_i < E_I$, with $p = 1, 2, \ldots q-1$, but if $q\hbar\omega_i > E_I$, we have a q-photon photoionization. The kinetic energy of the photoelectron, once it is far enough from the ion, is equal to $E_{\rm kin} = q\hbar\omega_i - E_I$.

4-b. Photoionization rate

We first assume the radiation is monochromatic, and use expressions (13) and (14) of Complement A_{XX} for the two-photon absorption probability amplitude, whose modulus squared yields the probability. As the final state b now belongs to a continuum, we must sum this probability over b and use Fermi's golden rule to compute the photoionization probability per unit time. Since the modulus squared of equation (14) is proportional to $n_i(n_i - 1)$, where n_i is the number of incident photons, the photoionization rate increases as the square of the incident radiation intensity (for $n_i \gg 1$). In a similar way, it can be shown that a q-photon ionization increases as the qth power of the incident radiation (for $n_i \gg 1$).

Consider now the case of non-monochromatic stationary radiation. As in section § 2-b, we assume the radiation spectral density is centered around a frequency ω_{ex} with a width $\Delta\omega_R$ much smaller than the spectral bandwidth $\Delta\omega_A$ of the detector. The field correlation function $G_R^N(t' - t'')$ that appears in the two-photon absorption probability is still given by relation (27) of Complement A_{XX}. The two operators $\bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R}, t')$ appearing in this equation have a predominantly exponential $e^{-i\omega_{ex}t'}$ time dependence. Similarly, the two operators $\bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R}, t'')$ have a predominantly exponential $e^{+i\omega_{ex}t''}$ time dependence. With the same reasoning as in § 2-b-β, which led to (14), we set:

$$G_R^N(t'' - t') = e^{2i\omega_{ex}(t'-t'')} C_R^N(t'' - t') \tag{39}$$

where $C_R^N(t'' - t')$ is an "envelope" function with a much slower dependence in $t'' - t'$, and on a time scale of the order of $1/\Delta\omega_R$. In the double integral of (8), the correlation

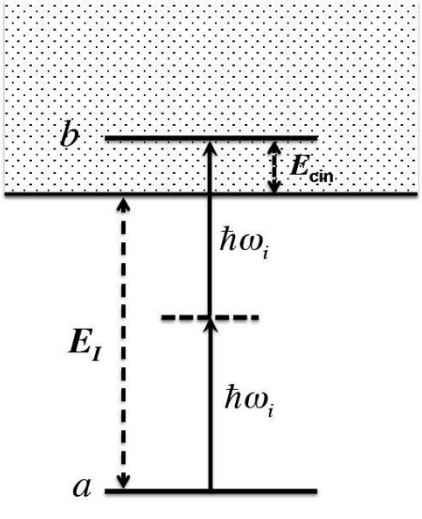


Figure 4: Two-photon photoionization. The atom goes from state a to state b, which is part of the ionization continuum, through the absorption of two photons, with energy $\hbar\omega_i$. The unbound electron produced at the end of that process leaves the atom with a kinetic energy equal, when the electron is far enough from the atom, to $E_{kin} = 2\hbar\omega_i - E_I$.

function $G_A(t''-t')$ is different from zero only for $|t''-t'| < \tau_{CA}$ where τ_{CA} is the correlation time of the atomic dipole, much shorter than $1/\Delta\omega_R$. We can therefore, as in § 2-b-β, take $t''=t'$ in the envelope function $C_R^N(t''-t')$ defined in relation (39) which yields G_R^N. We obtain for the field correlation function appearing in the two-photon ionization rate:

$$G_R^N(t''-t') = e^{2i\omega_{\rm ex}(t'-t'')}\times \langle\phi_{\rm in}^R|\left[\boldsymbol{e}_{ca}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{bc}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right] \left[\boldsymbol{e}_{bc}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{ca}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\rm in}^R\rangle \tag{40}$$

Note that the average value appearing in this equation is independent of t' since the radiation is supposed to be stationary.

4-c. Importance of fluctuations in the radiation intensity

Even when the radiation is monochromatic, i.e. when only a single mode i is populated, its intensity can take different values, spread out around an average value; the only case where the radiation intensity is well defined is when the radiation is in a Fock state $|n_i\rangle$. If we only consider stationary monochromatic radiation, the most general state is a statistical mixture of Fock states $|n_i\rangle$ with weight $p(n_i)$.

As an example, if the mode i is in thermal equilibrium at temperature T, the probability of its containing n_i photons is:

$$p(n_i) = \frac{1}{Z}\, e^{-\frac{(n_i+1/2)\hbar\omega_i}{k_B T}} \tag{41}$$

where k_B is the Boltzmann constant and Z the partition function given by:

$$Z = \sum_{n_i} e^{-\frac{(n_i+1/2)\hbar\omega_i}{k_B T}} \tag{42}$$

From these two equations one can easily compute the average value $\langle n\rangle$ of the number of photons n in this mode, as well as the average value $\langle n^2\rangle$ of n^2, for radiation at thermal equilibrium (see demonstration below). In particular, we can show that:

$$\langle n^2\rangle = \langle n\rangle + 2\langle n\rangle^2 \tag{43}$$

According to the previous results, the two-photon photoionization rate is proportional to $\langle n(n-1)\rangle = \langle n^2\rangle - \langle n\rangle$. If the radiation has a well-defined intensity, i.e. if it is in a Fock state $|n\rangle$, we have $\langle n^2\rangle = \langle n\rangle^2$ and the photoionization rate is proportional to $\langle n\rangle^2$ if $\langle n\rangle \gg 1$. On the other hand, if the radiation is in thermal equilibrium with the same average value $\langle n\rangle$, the photoionization rate is, according to (43), proportional to $2\langle n\rangle^2$, i.e. twice as large as for a state with no dispersion in intensity but same average value $\langle n\rangle$. An intensity fluctuation, keeping the average value $\langle n\rangle$ constant, considerably increases the photoionization rate. This result is to be expected for a nonlinear phenomenon: the values of n above the average value $\langle n\rangle$ contribute much more than the values below.

***Demonstration of relation* (43)**

This calculation has already been explained in §§ 2-b and 3-b of Complement B_{XV}. We briefly recall its principle. We note $x = \hbar\omega_i/k_BT$ and eliminate the irrelevant factor $e^{-x/2}$ from the partition function by setting:

$$Z(x) = e^{-x/2}K(x) \tag{44}$$

The function $K(x)$ is easily computed, since:

$$K(x) = 1 + e^{-x} + e^{-2x} + ... + e^{-nx} + ... = \frac{1}{1 - e^{-x}} \tag{45}$$

We also know that:

$$\begin{aligned}\langle n\rangle &= \sum_i n_i\ p(n_i) = \frac{1}{K(x)}\sum_i n_i\ e^{-n_i x} \\ &= \frac{-1}{K(x)}\frac{\mathrm{d}K(x)}{\mathrm{d}x}\end{aligned} \tag{46}$$

This leads to:

$$\langle n\rangle = \frac{1}{e^x - 1} \tag{47}$$

A similar calculation permits computing $\langle n^2\rangle$ using the second derivative $\mathrm{d}^2K(x)/\mathrm{d}x^2$, which leads to relation (43), the equivalent of relation (42b) of Complement B_{XV}.

If the radiation is no longer monochromatic, but still stationary and at thermal equilibrium, we can use the field correlation function (40) and Wick's theorem (Complement C_{XVI}) to rewrite this function as the sum of products of second order correlation functions:

$$\begin{aligned}G_R^N(t'' - t') =& e^{2i\omega_{\mathrm{ex}}(t'-t'')}\times \\ &\Big\{\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{ca}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{ca}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle \\ &\quad\times\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{bc}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{bc}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle \\ &+\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{ca}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{bc}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle \\ &\quad\times\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{bc}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{ca}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle\Big\}\end{aligned} \tag{48}$$

We thus get a sum of two terms, each proportional to the square of an average intensity.

5. Tunnel ionization by intense laser fields

As high power lasers became available, studies of multi-photon ionization processes led to the discovery of many new physical phenomena. In particular, when the instantaneous laser field becomes of the order of the Coulomb field binding the electron to the nucleus, ionization no longer results from a multi-photon ionization process, but from a tunnel effect. The laser field, yielding a potential that varies linearly as a function of the electron-nucleus distance, lowers the Coulomb potential sufficiently to allow the electron to escape

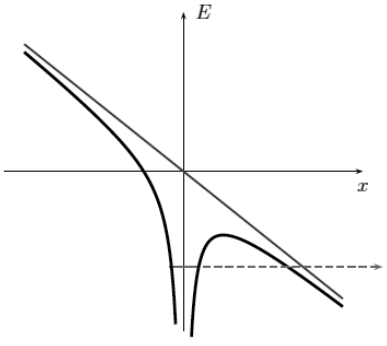


Figure 5: Effective potential seen by an electron undergoing tunnel ionization. The electron leaves the ion by tunneling through the potential barrier, sum of the ionic Coulomb potential and the linear potential $-qEx$ associated with the laser electric field, assumed to be linearly polarized along the x axis.

via a tunnel effect (see Fig.5). Once the electron has left the ion, it is accelerated by the laser field. As the oscillating laser field changes sign, the acceleration produced by the laser field is inverted, the electron comes back toward the ion and emits, as it passes close to the ion, a "bremsstrahlung" radiation (braking radiation). It can be shown that the frequency of this radiation is an odd harmonic of the laser frequency. The order of this harmonic is very high and can reach several hundred. As the fraction of the period of the laser field where the electron can escape by the tunnel effect is very small, the electron wave packet that leaves the ion has a very short time extension. The bremsstrahlung radiation it emits when it comes back to the ion also extends over a very short time, expressed in tens of attoseconds (one attosecond is equal to 10^{-18} sec). The interested reader can find an up to date review of these developments in Chapters 10 and 27 of reference [24].

Complement C_{XX}

Two-level atom in a monochromatic field. Dressed-atom method

1 Brief description of the dressed-atom method 2130
- 1-a State energies of the atom + photon system in the absence of coupling 2131
- 1-b Coupling matrix elements 2131
- 1-c Outline of the dressed-atom method 2133
- 1-d Physical meaning of photon number 2135
- 1-e Effects of spontaneous emission 2135

2 Weak coupling domain 2137
- 2-a Eigenvalues and eigenvectors of the effective Hamiltonian . . 2137
- 2-b Light shifts and radiative broadening 2138
- 2-c Dependence on incident intensity and detuning 2139
- 2-d Semiclassical interpretation in the weak coupling domain . . 2139
- 2-e Some extensions 2140

3 Strong coupling domain 2141
- 3-a Eigenvalues and eigenvectors of the effective Hamiltonian . . 2141
- 3-b Variation of dressed state energies with detuning 2142
- 3-c Fluorescence triplet 2144
- 3-d Temporal correlations between fluorescent photons 2145

4 Modifications of the field. Dispersion and absorption 2147
- 4-a Atom in a cavity 2147
- 4-b Frequency shift of the field in the presence of the atom 2148
- 4-c Field absorption 2149

Introduction.

The probability amplitude for an atom, subject to monochromatic radiation, to absorb a photon and go from a discrete state a to another discrete state b was calculated in § B of Chapter XX. We used, however, a perturbative treatment limited to lowest order with respect of the interaction Hamiltonian. The predictions of such an approximate calculation are, a priori, only valid for times that are sufficiently short for the higher order corrections to remain negligible. This complement presents another approach to atom-photon interactions, called the "dressed-atom approach", which does not have those limitations. It considers the atom and the mode of the quantum field it interacts with as a single quantum system. As this unified system is described by a time-independent Hamiltonian[1], one can study its energy diagram to obtain very useful information.

[1] This would obviously not be possible in a classical description of the radiation (even with a quantum treatment of the atom), since the field varies sinusoidally with time, and its coupling Hamiltonian with the atom is time-dependent.

This will allow us to improve the results of Chapter XX. The dressed-atom method yields a non-perturbative description of the physical processes under study, and hence remains valid even for intense fields. It provides new insights into several important physical phenomena: atoms' behavior in an intense electromagnetic field, spectral distribution of the light spontaneously emitted by an atom in such an intense field, time correlations between the emitted photons, origin of the forces exerted on an atom by radiation with a space-varying intensity.

As is usual in the literature, and as we already did in Complement C_{XIX}, we shall note g the atomic ground state and e the atomic excited state (instead of using our previous notation of a and b for the two atomic levels). For the same reason, rather than keeping the notation ω_i of Chapter XX for the angular frequency of the exciting beam, we shall use ω_L, that refers more explicitly to a laser beam, which can have a very high intensity.

We start in § 1 with a brief description of the dressed-atom method[2]. We assume the frequency ω_L to be close to resonance with the atomic frequency, noted $\omega_0 = (E_e - E_g)/\hbar$, but far enough from all the other atomic transition frequencies. The radiation-atom interaction will be characterized by a frequency called the "Rabi frequency". It is the equivalent, in this radiation quantum treatment, of the precession frequency of a spin turning around a classical radiofrequency field in a magnetic resonance experiment (Complement F_{IV}). Establishing the energy diagram of the unified atom-photon system will enable us to study both the weak coupling regime (Rabi frequency small compared to the natural width Γ of state e or to the detuning $|\omega_L - \omega_0|$ between the field and atomic frequencies) and the strong coupling regime (Rabi frequency large compared to the natural width and to the detuning).

The weak coupling regime is studied in § 2. We show that the ground state's energy undergoes a "light shift", by an amount that is proportional to the field intensity, and whose value as a function of the detuning follows a Lorentzian dispersion curve. The ground state also undergoes a radiative broadening, proportional to the field intensity, and which can be interpreted as a probability per unit time of leaving the ground state through the absorption of a photon.

We focus in § 3 on the strong coupling domain, where the Rabi oscillation between sates g and e appears, although damped because of the radiative instability of e. The energy diagram of the dressed-atom allows interpreting phenomena that are specific to the strong coupling regime, such as the fluorescence triplet and the temporal correlations between the photons emitted in the lateral components of that triplet.

The atom-field coupling perturbs, not only the atom, but also the field. We show in § 4 that the real and imaginary parts of the refractive index, associated with the atom's presence, are actually perturbations of the field, just as the light shifts and radiative broadening are perturbations of the atom.

1. Brief description of the dressed-atom method

Let us call "laser mode L" the quantum field mode that is populated by photons with frequency ω_L. In the absence of coupling between the atom A and this laser mode, the Hamiltonian of the total system $A + L$ is equal to $H_A + H_L$, where H_A is the atomic Hamiltonian, and H_L the Hamiltonian of the radiation in laser mode L. The energy levels

[2] A more detailed description can be found in Chapter VI of [21].

of $H_A + H_L$ are labeled by two quantum numbers: e or g for the atomic internal state[3], N for the photon number in the only radiation mode that is not empty and contains photons of frequency ω_L.

1-a. State energies of the atom + photon system in the absence of coupling

Consider the two states $|g, N\rangle$ and $|e, N-1\rangle$, whose energies with respect to the photon vacuum state are:

$$E_{g,N} = E_g + N\hbar\omega_L \qquad E_{e,N-1} = E_e + (N-1)\hbar\omega_L \tag{1}$$

Their energy difference is:

$$\begin{aligned} E_{g,N} - E_{e,N-1} &= \hbar\omega_L - (E_e - E_g) \\ &= \hbar(\omega_L - \omega_0) = \hbar\delta \end{aligned} \tag{2}$$

where:

$$\delta = \omega_L - \omega_0 \tag{3}$$

is the frequency detuning between the field frequency ω_L and the atomic transition frequency $\omega_0 = (E_e - E_g)/\hbar$. For a resonant field, i.e. when $\omega_L = \omega_0$, the two states are degenerate. Here we consider very small detunings (in absolute value) compared to ω_0:

$$|\delta| \ll \omega_0 \tag{4}$$

Consequently, even if the field is not exactly resonant, the two states $|g, N\rangle$ and $|e, N-1\rangle$ can be grouped in a two-dimensional multiplicity $\mathcal{E}(N)$:

$$\mathcal{E}(N) = \{|g, N\rangle, |e, N-1\rangle\} \tag{5}$$

which is far from all the other states of the system atom + field. There is an infinity of other multiplicities, for values of N going from 1 to infinity. As an example, Figure 1 shows the three multiplicities $\mathcal{E}(N-1)$, $\mathcal{E}(N)$ and $\mathcal{E}(N+1)$; there are $N-1$ others with lower energies, and an infinity with higher energies. Each multiplicity is separated from the next one by the distance $\hbar\omega_L$, and the spitting between the two levels inside one multiplicity equals $\hbar\delta$. The multiplicity $\mathcal{E}(1)$ corresponding to $N = 1$, includes the two states $|g, 1\rangle$ and $|e, 0\rangle$; on the other hand, the state $|g, 0\rangle$ is isolated.

1-b. Coupling matrix elements

The coupling H_I between the atom and the mode ω_L is proportional to the product of the atomic dipole moment $\boldsymbol{D}$ and the ω_L mode component of the radiation electric field $\boldsymbol{E}$. We can choose the origin of the coordinates so as to be able to write $e^{i\boldsymbol{k}_L\cdot\boldsymbol{R}} = -i$ where $\boldsymbol{R}$ is the position of the atomic center of mass. We then get for H_I:

$$H_I = -\sqrt{\frac{\hbar\omega_L}{2\varepsilon_0 L^3}}\ \boldsymbol{D}\cdot(\boldsymbol{\varepsilon}_L a_L + \boldsymbol{\varepsilon}_L^* a_L^\dagger) \tag{6}$$

[3]The atomic external degrees of freedom are treated classically by assuming that the atom is fixed at point $\boldsymbol{R}$.

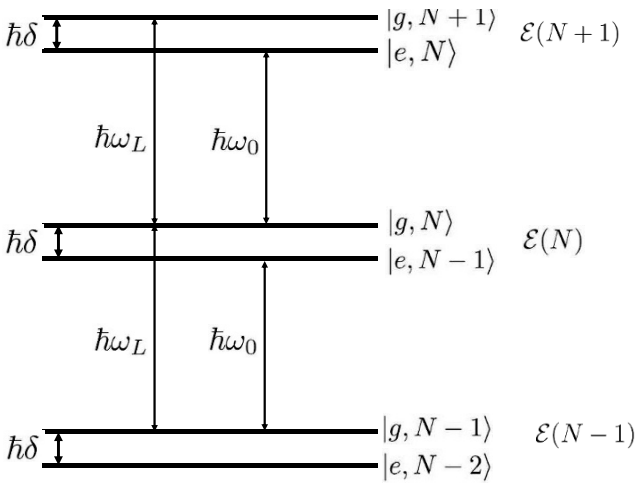


Figure 1: Energy levels of the system atom + photon in the absence of coupling. Only three adjacent multiplicities $\mathcal{E}(N-1)$, $\mathcal{E}(N)$ and $\mathcal{E}(N+1)$ are shown in the figure; many more exist below or above, corresponding respectively to smaller or larger values of N. Each vertical arrow links a pair of states having an energy difference indicated next to it.

The only non-zero matrix elements of the odd operator $\boldsymbol{D}$ are those between e and g. The annihilation and creation operators a_L and $a_L^\dagger$ change N by ∓ 1. It follows that $|g, N\rangle$ is coupled to $|e, N-1\rangle$ and $|e, N+1\rangle$, whereas $|e, N-1\rangle$ is coupled to $|g, N\rangle$ and $|g, N-2\rangle$. The two states $|g, N\rangle$ and $|e, N-1\rangle$ of the multiplicity $\mathcal{E}(N)$ are thus coupled to each other by the matrix element:

$$\langle e, N-1| H_I |g, N\rangle = \frac{\hbar\,\Omega_N}{2}, \tag{7}$$

where Ω_N is the "Rabi frequency" defined as:

$$\Omega_N = \sqrt{N}\ \ \Omega_1 \tag{8}$$

with:

$$\Omega_1 = \frac{2}{\hbar}\sqrt{\frac{\hbar\omega_L}{2\varepsilon_0 L^3}}\ \ \boldsymbol{\varepsilon}_L \cdot \langle e|\, \boldsymbol{D}\, |g\rangle \tag{9}$$

We assume Ω_1 is real and positive. If this is not the case, it suffices to change the relative phase[4] of the kets $|g\rangle$ and $|e\rangle$, which modifies the phase of the matrix element $\langle e|\, \boldsymbol{D}\, |g\rangle$; a suitable choice of that phase will make Ω_1 real and positive.

As operator H_I changes both the photon number by one unit and the atomic internal state, it does not have matrix elements between the kets of multiplicity $\mathcal{E}(N)$ and those of multiplicities $\mathcal{E}(N \pm 1)$. On the other hand, it can couple these kets with those of multiplicities $\mathcal{E}(N \pm 2)$ and, to higher order, to those of multiplicities even

[4]Such a phase change will affect the non-diagonal elements of the density matrix, leaving the physical predictions unchanged.

further away. However, the distances (in energy) between these multiplicities and $\mathcal{E}(N)$ are of the order of $2\hbar\omega_L$ (or of a multiple of that energy), whereas we have assumed that the interaction matrix elements are very small compared to $\hbar\omega_L$. The multiplicities other than $\mathcal{E}(N)$ have therefore energies too different from those of $\mathcal{E}(N)$ to play a significant role. We shall ignore their non-resonant coupling, which has a negligible effect for a quasi-resonant excitation.

1-c. Outline of the dressed-atom method

At the beginning of section § 1, we have described the quantum states of the system $A + L$ (atom + laser mode) in the absence of coupling; we showed they can be grouped into multiplicities $\mathcal{E}(N)$, with $N = 0, 1, 2, \ldots$ well separated from each other when condition (4) is satisfied. As an example, Figure 1 shows that the multiplicity $\mathcal{E}(N)$ for an atom with two levels g and e includes the states $|g, N\rangle$ and $|e, N-1\rangle$ separated by an energy $\hbar\delta$. Relation (7) tells us that these two states are coupled by an operator H_I describing the interaction between A and L, and that the corresponding matrix elements equal $\hbar\Omega_N/2$. We also discussed why, for a quasi-resonant excitation, the couplings between different multiplicities were negligible, so that one can separately study each multiplicity $\mathcal{E}(N)$.

α. *Dressed states and energies*

The first step in the dressed-atom approach is to study the energy levels of the system $A + L$ inside a multiplicity $\mathcal{E}(N)$, taking into account the coupling H_I restricted to $\mathcal{E}(N)$. We must diagonalize the Hamiltonian $H_A + H_L + H_I$ inside this sub-space. For a two-level atom, the restriction of that Hamiltonian to $\mathcal{E}(N)$, noted $H_{AL}^{(N)}$, is represented in the base of the kets written in (5) by a Hermitian 2×2 matrix equal to:

$$\begin{aligned}\left(H_{AL}^{(N)}\right) &= \begin{pmatrix} E_g + N\hbar\omega_L & \hbar\Omega_N/2 \\ \hbar\Omega_N/2 & E_e + (N-1)\hbar\omega_L \end{pmatrix} \\ &= [E_e + (N-1)\hbar\omega_L]\,\mathbb{1} + \hbar \begin{pmatrix} \delta & \Omega_N/2 \\ \Omega_N/2 & 0 \end{pmatrix} \end{aligned} \tag{10}$$

where $\mathbb{1}$ is the identity operator in $\mathcal{E}(N)$.

Because of the coupling created by the non-diagonal elements $\Omega_N/2$, the states $|g, N\rangle$ and $|e, N-1\rangle$ whose non-perturbed energies are separated by $\hbar\delta$ (left-hand side of Figure 2) are transformed into two states $|\psi_+(N)\rangle$ and $|\psi_-(N)\rangle$ with energies $\hbar\lambda_\pm$ (right-hand side of the figure):

$$\hbar\lambda_\pm = [E_e + (N-1)\hbar\omega_L] + \frac{\hbar\delta}{2} \pm \frac{\hbar}{2}\sqrt{\Omega_N^2 + \delta^2} \tag{11}$$

The new states $|\psi_+(N)\rangle$ and $|\psi_-(N)\rangle$ are linear superpositions of the initial states; they are called "dressed states", and their respective energies $\hbar\lambda_\pm$ the "dressed energies". This complement will show that a great number of interesting physical phenomena occurring when an atom is coupled to a laser mode can be interpreted in terms of these dressed states and their energies.

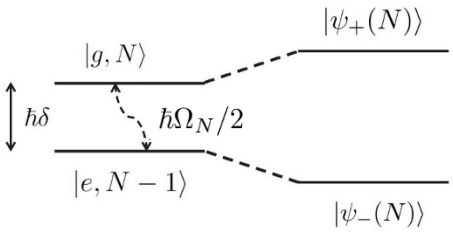


Figure 2: Energies of the states of the system $A+L$ within $\mathcal{E}(N)$, in the absence of coupling (left-hand side of the figure), and in the presence of a coupling of intensity $\hbar\Omega_N/2$ between the two initial states (right-hand side of the figure).

β. Rabi oscillation

Let us consider first a particularly simple application of the dressed-atom method. Imagine, for example, that the system is, at time $t=0$, in the state $|g,N\rangle$:

$$|\psi(t=0)\rangle = |\psi_{\text{in}}\rangle = |g,N\rangle \tag{12}$$

and let us try to find the probability that it will be found at a later time t in the state:

$$|\psi_{\text{fin}}\rangle = |e,N-1\rangle \tag{13}$$

We are dealing with the evolution of a system with two levels coupled by a static perturbation $\hbar\Omega_N/2$. This problem was studied in detail in § C of Chapter IV. We must first expand the initial state on the states $|\psi_\pm(N)\rangle$, and multiply each of them by an exponential $e^{-i\lambda_\pm t}$ whose argument is proportional to its energy $\hbar\lambda_\pm$:

$$|\psi(t)\rangle = e^{-i\lambda_+ t}\,\langle\psi_+(N)\,|g,N\rangle\quad|\psi_+(N)\rangle + e^{-i\lambda_- t}\,\langle\psi_-(N)\,|g,N\rangle\;\;|\psi_-(N)\rangle \tag{14}$$

The probability amplitude of finding the system in state $|e,N-1\rangle$ is then:

$$\begin{aligned}\langle e,N-1\,|\psi(t)\rangle \propto\; & e^{-i\lambda t/2}\,\langle\psi_+(N)\,|g,N\rangle\;\;\langle e,N-1\,|\psi_+(N)\rangle \\ & + e^{i\lambda t/2}\,\langle\psi_-(N)\,|g,N\rangle\;\;\langle e,N-1\,|\psi_-(N)\rangle\end{aligned} \tag{15}$$

where we introduced the Bohr frequency:

$$\lambda = \lambda_+ - \lambda_- = \sqrt{\Omega_N^2+\delta^2} \tag{16}$$

In (15), the sign $\propto$ simply means that we omitted a global phase factor $e^{-i(\lambda_+ +\lambda_-)t/2}$, with no physical significance. The probability of finding the system at time t in the final state (13) is therefore:

$$\begin{aligned}\mathcal{P}(t) &= |\langle e,N-1\,|\psi(t)\rangle|^2 \\ &= \left|C_+^g\right|^2\left|C_+^e\right|^2 + \left|C_-^g\right|^2\left|C_-^e\right|^2 + C_+^g\left(C_+^e\right)^*\left(C_-^g\right)^* C_-^e\;e^{-i\lambda t} + \text{c.c.}\end{aligned} \tag{17}$$

where the C are the scalar products:

$$C_{\pm}^{g} = \langle \psi_{\pm}(N) \, | g, N \rangle \qquad \text{and} \qquad C_{\pm}^{e} = \langle \psi_{\pm}(N) \, | e, N-1 \rangle \tag{18}$$

and c.c. means the complex conjugate.

We see that the probability $\mathcal{P}(t)$ is an oscillating function of time, with a frequency λ that is the only Bohr frequency $\sqrt{\Omega_N^2 + \delta^2}$ of the system within the multiplicity $\mathcal{E}_N$. This frequency can obviously be expanded to all orders of the perturbation Ω_N^2, but the result we obtained is not perturbative. The oscillation we found concerns the total system formed by a two-level atom placed in a monochromatic radiation field that can be intense and resonant: starting from state $|g, N\rangle$, the atom absorbs a photon and goes to state $|e, N-1\rangle$; it then comes back to state $|g, N\rangle$ by stimulated emission of a photon, and so on.

1-d. Physical meaning of photon number

The situation is different depending on whether the atom is placed in a real cavity or in free space.

If the atom is placed in a real cavity, as in some experiments, the field modes are the cavity eigenmodes. Such a situation will be discussed in § 4; the photon number N then has a perfectly clear physical meaning. The volume L^3, appearing in the modal expansion of the fields, and which is found in expressions (6) and (9) above, is simply the volume of the cavity containing the photons.

If the atom is in free space, the volume L^3, introduced to obtain discrete modes, is simply used in the computation, without any precise physical meaning. On the other hand, the energy density in the vicinity of the atom, proportional to $N\hbar\omega_L/L^3$, does have a physical meaning. Provided we keep N/L^3 constant, we can change N and L^3 arbitrarily without changing the coupling between the atom and the field; this is because the coupling is characterized by the Rabi frequency Ω_N, which depends on $\sqrt{N/L^3}$. In that case, the photon number N does not have an intrinsic physical meaning.

Imagine, for example, that the field is in a coherent state (Complement $\mathrm{G_V}$). The values of N are then distributed around an average value $\langle N \rangle$ in an interval of width $\Delta N = \sqrt{\langle N \rangle}$, very small in relative value compared to $\langle N \rangle$, but very large in absolute value. If both N and L^3 go to infinity, keeping the ratio N/L^3 constant, the Rabi frequency Ω_N will barely change in relative value even when N varies over a large interval around $\langle N \rangle$. The frequency Ω_N can thus be replaced in (10) by a constant Ω_R (which does not depend on N):

$$\Omega_N \simeq \Omega_R \tag{19}$$

This will be done in what follows and in § 3.

1-e. Effects of spontaneous emission

We have ignored, until now, all the field modes others than the laser mode. However, when the atom is in the excited state e, it can spontaneously emit a photon in another mode. This means that, in addition to the atom A and the laser mode L, we must take into account the system R including all the modes that, initially, did not contain any photons. As R is a very large system (sometimes called "reservoir" for that

reason), the coupling effects between $A + L$ and R must be described by a so-called "master equation"; this equation describes the evolution of the density operator σ_{A+L} of $A+L$ under the effect of the coupling with R (see part D of Chapter VI in [21]). Though we shall not introduce this master equation here, we shall merely discuss the physical interpretation of the results it leads to.

As the frequency spectrum of the reservoir R has a width $\Delta\omega$ of the order of the optical frequency, its associated correlation time $\tau_c \simeq 1/\Delta\omega$ is much shorter than all the other characteristic times of the problem. It is, in particular, much shorter than the radiative lifetime τ_R:

$$\tau_R = \frac{1}{\Gamma} \tag{20}$$

where Γ is the natural width of the excited state e; it is also shorter than the inverse of the Rabi frequency, which yields the characteristic time of the coupling to the laser mode. This means that, when the system $A + L$ is in the state $|e, N-1\rangle$ and a spontaneous emission occurs, it lasts for a time interval too short for the atom to have sufficient time to couple with L. The system then goes quasi-instantaneously from state $|e, N-1\rangle$, which belongs to $\mathcal{E}(N)$, to state $|g, N-1\rangle$ in the lower multiplicity $\mathcal{E}(N-1)$ – see Figure 1. As a consequence (see § C-3 of Chapter III in [21], as well as Complement D_{XIII}), the evolution within $\mathcal{E}(N)$ can still be described by the same equations as above, provided we simply add an imaginary term to the energy E_e of the excited state:

$$E_e \Longrightarrow E_e - i\hbar\frac{\Gamma}{2} \tag{21}$$

This means that, to describe the evolution of the system $A+L$ within[5] $\mathcal{E}(N)$ while taking into account spontaneous emission processes, we must replace the Hamiltonian $H_{AL}^{(N)}$ written in (10) by the effective non-Hermitian Hamiltonian:

$$\left(H_{\text{eff}}^{(N)}\right) = [E_e + (N-1)\hbar\omega_L]\,\mathbb{1} + \hbar\begin{pmatrix} \delta & \Omega_R/2 \\ \Omega_R/2 & -i\Gamma/2 \end{pmatrix} \tag{22}$$

Because of the imaginary term $-i\Gamma/2$ appearing in the matrix, the two eigenvalues of $H_{\text{eff}}^{(N)}$ also have an imaginary part: the two dressed states are now unstable as a result of spontaneous photon emission, which can occur in any of these states.

Within a constant factor, given by the term proportional to $\mathbb{1}$ on the right-hand side of (22), the eigenvalues $\hbar\lambda_\pm$ of H_{eff} are obtained by diagonalizing the 2-dimensional matrix that follows on the right-hand side of (22). The exact solution is written[6]:

$$\lambda_\pm = -i\frac{\Gamma}{4} + \frac{\delta}{2} \pm \frac{1}{2}\sqrt{\Omega_R^2 + \left(\delta + i\frac{\Gamma}{2}\right)^2} \tag{23}$$

We now discuss the physical meaning of these results in two limiting cases.

[5] The coupling with the reservoir induces other important effects leading to transitions between different multiplicities. This is what happens, for example, in the fluorescence phenomenon studied in § 3-c.

[6] For brevity, we use a slighlty incorrect mathematical notation, since the square root sign must in principle be applied on a real and positive number. What we mean with the square root sign written on the right-hand side of (23), is either one of the two complex numbers whose square is equal to the complex number under the root sign.

2. Weak coupling domain

We start with the weak coupling domain, which is more directly related with the results of Chapter XX.

2-a. Eigenvalues and eigenvectors of the effective Hamiltonian

Consider first the case where the non-diagonal coupling $\hbar\Omega_R/2$ between the two non-perturbed states of $\mathcal{E}(N)$ is small compared to the differences between the energies of these two states (including the imaginary term associated with the natural width of e). As this difference is complex, we must take its modulus:

$$\Omega_R \ll \left|\delta + i\frac{\Gamma}{2}\right| \tag{24}$$

This inequality is satisfied if:

$$\Omega_R \ll \Gamma \quad \text{or} \quad \Omega_{\mathrm{R}} \ll |\delta| \tag{25}$$

The weak coupling domain is thus obtained for low light intensities, or large frequency detunings.

For weak coupling, we can apply perturbation theory to obtain the energy corrections for the states $|g, N\rangle$ and $|e, N-1\rangle$, to order 2 in Ω_R. Starting from (22), we obtain in this way the correction $\delta E_{g,N}$ to the energy of state $|g, N\rangle$:

$$\delta E_{g,N} = \hbar\frac{(\Omega_R/2)^2}{\delta + i\Gamma/2} = \hbar\delta_g - i\hbar\frac{\gamma_g}{2} \tag{26}$$

where:

$$\delta_g = \frac{\delta}{4\delta^2 + \Gamma^2}\Omega_R^2 \qquad \text{and} \qquad \gamma_g = \frac{\Gamma}{4\delta^2 + \Gamma^2}\Omega_R^2 \tag{27}$$

A similar calculation yields for the correction to the energy of state $|e, N-1\rangle$:

$$\delta E_{e,N-1} = -\hbar\delta_g + i\hbar\frac{\gamma_g}{2} \tag{28}$$

We can write the approximate eigenvalues of the effective Hamiltonian (22) in the form:

$$\begin{aligned}\lambda_+ &\simeq \delta + \delta_g - i\frac{\gamma_g}{2} + ..\\ \lambda_- &\simeq -\delta_g - i\frac{\Gamma}{2} + i\frac{\gamma_g}{2} + ..\end{aligned} \tag{29}$$

which coincides with an expansion in powers of Ω_R/Γ of the exact result (23).

Perturbation theory also allows computing the eigenstates of H_{eff} to first order in Ω_R. The state $|\overline{g, N}\rangle$, which tends towards $|g, N\rangle$ when Ω_R goes to zero, is written:

$$|\overline{g, N}\rangle = |g, N\rangle + \frac{(\Omega_R/2)}{\delta + i\Gamma/2}|e, N-1\rangle \tag{30}$$

This means that state $|g, N\rangle$ is "contaminated" by state $|e, N-1\rangle$. A similar computation for state $|\overline{e, N-1}\rangle$, which tends towards $|e, N-1\rangle$ when Ω goes to zero, yields:

$$|\overline{e, N-1}\rangle = |e, N-1\rangle - \frac{(\Omega_R/2)}{\delta + i\Gamma/2}|g, N\rangle \tag{31}$$

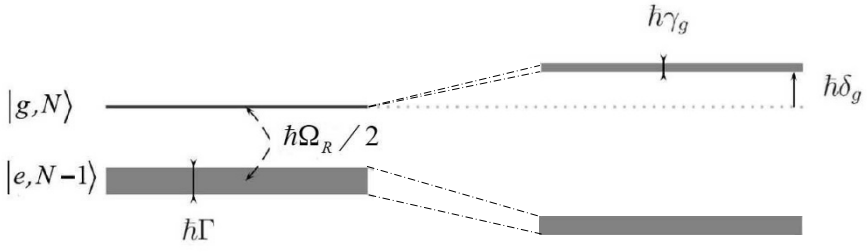


Figure 3: Non-perturbed states (left-hand side of the picture) and perturbed states (right-hand side) in the $\mathcal{E}(N)$ multiplicity. The coupling, characterized by the Rabi frequency Ω_R, shifts state $|g,N\rangle$ by a quantity $\hbar\delta_g$ (representing the light shift of the ground sate g); its wave function is "contaminated" by the unstable wave function of state $|e,N-1\rangle$, meaning that the ground state also becomes unstable as shown by its radiative broadening $\hbar\gamma_g$. State $|e,N-1\rangle$ is shifted in the opposite way, compared to $|g,N\rangle$; its width is reduced from $\hbar\Gamma$ to $\hbar(\Gamma-\gamma_g)$.

2-b. Light shifts and radiative broadening

The real parts of $\delta E_{g,N}$ and $\delta E_{e,N-1}$ represent shifts in the energy levels induced by the coupling with the light and called for that reason "light shifts". The imaginary part of $\delta E_{g,N}$ represents a radiative broadening of state $|g,N\rangle$, which becomes unstable under the coupling effect. The imaginary part of $\delta E_{e,N-1}$ describes a reduction of the radiative broadening $\hbar\Gamma$ of state $|e,N-1\rangle$.

Figure 3 shows, in its left-hand side, the non-perturbed states $|g,N\rangle$ and $|e,N-1\rangle$ in the $\mathcal{E}(N)$ multiplicity. They are separated by the gap $\hbar\delta$; if δ is positive, state $|g,N\rangle$ is above state $|e,N-1\rangle$; conversely, if δ is negative, state $|e,N-1\rangle$ is now above state $|g,N\rangle$. The thickness of the line representing state $|e,N-1\rangle$ symbolizes its natural width $\hbar\Gamma$. The right-hand side of the figure represents the states perturbed by the interaction with light. The two states $|g,N\rangle$ and $|e,N-1\rangle$ repel each other, meaning that they undergo light shifts of opposite signs. The $\hbar\delta_g$ shift of $|g,N\rangle$ is positive if $|g,N\rangle$ is above $|e,N\rangle$, i.e. if δ is positive; it is negative if δ is negative. The stable $|g,N\rangle$ state is also "contaminated" by the unstable state $|e,N-1\rangle$, which makes it unstable as shown by its radiative broadening $\hbar\gamma_g$. An atom in state g cannot stay there indefinitely: it will leave that state with a probability per unit time equal to γ_g, which can be interpreted as the photon absorption rate of an atom in state g. Conversely, due to the "contamination" of the unstable state $|e,N-1\rangle$ by the stable state $|g,N\rangle$, state $|e,N-1\rangle$ becomes less unstable and its width is reduced from $\hbar\Gamma$ to $\hbar(\Gamma-\gamma_g)$.

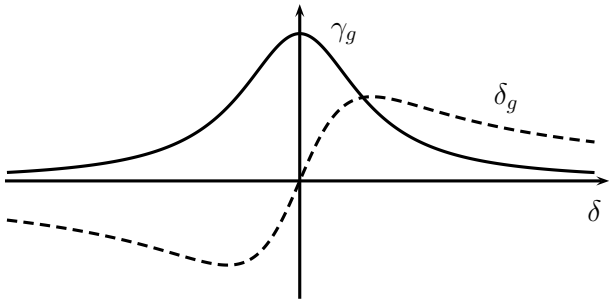


Figure 4: Plots of the light shift $\hbar\delta_g$ (dashed line curve) and of the radiative broadening $\hbar\gamma_g$ (solid line curve) as a function of the detuning δ between the laser frequency and the atomic frequency.

2-c. Dependence on incident intensity and detuning

The shifts $\hbar\delta_g$ and radiative broadenings $\hbar\gamma_g$ given by equation (27) are all proportional to Ω_R^2, hence to the number of incident photons, meaning to the light intensity. Their variations with the detuning δ follow respectively Lorentzian dispersion and absorption curves (Fig. 4).

When the detuning is very large (in absolute value) compared to the natural width of e ($|\delta| \gg \Gamma$), we can neglect Γ^2 compared to $4\delta^2$ in the denominators of expressions (27), which yields:

$$|\delta_g| = \frac{\Omega_R^2}{4|\delta|} \qquad \gamma_g = \frac{\Omega_R^2 \Gamma}{4\delta^2} \tag{32}$$

This leads to:

$$\gamma_g = |\delta_g| \frac{\Gamma}{|\delta|} \ll |\delta_g| \tag{33}$$

For large detunings, the light shifts are thus much larger than the radiative broadenings.

2-d. Semiclassical interpretation in the weak coupling domain

In this weak coupling domain, the atom responds linearly to the incident field; the results we just discussed can be interpreted semiclassically, in terms of a dipole induced by the incident field (see for example [50]). This dipole has a component in phase with the field and a quadrature component, related to the field by a dynamic polarizability $\alpha(\omega)$.

The quadrature component absorbs energy from the field. It varies with the detuning δ as an absorption curve; it is responsible for the absorption rate associated with the radiative broadening γ_g. The in-phase component of the dipole yields a polarization energy. Its variation with the detuning follows a dispersion curve. It is responsible for the light shift, just as the Stark shift results from the interaction of a static electric field with the static dipole it induces. This is why this light shift is often called a "dynamic Stark effect".

2-e. Some extensions

We now discuss some direct, important extensions of the previous study.

α. Non-monochromatic incident radiation

Imagine the radiation state is now a Fock state $|n_1, n_2, ..., n_i, ..\rangle$, or a statistical mixture of such states; the radiation spectral distribution is then described by the function $I(\omega_L)$. To second order perturbation theory (weak coupling domain), the processes that come into play in the light shifts and radiative broadenings are stimulated absorption and re-emission of photons. When several modes contain photons and the radiation state is a Fock state, the photon must be re-emitted by stimulated emission in the same mode it was absorbed from (otherwise the matrix element describing the second order coupling would be zero). This means that the effects of the different field modes can be added independently; we then get for δ_g and γ_g:

$$\delta_g \propto \int \mathrm{d}\omega_L |\mathcal{D}_{eg}|^2 I(\omega_L) \ \frac{\delta}{4\delta^2 + \Gamma^2}$$
$$\gamma_g \propto \int \mathrm{d}\omega_L |\mathcal{D}_{eg}|^2 I(\omega_L) \ \frac{\Gamma}{4\delta^2 + \Gamma^2} \tag{34}$$

β. Degenerate ground state

Assume the ground state g has a non-zero angular momentum J_g and therefore contains several Zeeman sublevels m_g; one can then show [51] that the sublevels of g having a well-defined light shift and radiative broadening are obtained by diagonalizing the Hermitian matrix whose elements are:

$$\sum_{m_e} \langle m_g | \boldsymbol{\varepsilon}_L^* \cdot \boldsymbol{D} | m_e \rangle \langle m_e | \boldsymbol{\varepsilon}_L \cdot \boldsymbol{D} | m_g' \rangle \tag{35}$$

where the states m_e are the sublevels of e. The eigenstates $|g_\alpha\rangle$ of this matrix, with eigenvalues μ_α, undergo light shifts proportional to $\mu_\alpha \delta_g$ and radiative broadenings proportional to $\mu_\alpha \gamma_g$ (where δ_g and γ_g are the shifts and broadenings for a two-level atom).

Reference [52] studies the symmetry properties of matrix (35), and discusses the equivalence between the light shifts and the effect of fictitious magnetic and electric fields acting on the ground multiplicity of the atom. We shall simply focus here on the simple case of a $J_g = 1/2 \leftrightarrow J_e = 1/2$ transition such as, for example, the hyperfine component $F = 1/2 \leftrightarrow F = 1/2$ of the $6^1S_0 \leftrightarrow 6^3P_1$ transition of the 199-isotope of mercury (λ_L= 253.7 nm). It is on such a transition that light shifts were observed for the first time [53].

The left-hand side of Figure 5 shows the components σ_+ and σ_- of this transition, that link respectively $m_g = -1/2$ to $m_e = +1/2$ and $m_g = +1/2$ to $m_e = -1/2$. If the beam polarization is σ_+, level $m_g = -1/2$ has a non-zero and well defined light shift, since the absorption and re-emission of a σ_+ photon can link sublevel $m_g = -1/2$ only to itself. On the other hand, level $m_g = +1/2$ is not shifted because there is no σ_+ optical transition starting from $m_g = +1/2$. We get opposite conclusions for a σ_- polarization of the light beam: the light shift of sublevel $m_g = +1/2$ is well-defined and sublevel $m_g = -1/2$ is not shifted. Now, by symmetry, the Clebsch-Gordan coefficients (Complement B_X) for the σ_+ and σ_- transitions are equal; the light shifts have the same

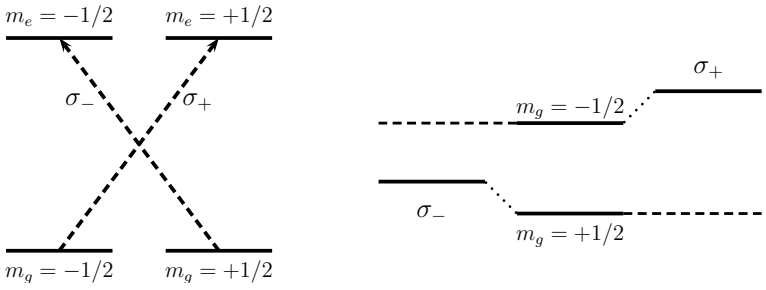


Figure 5: The left-hand side of the figure represents the $J_g = 1/2 \to J_e = 1/2$ transition, and the light beam polarizations that can induce transitions between its Zeeman sublevels. The diagram on the right-hand side plots in its center the ground state energy levels in the absence of any light beam (Zeeman levels in a static magnetic field); the two lateral extensions depict their light shifts by a non-resonant light beam with polarization σ_+ (on the right), or σ_- (on the left). In the first case, the light selectively shifts the sublevel $m_g = -1/2$, in the second case, it shifts the sublevel $m_g = +1/2$. This is why, depending on whether the beam polarization is σ_+ or σ_-, the variation in the gap between the two Zeeman sublevels changes sign.

value for a σ_+ excitation of sublevel $m_g = -1/2$ and for a σ_- excitation with same intensity of sublevel $m_g = +1/2$.

In the presence of a static magnetic field, there is an energy gap between the two atomic sublevels $m_g = \pm 1/2$ (Zeeman effect). The right-hand side of Figure 5 shows that a non-resonant light excitation changes this gap by the same amount, but in the opposite directions[7] depending on whether it has a σ_+ or σ_- polarization. The ground state magnetic resonance line, detected by optical methods using a resonant beam (Complement C_{XIX}, § 2-b), is thus shifted when a second non-resonant beam is applied; this shift has opposite directions, depending on whether that beam has a σ_+ or σ_- polarization. As relaxation times can be very long in the ground state, its magnetic resonance line is very narrow, which allows detecting very small light shifts, of the order of a few Hz. This is how the existence of light shifts were demonstrated in 1961, when laser sources were not yet available in laboratories [53]. With laser sources, one routinely observes shifts of the order of 10^6 Hz, and even more.

3. Strong coupling domain

We now examine how the previous results are modified in the strong coupling regime.

3-a. Eigenvalues and eigenvectors of the effective Hamiltonian

A strong coupling regime means that the non-diagonal element $\hbar\Omega_R/2$ of the effective Hamiltonian written in (22) is large compared to the difference between two diagonal elements:

$$\Omega_R \gg |\delta| \quad \text{and} \quad \Omega_R \gg \Gamma \tag{36}$$

[7] We assume the detuning δ is large compared to the Zeeman splitting.

For the sake of simplicity, we shall only consider the resonant case ($\delta = 0$). Equation (23), which yields the eigenvalues of the 2×2 matrix of (22) for any value of Ω_R, then becomes:

$$\lambda_{\pm} = -i\frac{\Gamma}{4} \pm \frac{1}{2}\sqrt{\Omega_R^2 - \frac{\Gamma^2}{4}} \tag{37}$$

where, as we did above, we use the concise notation for the square root of a number that is not always positive (see note 6).

As long as $\Omega_R < \Gamma/2$, the last term on the right-hand side of (37) is purely imaginary. The same is true for the two eigenvalues $\lambda_{\pm}$, which are equal to:

$$\lambda_{\pm} = -i\left[\frac{\Gamma}{4} \mp \frac{1}{2}\sqrt{\frac{\Gamma^2}{4} - \Omega_R^2}\right] \tag{38}$$

If, in addition, $\Omega_R \ll \Gamma$, a limited expansion of (38) in powers of Ω_R/Γ yields $\lambda_+ = -i\gamma_g/2$ and $\lambda_- = -i(\Gamma - \gamma_g)/2$; as expected, we confirm the results of the previous § for the weak coupling regime. As Ω_R increases, while remaining lower than $\Gamma/2$, the eigenvalue λ_+ increases whereas λ_- decreases, but their sum $(\lambda_+ + \lambda_-)$ remains constant and equal to $-i\Gamma/2$. When Ω_R reaches the value $\Gamma/2$, both eigenvalues $\lambda_{\pm}$ are equal to $-i\Gamma/4$.

As soon as Ω_R goes beyond $\Gamma/2$, the last term in (37) becomes real. The two eigenvalues $\lambda_{\pm}$ have opposite real parts and the same imaginary part, equal to $-i\Gamma/4$; the two dressed levels now always have the same width $\Gamma/4$. As the coupling becomes strong ($\Omega_R \gg \Gamma$), the energies are equal to:

$$\lambda_{\pm} \simeq -i\frac{\Gamma}{4} \pm \frac{\Omega_R}{2} \tag{39}$$

and the eigenvectors tend toward symmetric and antisymmetric linear combinations of $|g, N\rangle$ and $|e, N-1\rangle$:

$$|\psi_{\pm}(N)\rangle \to \frac{1}{\sqrt{2}}\left[|g, N\rangle \pm |e, N-1\rangle\right] \tag{40}$$

Such states can no longer be considered to be a result of a light mutual contamination of the non-perturbed states of multiplicity $\mathcal{E}(\mathcal{N})$. They are actually *entangled*, and hence impossible to consider as products of an atomic state and a field state. These states of the global atom + field system are commonly called *atom-field dressed states*.

3-b. Variation of dressed state energies with detuning

The solid lines in Figure 6 show the energies of the dressed states $|\psi_{\pm}(N)\rangle$ as a function of $\hbar\omega_L$. The energies are defined with respect to the energy of the non-perturbed state $|e, N-1\rangle$, chosen to be equal to $\hbar\omega_0$, and represented by a horizontal line with ordinate $\hbar\omega_0$. Compared to this energy, state $|g, N\rangle$ has an energy equal to $\hbar\omega_L$ which varies, as a function of $\hbar\omega_L$, as a straight line of slope unity. This line intersects the horizontal line representing the energy of $|e, N-1\rangle$ at a point with abscissa $\hbar\omega_0$.

At resonance ($\omega_L = \omega_0$), the two dressed states $|\psi_{\pm}(N)\rangle$ are separated by an energy $\hbar\Omega_R$, since we assume we are in the strong coupling domain where $\Omega_R \gg \Gamma$. Let

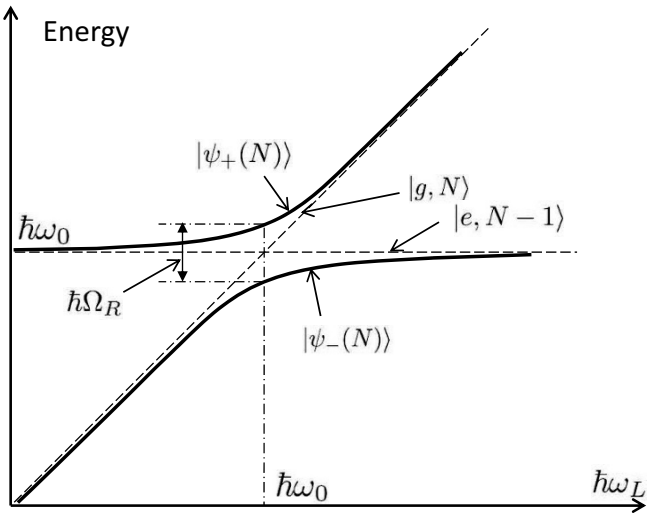


Figure 6: Energies of the dressed states $|\psi_\pm(N)\rangle$ (solid lines) and of the non-perturbed states $|g,N\rangle$ and $|e,N-1\rangle$ (dashed lines) as a function of $\hbar\omega_L$. The energies are defined with respect to the energy of the non-perturbed state $|e,N-1\rangle$, chosen to be equal to $\hbar\omega_0$. As $\hbar\omega_L$ varies, the energies of the dressed states follow a hyperbola whose asymptotes are the straight lines representing the energies of the non-perturbed states (anticrossing).

us first completely neglect Γ. Leaving resonance, and as the detuning becomes larger and larger (in absolute value), one finally reaches regions where $|\delta|$ is larger than Ω_R, which corresponds to a weak coupling regime. Varying the detuning, one then continuously goes from a strong coupling to a weak coupling region. The energies of the dressed levels $|\psi_\pm(N)\rangle$ follow a hyperbola whose asymptotes are the energies of the non-perturbed states $|g,N\rangle$ and $|e,N-1\rangle$ (Figure 6). As they come close to their asymptotes, the dressed states become very close to the non-perturbed corresponding states, and the distance between the hyperbola and its asymptote is simply the light shift δ_g defined in (27).

To take into account the natural width Γ of the excited level e, one should add a width to the dressed levels shown in Figure 6. Far away from the anticrossing center, close to the asymptotes, the width would be $\Gamma-\gamma_g$ for the dressed states that are close to the horizontal asymptote, and γ_g for the dressed states that are close to the asymptote with slope one. Following one hyperbola branch continuously, the width will progressively change from one of these values to the other, and take the value $\Gamma/4$ at the center of the anticrossing.

Another interesting phenomenon occurs when the system continuously follows one of the hyperbola branches. Imagine it follows the lower branch, from left to right, for instance because the excitation frequency is slowly varied. If the transit is slow enough to neglect any non-adiabatic transition to the other dressed state, i.e. to the other hyperbola branch, one continuously goes from state $|g,N\rangle$ to state $|e,N-1\rangle$. This is

another convenient way to go from g to e: instead of applying a resonant field during the time necessary for the Rabi oscillation to bring the system from g to e (π pulse), one slowly scans the field frequency through resonance, from a lower to a higher value. Note however that this scanning cannot be too slow, since it must occur on a time scale that is too rapid for the dissipative processes to be able to change the atomic internal state. Such a transit is often referred to as an "*adiabatic fast passage*", as it must be slow enough to remain adiabatic and fast enough to avoid any dissipation during the transit time. The dressed-atom approach allows clearly specifying the conditions for transferring the atom from one level to another.

3-c. Fluorescence triplet

With the dressed-atom approach, we can also simply explain the spectrum of the lines spontaneously emitted by an atom subjected to intense radiation. When studying elastic scattering in § E-1 of Chapter XX, we showed that, when the exciting radiation had an intensity low enough to allow a perturbation treatment, the radiation emitted spontaneously by the atom had the same frequency as the exciting radiation. We now show that the situation is different in the case of an intense excitation radiation: new frequencies appear in the light emitted by the atom[8].

We assume the exciting radiation to be resonant and intense, so that the two dressed states $|\psi_{\pm}(N)\rangle$ of multiplicity $\mathcal{E}(N)$ are separated by an energy interval $\hbar\Omega_R$ (Figure 7). These two states are linear superpositions of the states $|g, N\rangle$ and $|e, N-1\rangle$; consequently, they both have a non-zero projection onto $|e, N-1\rangle$. Similarly, the two states $|\psi_{\pm}(N-1)\rangle$ of multiplicity $\mathcal{E}(N-1)$ are linear superpositions of the states $|g, N-1\rangle$ and $|e, N-2\rangle$; they both have a non-zero projection onto $|g, N-1\rangle$. The lines emitted spontaneously by the atom are those that link two energy levels between which the atomic dipole operator $\boldsymbol{D}$ has a non-zero matrix element. Since $\boldsymbol{D}$ does not change the photon number N and can link e to g, the matrix element $\langle g, N-1|\boldsymbol{D}|e, N-1\rangle$ is non-zero; each of the two states $|\psi_{\pm}(N)\rangle$ can be linked via $\boldsymbol{D}$ to each of the two states $|\psi_{\pm}(N-1)\rangle$. The four radiative transitions represented by the curly arrows in Figure 7 are thus possible and correspond to three distinct frequencies: frequency $\omega_L + \Omega_R$ for the $|\psi_+(N)\rangle \rightarrow |\psi_-(N-1)\rangle$ transition; frequency ω_L for both the $|\psi_+(N)\rangle \rightarrow |\psi_+(N-1)\rangle$ and the $|\psi_-(N)\rangle \rightarrow |\psi_-(N-1)\rangle$ transitions; frequency $\omega_L - \Omega_R$ for the $|\psi_-(N)\rangle \rightarrow |\psi_+(N-1)\rangle$ transition. We get a frequency triplet for the spontaneously emitted light, which was first predicted by Mollow [54] by using a semiclassical treatment.

Autler-Townes doublet

Imagine that one of the two atomic states we considered until now, for example g, is connected via an allowed transition to a third state c, meaning that $\langle c|\boldsymbol{D}|g\rangle$ is non-zero. Let us also assume that the radiation frequency ω_L that is resonant for the $g \rightarrow e$ transition, is completely off-resonance for the $g \rightarrow c$ transition; consequently, it does not perturb state c, so that the sates $|c, N\rangle$ can be considered as eigenstates of the total Hamiltonian, even if they are slightly shifted. The two states $|\psi_{\pm}(N)\rangle$, which both have a non-zero projection onto $|g, N\rangle$, can therefore be connected via $\boldsymbol{D}$ to $|c, N\rangle$. This means that, because of the presence of an intense radiation exciting the $g \rightarrow e$ transition, the transition $g \rightarrow c$ is split into two lines separated by $\hbar\Omega_R$, called the "Autler-Townes doublet" [55].

[8] This is not Raman scattering as we assume there are no other atomic states except e and g.

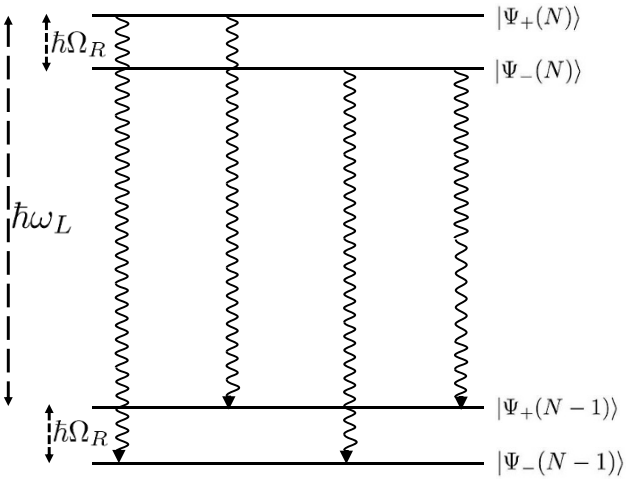


Figure 7: Radiative transitions between one of the two states $|\psi_\pm(N)\rangle$ of multiplicity $\mathcal{E}(N)$ to one of the two states $|\psi_\pm(N-1)\rangle$ of multiplicity $\mathcal{E}(N-1)$, whose energy is lower than that of $\mathcal{E}(N)$ by the quantity $\hbar\omega_L$. We assume the exciting radiation to be exactly at resonance, so that the energy interval between the two states $|\psi_\pm\rangle$ of each multiplicity is equal to $\hbar\Omega_R$.

3-d. Temporal correlations between fluorescent photons

We now study the characteristics of the radiation spontaneously emitted by an atom that interacts continuously with the electromagnetic field of a laser.

α. *Radiative cascade of the dressed atom*

We saw that a dressed atom, spontaneously emitting a photon, goes from the multiplicity $\mathcal{E}(N)$ to the one just below, $\mathcal{E}(N-1)$, located at an energy distance $\hbar\omega_L$. We shall not study here the precise evolution of the physical system as it leaves multiplicity $\mathcal{E}(N)$, which requires the master equation, already mentioned in § 1-e. Our discussion will remain qualitative, but the interested reader will find a more detailed approach in § D of Chapter VI in [21].

Once it reaches $\mathcal{E}(N-1)$, the atom can spontaneously emit a new photon, which brings the dressed atom to $\mathcal{E}(N-2)$, and so on. The series of photons spontaneously emitted by the atom in continuous interaction with the laser radiation can be viewed as a "radiative cascade" of the dressed atom descending its energy diagram.

This image of a radiative cascade permits studying the time correlations between the photons emitted by the atom. As we shall see, the observed correlations depend on the spectral resolution of the photodetectors used.

β. High spectral resolution photodetection

With a high enough spectral resolution, one can observe the time-correlations between photons emitted in the two side-bands of the triplet. Suppose we place filters in front of the photodetectors, so that each can receive only one of the components of the fluorescent triplet, centered at frequencies ω_L, $\omega_L + \Omega_R$ and $\omega_L - \Omega_R$. This means that the spectral resolution of the apparatus is better than the splitting frequency Ω_R of this triplet, but it does not imply that it is lower than the natural width Γ of each components. If we call $\delta\omega$ this spectral resolution, we then have:

$$\Gamma < \delta\omega < \Omega_R \tag{41}$$

Imagine that at a given time, a detector registers a photon emitted for example in the lateral band centered at $\omega_L + \Omega_R$, as the system undergoes the transition $|\Psi_+(N)\rangle \rightarrow |\Psi_-(N-1)\rangle$ (curly arrow on the left-hand side of the Figure 7). The next photon is emitted as the system, starting from $|\Psi_-(N-1)\rangle$, undergoes either the $|\Psi_-(N-1)\rangle \rightarrow |\Psi_-(N-2)\rangle$ transition, emitting a photon of frequency ω_L, or the $|\Psi_-(N-1)\rangle \rightarrow |\Psi_+(N-2)\rangle$ transition, emitting a photon of frequency $\omega_L - \Omega_R$. This means that a second photon with the same frequency $\omega_L + \Omega_R$ as the first one, cannot be emitted right after the first one.

If the frequency of that second photon is ω_L, the system ends up in state $|\Psi_-(N-2)\rangle$; from that state, it can emit either a third photon with frequency ω_L, or a third photon with a lower frequency $\omega_L - \Omega_R$. If, on the other hand, the frequency of that second photon is $\omega_L - \Omega_R$, the system ends up in state $|\Psi_+(N-2)\rangle$; from that state, it can emit either a photon with a higher frequency $\omega_L + \Omega_R$, or a photon of frequency ω_L. As opposed to the second photon, the third photon may thus have the same frequency $\omega_L + \Omega_R$ as the first one. Following the same line of reasoning one can argue that, if the first photon has a frequency $\omega_L - \Omega_R$, this cannot be the case for the second photon; one must wait until the third photon to eventually obtain the same frequency $\omega_L - \Omega_R$. This means that, if photons with only the two extreme frequencies $\omega_L \pm \Omega_R$ are selectively observed, the detected emission processes will necessarily alternate in time (but these events may be separated by any number of photon emissions at the central frequency ω_L).

Comment:

Taking a Fourier transform to return to the time domain, relation (41) imposes a limit to the temporal resolution δt of the detection system: $\delta t \geq 1/\Omega_R$. This means that it is not possible to measure the exact time at which a photon is emitted with a precision of the order of the Rabi precession period.

γ. Photodetection with high temporal resolution

We now study the opposite case where the detectors have a temporal resolution better than the Rabi precession period. This allows a precise determination of the time at which the photon is emitted, but one can no longer distinguish the frequencies of the three triplet components.

We have seen above (§ 1-e) that the elementary spontaneous emission processes have a correlation time that is very short compared to all the other characteristic times

of the problem (because of the large spectral width $\Delta\omega$ of the empty modes' reservoir). A spontaneous emission from $\mathcal{E}(N)$ thus corresponds to a very short "quantum jump" taking the system from state $|e, N-1\rangle$ in $\mathcal{E}(N)$ into state $|g, N-1\rangle$ in $\mathcal{E}(N-1)$. Once the atom has reached this second state, it cannot emit a second photon right away, since no spontaneous emission can occur from a ground state g. A certain time must elapse for the atom-laser interaction to bring the system from state $|g, N-1\rangle$ to the state $|e, N-2\rangle$ it is coupled with, and from which another photon can be spontaneously emitted. The system then falls back to state $|g, N-2\rangle$, and the previous process repeats itself (with an N value lowered by one unit). It therefore becomes clear why one observes a "temporal antibunching" of the photons emitted by a single atom, as they must be separated by a time interval at least of the order of $1/\Omega_R$; this antibunching was already referred to in § 3-b-γ of Complement B_{XX}.

4. Modifications of the field. Dispersion and absorption

We now study how the field is modified by its interaction with the atom.

4-a. Atom in a cavity

The atom-field interaction does not solely perturb the atom; it also changes the field. In order to study this effect, it is convenient to imagine the atom being placed in a real cavity assumed to be perfect, meaning that its losses can be ignored (they occur on a time scale much longer than all the other relevant times of the experiment). As opposed to what we did before, we shall keep the N dependence of the Rabi frequencies Ω_N given by equations (8) and (9), since in a cavity the photon number N has a physical meaning (§ 1-d).

Figure 8 shows the first multiplicities $\mathcal{E}(N)$ of the system atom + field for low and increasing values of the photon number N, starting at $N = 0$. Multiplicity $\mathcal{E}(0)$

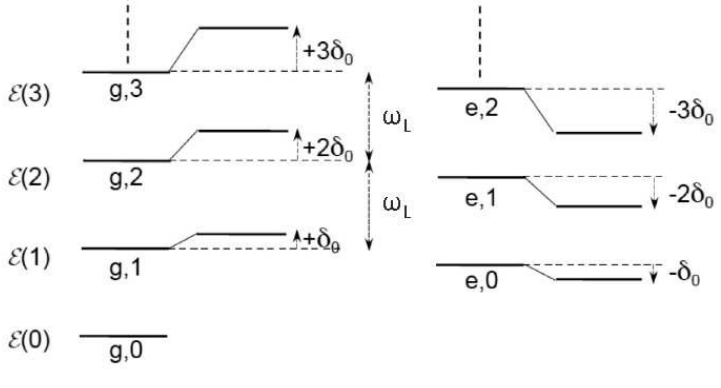


Figure 8: Energy levels of the system atom + field for low values of the photon number N (in angular frequency units, meaning the energies are divided by $\hbar$). States $|g, N\rangle$ and $|e, N-1\rangle$ of $\mathcal{E}(N)$ undergo opposite shifts, proportional to N. State $|g, 0\rangle$ is not shifted.

contains a single state $|g,0\rangle$. Multiplicity $\mathcal{E}(1)$ contains the two states $|g,1\rangle$ and $|e,0\rangle$. Multiplicity $\mathcal{E}(2)$ contains the two states $|g,2\rangle$ and $|e,1\rangle$, and so on.

We shall assume we are in the weak coupling regime, so that we can use the perturbative results of § 2 for the light shifts of the different energy levels. State $|g,0\rangle$ is not shifted as it is not coupled to any other state[9]. States $|g,1\rangle$ and $|e,0\rangle$ of $\mathcal{E}(1)$ undergo opposite light shifts, respectively $+\hbar\delta_g$ and $-\hbar\delta_g$, where δ_g is given by equation (27), where we have replaced Ω_R by Ω_1, the Rabi frequency for $N=1$ (see (8)). Setting:

$$\delta_0 = \Omega_1^2 \frac{\delta}{4\delta^2+\Gamma^2} \tag{42}$$

the light shifts of states $|g,1\rangle$ and $|e,0\rangle$ are, respectively, $+\hbar\delta_0$ and $-\hbar\delta_0$. According to (8), the squares of the Rabi frequencies Ω_N characterizing the atom-field coupling in the multiplicities $\mathcal{E}(N)$, are proportional to N; this means that states $|g,2\rangle$ and $|e,1\rangle$ of $\mathcal{E}(2)$ undergo light shifts respectively equal to $+2\hbar\delta_0$ and $-2\hbar\delta_0$. More generally, states $|g,N\rangle$ and $|e,N-1\rangle$ of $\mathcal{E}(N)$ undergo shifts respectively equal to $+N\hbar\delta_0$ and $-N\hbar\delta_0$.

A similar reasoning can be applied to the radiative broadening. It shows that the radiative broadenings of states $|g,N\rangle$ and $|e,N-1\rangle$ of $\mathcal{E}(N)$ are respectively equal[10] to $+N\gamma_0$ and $\Gamma - N\gamma_0$, where γ_0 is given, according to (27), by:

$$\gamma_0 = \Omega_1^2 \frac{\Gamma}{4\delta^2+\Gamma^2} \tag{43}$$

4-b. Frequency shift of the field in the presence of the atom

Consider the left column in Figure 8. The gap between the perturbed energies of states $|g,1\rangle$ and $|g,0\rangle$ is equal to $\hbar(\omega_L+\delta_0)$; the gap between the perturbed energies of states $|g,2\rangle$ and $|g,1\rangle$ is equal to $\hbar(\omega_L+2\delta_0-\delta_0)=\hbar(\omega_L+\delta_0)$, and so on. As the light shifts of the states $|g,N\rangle$ are proportional to $N\delta_0$, increasing linearly with N, the perturbed levels in the left column of Figure 8 have a constant gap between them, even in the presence of the coupling; the distance between consecutive levels simply goes from $\hbar\omega_L$ to $\hbar(\omega_L+\delta_0)$. In other words, the presence in the cavity of an atom in its ground state g changes the field frequency from ω_L to $\omega_L+\delta_0$. As the light shifts of the levels in the right column of Figure 8 have an opposite sign, a similar argument shows that the presence in the cavity of an atom in the excited state e changes the field frequency from ω_L to $\omega_L-\delta_0$.

The atom-field interaction thus shifts the field frequency inside the cavity by a quantity that changes sign, depending on whether the atom is in the internal state g or e. Let us assume this interaction lasts a time T, as will be the case if an atom is introduced into the cavity and takes that time to traverse it. Compared to the free oscillation in the absence of the atom, the field oscillation will be out of phase by an amount ϕ; this phase shift is equal to $\phi=+\delta_0 T$ if the atom is in state g, and to $\phi=-\delta_0 T$ if the atom is in state e.

This change in the field frequency is a phenomenon similar to that described by the real part of the refractive index: a light beam going through an atomic media changes

[9] There is actually a coupling between state $|g,0\rangle$ and state $|e,1\rangle$, but is highly non-resonant; we shall ignore it since, as stipulated above, our computation is to zeroth order in Ω_R/ω_L.

[10] State $|g,0\rangle$ does not undergo any radiative broadening.

its propagation velocity without any changes in its frequency. In a cavity, the field wavelength cannot change as it is fixed by the boundary conditions on the cavity walls, and hence by the cavity size. The phase shift compared to the free evolution cannot accumulate in space but must accumulate in time (resulting in a frequency change of the field). Note that if one varies the ω_L frequency of the field around the atomic frequency ω_0, the sign change of the light shift δ_0 is reminiscent of the sign change of the real part of the refractive index in the vicinity of an atomic resonance[11].

4-c. Field absorption

Consider an atom in its ground state g in the presence, at time $t = 0$, of an ω_L mode of the field in a quasi-classical coherent state $|\alpha\rangle$ (Complement G_V). The state of the total system reads:

$$|\psi(0)\rangle = |g\rangle \otimes |\alpha\rangle = \sum_{N=0}^{\infty} \frac{\alpha^N}{\sqrt{N!}}\, e^{-|\alpha|^2/2}\, |g, N\rangle. \tag{44}$$

The time evolution in the presence of coupling changes the energies of the states $|g, N\rangle$ to:

$$\tilde{E}_N = N\hbar(\omega_L + \delta_0 - i\gamma_0/2). \tag{45}$$

and the state of the system at time t becomes:

$$\begin{aligned} |\psi(t)\rangle &= \sum_{N=0}^{\infty} \frac{\alpha^N}{\sqrt{N!}} e^{-|\alpha|^2/2} \exp[-iN(\omega_L + \delta_0 - i\gamma_0/2)t]\ |g, N\rangle \\ &\propto |g\rangle \otimes |\alpha \exp[-i(\omega_L + \delta_0 - i\gamma_0/2)t]\rangle. \end{aligned} \tag{46}$$

The atom is still in the presence of a quasi-classical coherent state. However, compared to the free field evolution of that state in the absence of coupling, the atom-field interaction has introduced a phase shift $\delta_0 t$ (as already discussed above) as well as a decrease in amplitude $e^{-\gamma_0 t/2}$, resulting in an attenuation of the amplitude of the field. This is reminiscent of the radiation absorption described by the imaginary part of the refractive index.

To sum up, we showed that the atom-field coupling produces light shifts and radiative broadening of the atomic levels, corresponding to the well-know field dispersion and absorption phenomena in optics.

Conclusion.

In conclusion, we showed for many various situations that the dressed-atom approach brings strong clarifications while keeping the calculations simple. Considering the atom and the field mode with which it interacts as a quantum system described by a time-independent Hamiltonian allows introducing true energy levels for the global system; this leads to a new, broad overview of the stimulated absorption and emission of photons.

As an example, this approach makes it very clear how the atom-photon coupling changes the energy diagram of the dressed atom at high field intensity; this leads to a

[11] This effect is sometimes referred to as "*anomalous dispersion*."

very simple interpretation of the new frequencies appearing in the atomic fluorescence spectrum in the strong coupling domain. As the energy diagram of the dressed atom is a succession of multiplicities separated by an energy equal to a photon energy, the spontaneous emission of a photon is viewed in this approach as a quantum jump of the dressed atom from one multiplicity to the one just below (radiative cascade). This approach allows a simple calculation of the delay function yielding the distribution of the time intervals between two successive quantum jumps; this permits studying the time correlations between fluorescent photons. Let us also mention that this delay function allows simulating the temporal evolution of an atom, hence obtaining individual quantum trajectories, which can be used to get an averaged atomic evolution. Several experimental applications of the dressed-atom method are presented in the next complement.

Complement D_{XX}

Light shifts: a tool for manipulating atoms and fields

1 Dipole forces and laser trapping 2151
2 Mirrors for atoms 2153
3 Optical lattices 2153
4 Sub-Doppler cooling. Sisyphus effect 2155
4-a Laser configurations with space-dependent polarization . . . 2156
4-b Atomic transition 2156
4-c Light shifts 2156
4-d Optical pumping 2157
4-e Sisyphus effect 2158
5 Non-destructive detection of a photon 2159

Light shifts studied in § 2-b of Complement C_{XX} exhibit a number of important properties leading to numerous applications; these will be briefly discussed in this complement.

As these light shifts are proportional to the laser intensity, their magnitude can be space-dependent if the laser intensity is not homogeneous in space. These shifts can be used to create either potential wells (§ 1) to trap atoms once they are cold enough (laser trapping), or potential barriers (§ 2) reflecting atoms (mirrors for atoms). A particularly interesting example involves periodic optical potential wells created at the nodes and antinodes of a laser standing wave in an off-resonant condition (§ 3). This situation is reminiscent of that encountered by electrons trapped in the periodic potential of a crystal lattice. Neutral atoms trapped in optical lattices can thus serve as models for condensed matter problems.

For low enough values of the detuning δ between the laser frequency and the atomic frequency, and if the ground state has several Zeeman sublevels, non-dissipative effects, such as light shifts, can coexist with dissipative effects, such as optical pumping between Zeeman sublevels. We explain in § 4 how correlations between these two types of effects can lead to new cooling mechanisms, such as Sisyphus cooling, allowing the atoms to reach temperatures much lower than with Doppler cooling.

Finally, we show in § 5 how the light shifts undergone by an atom crossing a highly detuned cavity allows determining the number of photons present in the cavity, by performing measurements on the atoms at the cavity exit, without absorbing any of the cavity photons.

1. Dipole forces and laser trapping

When the light intensity varies in space, as with a focalized laser beam or a standing wave, the light shifts also become space-dependent. If the detuning δ between the laser frequency and the atomic frequency is large compared to the natural width Γ of the excited level, it is then justified to ignore the dissipation due to spontaneous emission, on

the characteristic time scales of the experiment. The light shift $\hbar\delta_g(\boldsymbol{R})$ of ground state g depends, as does the light intensity, on the position $\boldsymbol{R}$ of the atomic center of mass; it can therefore be considered as a potential energy $V(\boldsymbol{R}) = \hbar\delta_g(\boldsymbol{R})$ that affects the atomic motion. This potential has the same sign as the light shift, and hence depends on the sign of the frequency detuning δ.

The potential $V(\mathbf{r})$ gives rise to a force:

$$\mathbf{F}_{\text{dip}}(\mathbf{R}) = -\boldsymbol{\nabla}_{\mathbf{R}} V(\mathbf{R}) \tag{1}$$

called the "dipole force", or sometimes the "reactive force" (§ 11-4 in [24]). It is different from the radiation pressure forces studied in § 1-d of Complement A_{XIX}, which come from momentum exchanges as the atom absorbs photons that are spontaneously reemitted. The dipole forces introduced here arise from the spatial variations of the light shifts undergone by the dressed-atom levels. One could say they are caused by the redistribution of photons between the different plane waves composing the laser beam[1]: the atom absorbs a photon from one plane wave and re-emits it, by stimulated emission, in another plane wave; this process changes the atom's momentum, and hence giving rise to a force.

Comment:

As is the case for light shifts, the intensity of the dipole forces, as a function of the frequency detuning δ between the laser frequency and the atomic frequency, follows a dispersion curve. In addition, the light shifts of the two dressed levels in multiplicity $\mathcal{E}(N)$ have an opposite sign for a given detuning δ; the dipole force thus changes sign from one dressed state $|\Psi_+(N)\rangle$ to its associated state $|\Psi_-(N)\rangle$. When the detuning δ is not too large, and if spontaneous emission processes can occur, the dressed-atom radiative cascade can lead to sign changes of the dipole force, as the atom goes from states $|\Psi_\pm(N)\rangle$ to $|\Psi_\mp(N-1)\rangle$; this is the origin of the fluctuations of the dipole forces.

An important application of dipole forces is the implementation of laser traps. Consider first a laser beam detuned toward the red ($\omega_L < \omega_0$) and focalized at point O. The light shift, zero outside the laser beam, is negative inside the laser beam; it increases in absolute value as one gets closer to the focal point, where it reaches its maximum value. This creates a potential well that could trap a neutral atom; this will indeed happen if the atom's kinetic energy, of the order of $k_B T$, is lower than the depth U_0 of the potential well. This is why these laser traps have been built only since the 1980's, once atomic cooling techniques (Complement A_{XIX}, § 2) allowed slowing down atoms to temperatures of the order of a microkelvin [56].

Comment:

The trapping forces involved in laser traps are of the order of an atomic dipole multiplied by the gradient of a laser field. They are much weaker than the forces exerted by a static electric field on a charged particle. This explains why laser traps for neutral atoms are much shallower than ion or electron traps. There exist, however, other types of traps for

[1] A single plane wave does not have an intensity gradient, and cannot exert a dipole force. These forces, due to intensity gradients, require the presence of several plane waves with different wave vectors.

neutral atoms, using different physical mechanisms (for a short review, see for example § 2-c of Complement A_{XIX}, and Chapter 14 of reference [24]).

2. Mirrors for atoms

A laser detuned toward the blue ($\omega_L > \omega_0$) gives rise to repulsive potentials. Imagine for example that the laser wave propagates within a bloc of glass, and undergoes total internal reflection at the boundary between the glass and the vacuum (Figure 1-a). An evanescent wave appears outside the glass, with an amplitude decaying exponentially in a direction perpendicular to the boundary, becoming negligible over a distance of the order of the laser wavelength. This evanescent wave creates a potential barrier of height U_0, which reflects atoms[2] arriving with an energy $E < U_0$ (Figure 1-b). This set-up can be used as a mirror for neutral atoms [57].

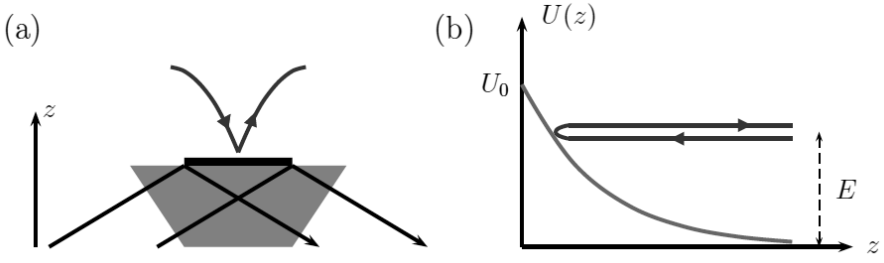


Figure 1: (a) A laser beam traveling within a block of glass (shaded in grey in the figure) undergoes total internal reflection at the boundary. Outside the glass, an evanescent wave appears. (b) If the laser is detuned toward the blue ($\omega_L > \omega_0$), this evanescent wave creates a potential barrier of height U_0. Atoms falling on this barrier with an energy E lower than U_0 are reflected by the barrier and turn around.

3. Optical lattices

When laser beams form a standing wave, the light intensity is modulated in space, with a periodicity $\lambda/2$: the intensity is zero at the nodes, and maximal at the antinodes. This creates periodic potential wells, located at the antinodes of the wave for a negative detuning ($\omega_L < \omega_0$), and at the nodes for a positive detuning ($\omega_L > \omega_0$). Figure 2 shows a two-dimensional optical lattice created by two standing waves, along two orthogonal axes[3].

The study of optical lattices is interesting for several reasons, in particular because the motion of a neutral atom in an optical lattice is reminiscent of that of an electron in

[2] Atoms falling on a solid surface would stick to it, rather than being reflected.

[3] The frequencies ω_{L1} and ω_{L2} of the two standing waves are in general sufficiently far apart for the interference terms between the two waves to have a negligible effect on the atom's motion; the potentials created by the two waves can then be independently added. This requires $|\omega_{L1} - \omega_{L2}|$ to be large compared to all the characteristic frequencies of the atom's motion, such as it's vibrational motion inside a well; in that case, the interference terms oscillate too fast to have a significant effect.

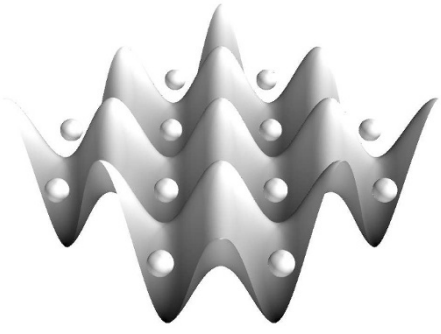


Figure 2: Schematic representation of a two-dimensional optical lattice: placed in a superposition of two standing laser waves along two orthogonal axes, the atom is subjected to a potential periodic in space, represented by the undulating surface in the figure. These periodic potential wells, located at the antinodes of the standing waves for $\omega_L < \omega_0$, and at the nodes for $\omega_L > \omega_0$, form an optical lattice. The spheres above the surface indicate the positions where the atoms can be trapped.

a crystal lattice. Granted the order of magnitudes involved are quite different, since the spatial period of an optical lattice is of the order of a micron, whereas the period of a crystal lattice is of the order of a fraction of a nanometer. Nevertheless, optical lattices offer a large number of possibilities not available to crystalline lattices:

- One can easily change the intensity of the laser waves forming the standing waves, hence modifying the depth of the potential wells; this allows controlling the tunnel effect between adjacent wells. This method was used to explore the transition between a deep well regime where the atoms are localized at the bottom of the wells, and a shallow well regime where the atoms' wave functions are delocalized over the entire lattice [58].

- One can abruptly switch off the trapping laser beams (which obviously cannot be done for a crystal lattice) and study the resulting behavior of the liberated atoms. Studying the expansion velocity of the clouds of atoms yields information on their velocity distribution, and hence on their temperature (time-of-flight method). Studying their spatial distribution and the possible appearance of a diffraction pattern allows determining whether the matter waves trapped in distinct potential wells of the optical lattice were coherent or not.

- One can use two different frequencies ω_{L1} and ω_{L2} for the two laser waves counterpropagating to form the one-dimensional standing laser wave. This leads to a "standing" laser wave, moving with constant velocity if $\omega_{L1} - \omega_{L2}$ is fixed, or with an acceleration if $\omega_{L1} - \omega_{L2}$ increases linearly with time. In this latter case, the atom experiences a constant inertial force in the rest frame of the standing wave; its motion is then similar to that of an electron in a crystal lattice periodic potential, subjected in addition to a static electric field. The motion of such a particle, in a periodic lattice and subjected to a constant force, is predicted to be oscillatory,

following the so-called *Bloch oscillations*; the experimental observation of such oscillations is facilitated in an optical lattice, as the atom's relaxation time can be much longer than the oscillations' period [59].

Cold atoms trapped in an optical lattice are a model system for "simulating" a number of situations encountered in solid state physics. Cold atom studies involve interactions between atoms much weaker than the Coulomb interactions between electrons. Furthermore, they can be controlled thanks to resonance effects occurring as atoms collide with each other.

Note finally that optical lattices are a good example highlighting the importance of light shifts. One may wonder if it might not be simpler to shift atomic levels by Zeeman or Stark effects in static magnetic or electric fields, rather than using an off-resonance light beam to produce a light shift. The advantage of the light shifts is that they can be used to form potentials varying over very short distances, of the order of an optical wavelength, which is much more difficult to attain with static fields.

4. Sub-Doppler cooling. Sisyphus effect

We described, in § 2-b of Complement A_{XIX}, a cooling mechanism for the atoms, based on the Doppler effect, and called for that reason "Doppler cooling". We computed the friction and diffusion coefficients associated with that mechanism and showed that the lowest temperature T_D that could be reached by Doppler cooling was of the order of $\hbar\Gamma/k_B$ (where Γ is the natural width of the atoms' excited states, and k_B the Boltzmann constant). Actually, the first measurements of the temperatures reached by laser cooling, and based on the time-of-flight method [60], showed that temperatures much lower than T_D could be obtained; furthermore, their dependence on the detuning δ between the laser beams' frequency and the atomic frequency did not follow the prediction of the Doppler cooling theory. This implied the existence of other cooling mechanisms for the atoms, leading to temperatures lower than the Doppler limit T_D; as expected, these mechanisms were called "sub-Doppler" mechanisms. One of them, called the "Sisyphus effect", will be described in this section.

The theory of Doppler laser cooling, exposed in § 2-b of Complement A_{XIX}, does not take into account several important characteristics of laser cooling experiments.

- In most experiments performed in three-dimensional space, the polarization of the laser field cannot be uniform. The spatial variations of this polarization should not be ignored.

- The atoms under study have several sublevels, in the lower g state and in the excited e state. The two-level atom approximation of § 2-b in Complement A_{XIX} is therefore not sufficient.

- As there are several sublevels in the lower state g, one should include the effects of the optical pumping between these sublevels, effects whose characteristic time constants (pumping times) are longer than the lifetime $1/\Gamma$ of the excited state.

- As the detuning δ between the laser beams' frequency and the atomic frequency is different from zero, one must take into account the light shifts of the lower level g, which can take on different values for the different sublevels.

Before describing the Sisyphus effect, we first show on a simple example how these different effects come into play.

4-a. Laser configurations with space-dependent polarization

Laser configurations with a space-dependent polarization do not necessarily involve three pairs of laser beams counterpropagating along the Ox, Oy and Oz axes. They can be achieved in one-dimension, and are easier to study, as long as the two counterpropagating laser waves have different polarizations. As an example, Figure 3 represents two laser waves propagating in opposite directions along the Oz axis and having linear orthogonal polarizations $\mathbf{e}_x$ and $\mathbf{e}_y$. The polarization of the total field changes from right-hand circularly polarized (σ_+ with respect to the quantization axis Oz) to left-hand circularly polarized (σ_-) in planes separated by a distance $\lambda/4$, and is linear at $\pm 45°$ of the Ox and Oy axes, half-way between these planes.

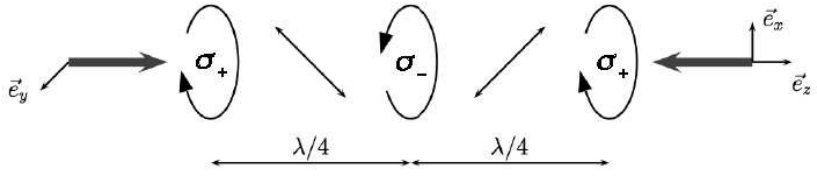


Figure 3: Laser configuration with a space-dependent polarization: two laser waves propagate in opposite directions along the Oz axis, having linear orthogonal polarizations $\mathbf{e}_x$ *and* $\mathbf{e}_y$.

4-b. Atomic transition

Many of the laser cooling experiments use transitions between a lower state g with angular momentum J_g and an excited state e with an angular momentum equal to $J_e = J_g + 1$. Here, we shall consider the simplest possible case $J_g = 1/2$, where the lower state contains only 2 sublevels $g_{\pm 1/2}$. We then have $J_e = 3/2$ and the excited state has 4 sublevels $e_{\pm 1/2}$ and $e_{\pm 3/2}$.

4-c. Light shifts

Consider first a point in space where the laser field polarization is σ_+ (with respect to the quantization axis Oz). We saw in § 1-b of Complement C_{XIX} that photons with a σ_+ (σ_-) polarization have a spin angular momentum $+\hbar$ ($-\hbar$) along the Oz axis. Conservation of total angular momentum in the photon absorption process leads to the selection rule $m_e - m_g = +1(-1)$ for the absorption of a $\sigma_+(\sigma_-)$ photon, where m_e and m_g are the magnetic quantum number of the states involved in the transition. Figure 4 represents the 2 transitions $g_{-1/2} \leftrightarrow e_{+1/2}$ and $g_{+1/2} \leftrightarrow e_{+3/2}$ ($m_e - m_g = +1$) that can be excited by the laser field. The numbers 1/3 and 1 shown next to these 2 transitions are the squares of the Clebsch-Gordan coefficients of these transitions (Complement B_X); they indicate that the $g_{+1/2} \rightarrow e_{+3/2}$ transition is 3 times more intense than the $g_{-1/2} \leftrightarrow$

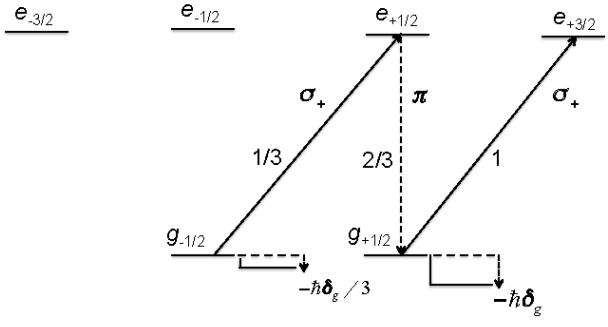


Figure 4: Transition $1/2 \leftrightarrow 3/2$. *The oblique upwards arrows show the transitions excited at a point where the laser field polarization is* σ_+, *the vertical downward arrow indicates the spontaneous emission transition from sublevel* $e_{+1/2}$ *toward sublevel* $g_{+1/2}$. *The light shifts of states* $g_{-1/2}$ *and* $g_{+1/2}$ *are noted* $-\hbar\delta_g/3$ *and* $-\hbar\delta_g$. *At a point where the laser polarization becomes* σ_- *instead of* σ_+, *the shifts of the two sublevels are interchanged, by symmetry.*

$e_{+1/2}$ transition. As the detuning δ between the laser beams' frequency and the atomic frequency is negative in a laser cooling experiment, both states $g_{\pm 1/2}$ have a negative light shift, but with a modulus 3 times larger for state $g_{+1/2}$ than for state $g_{-1/2}$. These light shifts are written in the figure as $-\hbar\delta_g$ and $-\hbar\delta_g/3$, where δ_g is positive.

At a point in space where the polarization is σ_-, the previous results are interchanged. It is now the $g_{-1/2} \leftrightarrow e_{-3/2}$ transition that is 3 times more intense than the $g_{+1/2} \leftrightarrow e_{-1/2}$ transition, yielding light shifts equal to $-\hbar\delta_g$ and $-\hbar\delta_g/3$ for the states $g_{-1/2}$ and $g_{+1/2}$, respectively.

Finally, at a point where the polarization is linear, the two light shifts are identical for symmetry reasons, and proportional to the square of the Clebsch-Gordan coefficient (equal to 2/3), indicated in the figure for the $g_{+1/2} \leftrightarrow e_{+1/2}$ transition. Consequently, they are both equal to $-2\hbar\delta_g/3$.

This means that as one moves along z axis, the positions of the 2 Zeeman sublevels $g_{-1/2}$ and $g_{+1/2}$ oscillate, with opposing phases, between the values $-\hbar\delta_g$ and $-\hbar\delta_g/3$ (taking the energy of the unperturbed ground state equal to zero).

4-d. Optical pumping

Let us focus on a point where the laser field polarization is σ_+ and there is an atom in state $g_{+1/2}$. The atom can absorb a σ_+ photon and end up in state $e_{+3/2}$. From this state, it can only fall, by spontaneous emission, back to its initial state $g_{+1/2}$; optical pumping (§ 1-b of Complement C_{XIX}) does not lead, in this case, to any population change. On the other hand, if the atom is initially in state $g_{-1/2}$ and absorbs a photon σ_+ that brings it to state $e_{+1/2}$, it can then fall back, by spontaneous emission, into state $g_{+1/2}$; optical pumping takes place from the least shifted sublevel $g_{-1/2}$ towards the most shifted sublevel $g_{+1/2}$. A comparable situation is found at a point where the laser field polarization is σ_-. Optical pumping can only occur from the least shifted sublevel $g_{+1/2}$ toward the most shifted sublevel $g_{-1/2}$. As for a point where the laser field polarization

is linear, since the Clebsch-Gordan coefficients of the $g_{-1/2} \leftrightarrow e_{-1/2}$ and $g_{+1/2} \leftrightarrow e_{+1/2}$ transitions are equal, as are those of the $e_{-1/2} \leftrightarrow g_{+1/2}$ and $e_{+1/2} \leftrightarrow g_{-1/2}$ transitions, optical pumping cannot favor one of the the populations of the 2 sublevels $g_{-1/2}$ and $g_{+1/2}$. To sum up, optical pumping can only transfer population from the least shifted sublevel to the most shifted sublevel, with a maximum efficiency at points where the laser field polarization is circular.

4-e. Sisyphus effect

We now show how the correlations between the light shifts and the optical pumping effects studied in the last two sections can reduce the atom's kinetic energy, and hence cool it down.

Figure 5 shows, for an atom moving along the z axis, the energies of its 2 sublevels $g_{-1/2}$ and $g_{+1/2}$, shifted by the light. Let us assume the atom starts from the bottom of a potential valley, at a point where the laser field polarization is σ_+, and is initially in its most shifted state $g_{+1/2}$. As it moves toward the right, it climbs a potential well, and looses some kinetic energy. If the optical pumping time is long enough, it will have time to reach the top of the hill, where the laser field polarization is σ_-; it then has a high probability to undergo an optical pumping cycle and be transferred to the most shifted sublevel, which is now sublevel $g_{-1/2}$. The whole cycle we just described can repeat itself, and each time the kinetic energy of the atom is lowered by a quantity of the order of the maximum energy difference between the two sublevels in Figure 5, equal to $(2/3)\hbar\delta_g$. The atom is facing a situation similar to that of the hero of Greek mythology, Sisyphus: it must endlessly climb a potential hill since it is sent back to the bottom as soon as it reaches the top, hence the name Sisyphus effect given to this mechanism.

The temperature reached by such a mechanism can be estimated by a simple

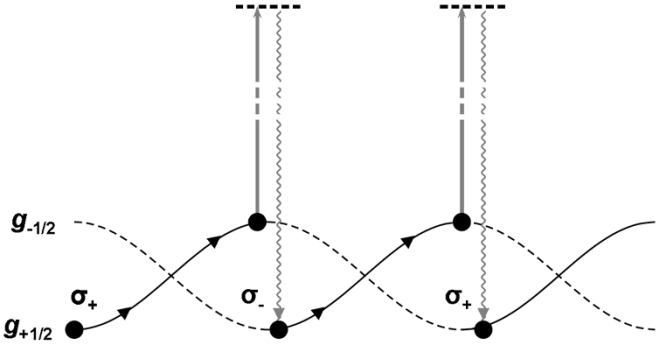


Figure 5: Principle of Sisyphus cooling: an atom in state $g_{+1/2}$ moving from a point where the laser field polarization is σ_+ must climb a potential hill of height $2\hbar\delta_g/3$, which decreases its kinetic energy. When it reaches the top of the hill, where the laser field polarization is σ_-, it has a strong probability to fall back, by spontaneous emission, to the state $g_{-1/2}$. As the cycle repeats itself, the atom is for ever climbing potential hills, like the hero Sisyphus in Greek mythology. Its kinetic energy diminishes constantly.

argument. The atom's kinetic energy decreases during each Sisyphus cycle, until it is low enough for the atom to be trapped at the bottom of a potential well. At that point its kinetic energy must be of the order of $\hbar\delta_g$. The ultimate temperature T_S that can be reached by Sisyphus cooling is expected to be $T_S \simeq \hbar\delta_g/k_B$. In laser cooling experiments, the laser intensities are generally low and the atomic transitions are not saturated; consequently, $\hbar\delta_g \ll \hbar\Gamma$ and hence $T_S \ll T_D$. This explains why the measured temperatures can be two orders of magnitude lower than the Doppler temperature, and reach values of the order of $10^{-6}K$, opening the way to numerous applications.

All these qualitative predictions have been confirmed by more quantitative models, see ([61]) and ([62]). Experiments have confirmed the theoretical predictions, in particular those concerning the dependence of T_S on the various experimental parameters, such as laser intensity and detuning [63].

5. Non-destructive detection of a photon

Consider now an experiment where the atoms of a beam cross, one after the other, a cavity containing radiation whose quantum state is described by a Fock state; the number of photons in the cavity mode is fixed, equal for example to 0 (radiation vacuum) or 1 (single photon). The atoms are prepared in a coherent superposition of the two states g and e:

$$|\psi_{\text{in}}\rangle = \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle) \tag{2}$$

While each atom interacts with the photons, its levels undergo light shifts resulting in different phases for the two atomic states as the atom crosses the cavity; note, however, that if the detuning δ between the laser frequency and the atomic frequency is sufficiently large, no photon will be absorbed or emitted. As the atom exits the cavity, the radiation state is the same initial Fock state, whereas the atomic state is modified by this phase factor. The final atomic state can be written (within a global phase factor of no physical significance):

$$|\psi_{\text{fin}}\rangle = \frac{1}{\sqrt{2}}(|g\rangle + e^{-i\phi}|e\rangle) \tag{3}$$

The phase ϕ is simply the integral over time of the energy difference between the dressed-atom levels that come into play as the atom crosses the cavity. It is given by the energy diagram of the dressed-atom.

Figure 8 of Complement C_{XX} shows that the gap between states $|e\,,0\rangle$ and $|g\,,0\rangle$ is reduced from $\hbar\omega_0$ to $\hbar(\omega_0 - \delta_0)$ by the light shifts. For a cavity with no photons ($N = 0$), when the atom exits the cavity, the coherence between its states e and g has been dephased by:

$$\phi_0 = -\int \delta_0(t')\ \mathrm{d}t' \tag{4}$$

where $\delta_0(t')$ is obtained by replacing in (42) of Complement C_{XX} the Rabi frequency Ω_1 by a function of time that accounts for the motion of the atom in the cavity mode, where it is subjected to a time-dependent light intensity; remember that we assumed the

detuning δ between the atomic frequency and the laser frequency to be large enough for no real photon absorption by the atom to occur[4]. If now the cavity contains one photon ($N = 1$), Figure 8 indicates that the gap between states $|e\,,1\rangle$ and $|g\,,1\rangle$ is reduced by the light shifts to $\hbar(\omega_L - 2\delta_0 - \delta_0)$, i.e. to $\hbar(\omega_L - 3\delta_0)$. When the atom exits the cavity, the coherence between its states e and g is now shifted by three times the amount obtained in (4). This means that an atom, traversing the cavity in a superposition of states g and e, keeps in the phase of that coherent superposition a trace of the number of photons present in the cavity; this occurs without any photon absorption (since the detuning δ is too large). To sum up, if $N = 0$, the state of the atom at the cavity exit is:

$$|\psi_{\text{fin}}(N = 0)\rangle = \frac{1}{\sqrt{2}}(|g\rangle + e^{-i\phi_0}|e\rangle) \tag{5}$$

whereas if $N = 1$, this state is:

$$|\psi_{\text{fin}}(N = 1)\rangle = \frac{1}{\sqrt{2}}(|g\rangle + e^{-i\phi_1}|e\rangle) \tag{6}$$

with $\phi_1 = 3\phi_0$.

How can we make use of this trace left on the atom by the possible presence of a photon in the cavity, and determine if this cavity contains zero or one photon? The time taken by the atom to cross the cavity can be adjusted by changing the atom's speed. Imagine that this time is tuned so that $\phi_0 - \phi_1 = \pi$; this means that the two states (5) and (6) are now orthogonal. As the atom exits the cavity, we can apply to it a $\pi/2$ laser pulse adjusted to transform $|\psi_f(N = 0)\rangle$ into $|g\rangle$. That same pulse will transform $|\psi_f(N = 1)\rangle$ into the state orthogonal to $|g\rangle$, that is to $|e\rangle$. This means that measuring the atomic state after this $\pi/2$ laser pulse allows concluding that $N = 0$ if the atom is found in state g, and that $N = 1$ if the atom is found in state e. The measurement can be repeated several times by sending a stream of atoms, one after the other, and applying to each of them the same procedure; one can measure several times in a row the same value N, which proves the number of photons in the cavity did not change during the measurements. As opposed to photoionization where a photon is absorbed giving rise to a photoelectron (Complement A_{XX}), this method is non-destructive: the presence of the photon is detected without it being absorbed. This experiment, generalized to the case where the photon number is larger than one, is described in more detail in reference [64].

Conclusion.

For a long time, light shifts have been considered as an interesting physical phenomenon without specific applications, and even as an undesired perturbation for high resolution spectroscopy, since they modify the atomic transition frequencies one is trying to measure with the highest possible precision. These shifts must be taken into account to extract from the measurements the non-perturbed frequencies of atomic transitions; most of the

[4]We also assume that the field variation encountered by the atom as it crosses the cavity is slow enough for non-adiabatic transitions from e to g, or from g to e, to be highly improbable. We also suppose that the natural width Γ of the excited state e, and the time T the atom takes to cross the cavity, are small enough for $\Gamma T \ll 1$, meaning spontaneous emission from state e does not have time to occur while the atom crosses the cavity.

time, several measurements at different light intensities must be performed to extrapolate the results to zero light intensity.

This complement clearly shows how much the situation has changed, by presenting the large variety of experimental methods using light shifts of atomic energy levels, and their great number of applications. These methods were implemented more than 20 years after these shifts were theoretically predicted and experimentally demonstrated; this illustrates the long term practical impact of fundamental research. These methods allow acting both on the internal and external atomic variables; they also permit using atoms as a very sensitive non-destructive probe for the properties of a field composed of only a few photons. These methods made it possible to trap atoms in a standing laser wave, or to obtain periodic lattices of neutral atoms trapped in such a wave. It also led to laser cooling methods that allowed reaching temperatures previously totally inaccessible for atomic gases, millions of times lower than the lowest temperatures found in the interstellar or intergalactic space of the Universe.

Complement E_{XX}

Detection of one- or two-photon wave packets, interference

1 One-photon wave packet, photodetection probability 2165
- 1-a Photoionization of a broadband detector 2165
- 1-b Detection probability amplitude 2166
- 1-c Temporal variation of the signal 2167

2 One- or two-photon interference signals 2167
- 2-a How should one compute photon interference? 2168
- 2-b Interference signal for a one-photon wave packet in two modes 2168
- 2-c Interference signals for a product of two one-photon wave packets 2170

3 Absorption amplitude of a photon by an atom 2174
- 3-a Computation of the amplitude 2174
- 3-b Properties of that amplitude 2175

4 Scattering of a wave packet 2176
- 4-a Absorption amplitude by atom B of the photon scattered by atom A 2176
- 4-b Wave packet scattered by atom A 2177

5 Example of wave packets with two entangled photons . . . 2181
- 5-a Parametric down-conversion 2181
- 5-b Temporal correlations between the two photons generated in parametric down-conversion 2183

Introduction

In Chapter XX the initial and final states of the atom + photon(s) system were chosen as states that, in the absence of interaction, had a well defined energy; before the interaction, such states do not evolve in time, as if the photon were not propagating in space. As an example, in the scattering process of a photon by an atom, the chosen initial radiation state is a photon with momentum $\hbar\boldsymbol{k}_i$ and energy $\hbar\omega_i = \hbar ck_i$, which spreads over the entire space; similarly, the final state is also a photon with momentum $\hbar\boldsymbol{k}_f$ and energy $\hbar\omega_f = \hbar ck_f$. The interaction was "turned on" at time t_i, which allowed computing the probability amplitude for the atom + photon(s) system to go from one state to the other between t_i and t_f. This is clearly a phenomenological approach: what actually happens is that the interaction operator remains constant but only comes into play to change the state vector when the atom is in the presence of radiation. A more realistic description of the process should involve the propagation of wave packets, with the incident radiation being described by a wave packet initially very far away from the atom, but going towards it. Their interaction then gives rise to a scattered wave packet moving away towards infinity, while the incident wave packet, modified by the interaction, also continues on its way.

Note however that introducing a wave packet for a photon cannot be done by the standard method used for a massive particle. As already pointed out at the end of § B-2 in Chapter XIX, a photon does not have a position operator. One cannot obtain its spatial wave function by projecting its state vector onto the eigenvectors of that operator, and then squaring this wave function's modulus to get the probability of finding the photon in any given region of space. One could then imagine using the spatial variations of the electric and magnetic fields to infer the photon localization. But for radiation states with exactly one, two, etc. photons, the average value of theses fields at any point in space is zero (it is the sum of zero average value creation and annihilation operators in each mode). Consequently, for a single photon, this average value cannot be directly used for building a wave packet localized in space. This is why we shall use another approach: we shall assume the photon interacts with detectors, well localized in space, and compute the probability of its detection by these apparatus. This will lead us to introduce an amplitude for the photon detection (by photoionization) at a given point, which presents close analogies with the spatial wave function of a massive particle in non-relativistic quantum mechanics.

We start in § 1 by exposing the general idea of this approach; we introduce a function $\mathcal{E}(\boldsymbol{r}, t)$ that allows localizing a single photon in space through its probability of being absorbed by a broadband detector. This leads to the concept of wave packets, even though the average value of the electric field remains zero throughout the entire space. In the perturbative computations, one can also introduce initial and final radiation states that are wave packets, described by linear superpositions of photon states with different momenta and energies.

In § 2, we show how the detection amplitude $\mathcal{E}(\boldsymbol{r}, t)$ allows studying light interference phenomena involving one or two photons. These phenomena are interpreted in terms of interference between the transition amplitudes associated with different paths leading the quantum field from a given initial state to a given final state. Starting first (§ 2-a) with a general discussion of the interference signals, we then focus in § 2-b on the interference involving one photon being simultaneously in two modes of the field. Finally, we examine interference involving two photons in the simple case where the system is described by the product of two one-photon wave packets (§ 2-c).

In § 3, we replace the broadband detector by an atom with discrete energy levels. Without having to assume that the coupling between the atom and the field is turned on abruptly (which is hard to justify from a physical point of view), a number of results of Chapter XX are confirmed with, in addition, the possibility of studying the temporal aspect of the absorption phenomenon. In § 4, we extend this method to study the scattering of photons by an atom. Here again, we will confirm the results of Chapter XX, while enriching our understanding of the temporal aspects of the physical process.

Finally, in § 5, we consider "real" two-photon wave packets that are entangled wave packets. Parametric down-conversion is an example of a situation leading to strong temporal correlations between two entangled photons. Such correlations are impossible to understand in terms of a classical treatment of the radiation.

In this entire complement, we have limited our studies to one- or two-photon wave packets, but the computations can be extended to wave packets containing a larger number of photons[1].

[1] Or even an undetermined number, as in a coherent state.

1. One-photon wave packet, photodetection probability

A one-photon wave packet is described by a state $|\psi\rangle$, where the photon number is precisely equal to one (eigenstate of the photon number operator, with eigenvalue equal to 1). We build this wave packet[2] as a linear superposition of states with different momenta $\hbar\boldsymbol{k}$:

$$|\psi\rangle = \int \mathrm{d}^3k\, c(\boldsymbol{k})\, a^\dagger_{\boldsymbol{k}}|0\rangle = \int \mathrm{d}^3k\, c(\boldsymbol{k})\, |\boldsymbol{k}\rangle \tag{1}$$

State $|\psi\rangle$ is not stationary (it is not an energy eigenstate). It is of the type considered in § B-3-c of Chapter XIX, an eigenket of operator $\hat{N}$ (total number of photons) with an eigenvalue equal to one. It is assumed to be normalized:

$$\int \mathrm{d}^3k\, |c(\boldsymbol{k})|^2 = 1, \tag{2}$$

which allows interpreting $|c(\boldsymbol{k})|^2$ as the probability density for the photon momentum to be equal to $\hbar\boldsymbol{k}$.

1-a. Photoionization of a broadband detector

Imagine we place an atom playing the role of a photodetector at point $\boldsymbol{r}$ in the radiation field described by state (1). According to relation (26) of Complement $\mathrm{B_{XX}}$, the probability $W_I(\boldsymbol{r},t)\mathrm{d}t$ for observing a photoelectron emission between times t and $t+\mathrm{d}t$ is given (to the interaction's lowest order) by:

$$W_I(\boldsymbol{r},t) = s\langle\psi|E^{(-)}(\boldsymbol{r},t)E^{(+)}(\boldsymbol{r},t)|\psi\rangle \tag{3}$$

where s is a constant depending on the photodetector sensitivity. $E^{(-)}(\boldsymbol{r},t)$ and $E^{(+)}(\boldsymbol{r},t)$ are the negative and positive frequency components of the electric field operator appearing in its plane wave expansion as given by (A-7) in Chapter XIX:

$$E^{(+)}(\boldsymbol{r},t) = \left[E^{(-)}(\boldsymbol{r},t)\right]^\dagger = i\int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}}\,\sqrt{\frac{\hbar\omega_k}{2\epsilon_0}}\; a(\boldsymbol{k})\, e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)} \tag{4}$$

with:

$$\omega_k = c\times|\mathbf{k}| = c\times k \tag{5}$$

where c is the speed of light. Remember that expression (3) was established in the interaction representation, where the state vector $|\psi\rangle$ evolves only under the effect of the atom-radiation interaction; the operators evolve freely only under the effect of the atomic or radiation Hamiltonians (i.e. without mutual interaction). However, as we are performing a computation to lowest order, we can consider in (3) that $|\psi\rangle$ is actually

[2]For the sake of simplicity, we ignore in this complement the degrees of freedom of the radiation polarization, which do not play a significant role in the effects under study. This amounts to assuming that all the vectors $\boldsymbol{k}$ appearing in (1) have almost the same directions and that they are all associated with the same polarization $\boldsymbol{\varepsilon}$.

constant. The annihilation operator $E^{(+)}(\boldsymbol{r},t)$ in (3) acting on the one-photon state $|\psi\rangle$ yields the vacuum, and we can rewrite (3) as:

$$W_I(\boldsymbol{r},t) = s\langle\psi|E^{(-)}(\boldsymbol{r},t)|0\rangle\langle 0|E^{(+)}(\boldsymbol{r},t)|\psi\rangle = s|\langle 0|E^{(+)}(\boldsymbol{r},t)|\psi\rangle|^2 \tag{6}$$

that is:

$$W_I(\boldsymbol{r},t) = s\left|\int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}}\,\sqrt{\frac{\hbar\omega_k}{2\epsilon_0}}\,c(\boldsymbol{k})e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)}\right|^2 \tag{7}$$

For a massive particle of mass m, the probability to find it at point $\boldsymbol{r}$ and at time t is given by the squared modulus $|\Psi(\boldsymbol{r},t)|^2$ of its wave function $\Psi(\boldsymbol{r},t)$. This wave function is the Fourier transform of the probability amplitude $g(\boldsymbol{k})$ that a measurement of the particle's momentum gives the value $\hbar\boldsymbol{k}$. For a free particle, this probability amplitude $g(\boldsymbol{k})$ has a time variation in $e^{-i\omega_k t}$. Equality (7) is thus reminiscent of this probability for a massive free particle; however, in view of the $\sqrt{\omega_k}$ factor (proportional to $\sqrt{k}$) in front of $c(\boldsymbol{k})$ in the integral of (7), $W_I(\boldsymbol{r},t)$ is not proportional (at a given instant t) to the modulus squared of the spatial Fourier transform of $c(\boldsymbol{k},t) = c(\boldsymbol{k})e^{-i\omega_k t}$. This means that, limiting ourselves to one-photon states, we can indeed consider the function $c(\boldsymbol{k})$ appearing in (1) as a wave function in momentum space, since $|c(\boldsymbol{k})|^2$ is a probability density for the photon momentum. However, the probability to detect a photon at point $\boldsymbol{r}$ and at time t with a photodetector is not simply the modulus squared of the Fourier transform of that "wave function in momentum space" $c(\boldsymbol{k})e^{-i\omega_k t}$. This confirms that it is not possible, for a photon, to introduce a spatial wave function that is exactly equivalent to that of a massive particle.

1-b. Detection probability amplitude

The right-hand side of (6) is the squared modulus of the function:

$$\mathcal{E}(\boldsymbol{r},t) = \langle 0|E^{(+)}(\boldsymbol{r},t)|\psi\rangle \tag{8}$$

which plays an important role in all the computations to follow; it has the dimensions of an electric field. For the wave packet written in (1), its expression is:

$$\mathcal{E}(\boldsymbol{r},t) = i\int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}}\,\sqrt{\frac{\hbar\omega_k}{2\epsilon_0}}\,c(\boldsymbol{k})e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)} \tag{9}$$

It should not be confused with the average value in state $|\psi\rangle$ of the operator $E^{(+)}(\boldsymbol{r},t)$ written in (4), since that average value is zero, as we mentioned above. Nor is it, as already pointed out, a wave function for the photon in position space; it is a probability amplitude for the detection (and not the presence) of the photon at point $\boldsymbol{r}$ and time t. When we mention, in this complement, the space time wave packet associated with the photon in state (1), we will always be referring to the amplitude (8).

Note, however, that in the particular case where the function $c(\boldsymbol{k})$ in momentum space is well centered around a value k_m with a dispersion Δk very small compared to k_m, one can neglect in (7) the k variation of $\sqrt{k}$ and replace $\sqrt{k}$ by $\sqrt{k_m}$; the integral in (7) will therefore involve the spatial Fourier transform of $c(\boldsymbol{k})e^{-i\omega_k t}$. This approximation will often be used in what follows.

Comments:

(i) In this section, we have ignored the radiation polarization, assuming that all the plane waves in relation (1) are associated with the same polarization vector $\boldsymbol{\varepsilon}$. When this is not the case, the detection amplitude becomes a three-component vector function. Its component along any given axis yields the detection probability amplitude by a photodetector preceded by a polarization analyser letting through only the light polarized linearly along that axis. This vector detection amplitude is similar to the wave function of a spin 1 particle, which also has three components.

(ii) In this complement, we shall study only wave packets containing a well defined photon number, 1 or 2, which allows directly generalizing the computations of Chapter XX. Wave packets can, however, be built many different ways, without exactly fixing the photon number. It often happens, for example, that one wishes to reproduce a classical field for which each field mode $\mathbf{k}$ has a given amplitude $\alpha(\mathbf{k})$; it is then natural to use a state where each quantum mode is in a coherent state with eigenvalue $\alpha(\mathbf{k})$. In that case, only the average photon number is fixed, not its exact value.

1-c. Temporal variation of the signal

When the photodetector is placed at $\mathbf{r} = 0$, it delivers a signal that is given, according to (7), by:

$$W_I(\mathbf{r} = \mathbf{0}, t) \propto \left| \int \mathrm{d}^3k \, \sqrt{k} \, c(\boldsymbol{k}) e^{-ickt} \right|^2 \tag{10}$$

This signal is proportional to the squared modulus of the Fourier transform of $\sqrt{k}\, c(\boldsymbol{k})$. Let us assume that $c(k)$ is a real positive function of k, and that $|c(k)|^2$ is a function of k centered at $k = k_m$, with a width Δk. If this width Δk of the wave packet is very small compared with the average wave number k_m, one can replace $\sqrt{k}$ by $\sqrt{k_m}$, and the signal becomes proportional to the modulus squared of the Fourier transform of $c(\boldsymbol{k})$. At time $t = 0$, and since we assumed all the $c(k)$ to be positive, all the waves forming the wave packets are in phase and the signal observed on the photodetector is maximum; it is zero for $t = \pm\infty$, and takes on significant values only during a time interval $\Delta t \simeq 1/c\Delta k$ around $t = 0$. This signal describes, in a way, the passage of the wave packet at the detector's position.

To study the detection probability at a point $\mathbf{r} \neq \mathbf{0}$, we just have to replace e^{-ickt} by $e^{i(\mathbf{k}\cdot\mathbf{r}-kct)}$ in the integral of relation (10). As an example, imagine we have a one-dimensional wave packet, all the wave vectors $\boldsymbol{k}$ being parallel to the Oz axis ($k_x = k_y = 0$); the exponential reduces to $e^{-ik(z-ct)}$. The phenomena observed at a point in space of coordinate $z \neq 0$ are thus deduced from those observed at $z = 0$ by a simple time shift equal to z/c: the wave packet moves along the Oz direction with velocity c and without any deformation.

2. One- or two-photon interference signals

We now discuss in terms of wave packets what happens in light interference experiments involving one or two photons.

2-a. How should one compute photon interference?

In non-relativistic quantum mechanics, a particle with a non-zero mass m is described by a wave function $\psi(\boldsymbol{r},t)$ whose squared modulus $|\psi(\boldsymbol{r},t)|^2$ yields the probability density of finding the particle at point $\boldsymbol{r}$ and time t. In a Young's type interference experiment, the wave function, after going through the two slits pierced into a screen, is a linear superposition of two wave functions $\psi_1(\boldsymbol{r},t)$ and $\psi_2(\boldsymbol{r},t)$ originating from the two slits. These two waves overlap in a region of space where the probability density of finding the particle at point $\boldsymbol{r}$ and time t, which is equal to $|\psi_1(\boldsymbol{r},t)+\psi_2(\boldsymbol{r},t)|^2$, contains a term $2\mathrm{Re}\{\psi_1(\boldsymbol{r},t)\psi_2^*(\boldsymbol{r},t)\}$ oscillating in space and time; this results in interference fringes.

However, we recalled in § 1 why we cannot, in general, introduce a spatial wave function for a photon that would be strictly analogous to $\psi(\boldsymbol{r},t)$, and whose squared modulus would yield the probability density for the presence of the photon at a given point. This led us to define an amplitude $\mathcal{E}(\boldsymbol{r},t)$ in (8), whose squared modulus yields the probability density for photodetecting the photon at point $\boldsymbol{r}$ and time t. We are going to show in § 2-b that such amplitudes can actually be used to interpret interference fringes; as an example, we shall study the fringes appearing in the single photodetection signal $w_I(\boldsymbol{r},t)$ observed on a one-photon wave packet after it goes through a screen pierced with two slits. As already underlined, it is important not to confuse $\mathcal{E}(\boldsymbol{r},t)$ with the average value of the electric field in the quantum state under study – which in any case is zero in a one-photon state. In classical electromagnetism, the electric (or magnetic) fields directly interfere; in quantum electromagnetism, one must reason in terms of probability amplitudes.

For field states containing at least two photons, the double photodetection signal $w_{II}(\boldsymbol{r}_B,t_B,\boldsymbol{r}_A,t_A)$ is different from zero. To interpret it in the simplest possible case, we assume (in § 2-c) that the radiation is described by a tensor product of two one-photon wave packets[3]. We will show that interference fringes observable on w_{II} can also be interpreted in terms of products of detection probability amplitudes; these fringes result from interference between transition amplitudes associated with two different paths leading the field from its initial state (where it contains two photons) to the vacuum. Here again, one should reason in terms of interference not directly between average values of electric or magnetic fields, but between paths.

2-b. Interference signal for a one-photon wave packet in two modes

We start with the simplest photon interference experiment, the well-know Young's double slit experiment, but in a case when only one photon at a time passes through the screen pierced with the two slits. The state vector of this single photon is then the sum of two components associated with the passage through one or the other of the two slits. When the photon reaches the interference region, these two components are associated with two different radiation modes.

[3] A simple example of a two-photon entangled state, which is not a product of two one-photon states, will be studied in § 5.

α. One-photon wave packets, after passage through a two-slit screen

We now focus on the radiation state after it passed the two-slit screen. This state is described by a one-photon wave packet, which, in the interaction representation, is of the form:

$$|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle \qquad \text{with:} \quad |c_1|^2 + |c_2|^2 = 1 \tag{11}$$

In this expression, the ket $|\psi_1\rangle$ describes the wave packet emerging from the first slit; as in (1), it can therefore be written with a function $c_1(\mathbf{k})$ that is peaked around the value $\boldsymbol{k}_1$. The ket $|\psi_2\rangle$ describes the wave packet emerging from the second slit, and its function $c_2(\mathbf{k})$ is peaked around the value $\boldsymbol{k}_2$. Since before going through the slits the two wave packets came from the same source, they must be centered around a common frequency $\omega_m = ck_1 = ck_2$; consequently $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ have the same modulus, but their directions can be different. We shall finally assume that the wave packets emerging from the two slits arrive at the same time in the interference region (meaning the optical paths along the two trajectories are equal) and that each wave packet is sufficiently long for the frequency ω_m to be well defined.

As in (8), for each wave packet $|\psi_i\rangle$ we introduce a detection amplitude $\mathcal{E}_i(\boldsymbol{r}, t)$:

$$\langle 0|E^{(+)}(\boldsymbol{r}, t)|\psi_i\rangle = \mathcal{E}_i(\boldsymbol{r}, t) \qquad \text{where:} \quad i = 1, 2 \tag{12}$$

In the interference region, we assume the two modes[4] to be close to plane waves with wave vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. We then set:

$$\mathcal{E}_i(\boldsymbol{r}, t) = \mathcal{F}_i(\boldsymbol{r}, t)\ e^{i(\boldsymbol{k}_i \cdot \boldsymbol{r} - \omega_m t)} \tag{13}$$

where the function $\mathcal{F}_i(\boldsymbol{r}, t)$ has a much slower space and time variation than the exponential $e^{i(\boldsymbol{k}_i \cdot \boldsymbol{r} - \omega_m t)}$.

β. Calculation of the single photodetection signal

We assume that the field is contained in a box of volume L^3; we use a complete orthonormal set of field modes, with wave vectors $\{\boldsymbol{k}_i\}$, which includes both $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. Relation (B-3) of Chapter XIX indicates that the positive frequency component of the electric field can be written [5] as:

$$E^{(+)}(\boldsymbol{r}, t) = i \sum_i \sqrt{\frac{\hbar\omega_i}{2\varepsilon_0 L^3}}\ e^{i(\boldsymbol{k}_i \cdot \boldsymbol{r} - \omega_i t)}\ a_i \tag{14}$$

with $\omega_i = c \times k_i$. When this operator acts on the ket (11), all the i terms lead to a zero result, except for the $i = 1$ and $i = 2$ terms. For these two terms, we have $a_i|\psi_{1,2}\rangle = |0\rangle$, so that (11) and (12) lead to:

$$E^{(+)}(\boldsymbol{r}, t)|\psi\rangle = \big[c_1\ \mathcal{E}_1(\boldsymbol{r}, t) + c_2\ \mathcal{E}_2(\boldsymbol{r}, t)\big]|0\rangle \tag{15}$$

[4] Another possibility would be to use Gaussian wave packets with the same focal point, having in the vicinity of that point plane wave structures with wave vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$, and lateral extensions very large compared to the wavelengths $2\pi/k_1$ and $2\pi/k_2$.

[5] For the sake of simplicity, we ignore polarization variables of the field.

The probability for detecting the photon at point $\boldsymbol{r}$ and time t is proportional to the square of norm of this ket, written as:

$$\langle\psi|E^{(-)}(\boldsymbol{r},t)E^{(+)}(\boldsymbol{r},t)|\psi\rangle = \left|c_1\mathcal{E}_1(\boldsymbol{r},t) + c_2\mathcal{E}_2(\boldsymbol{r},t)\right|^2 \tag{16}$$

The equality includes square and cross terms. The square terms can be written, taking (13) into account:

$$|c_i\mathcal{E}_i(\boldsymbol{r},t)|^2 = |c_i\mathcal{F}_i(\boldsymbol{r},t)|^2 \tag{17}$$

and they vary slowly as a function of $\boldsymbol{r}$ and t. The crossed terms are expressed as:

$$c_1c_2^*\ \mathcal{E}_1(\boldsymbol{r},t)\mathcal{E}_2^*(\boldsymbol{r},t) + \text{c.c.} = c_1c_2^*\ \mathcal{F}_1(\boldsymbol{r},t)\mathcal{F}_2^*(\boldsymbol{r},t)\exp\{i[(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot\boldsymbol{r})]\} + \text{c.c.} \tag{18}$$

and exhibit spatial modulations characteristic of interference phenomena (c.c. stands for complex conjugate).

γ. Discussion

Relation (16) shows that the photodetection signal is the squared modulus of the sum of two amplitudes, $c_1\mathcal{E}_1(\boldsymbol{r},t)$ and $c_2\mathcal{E}_2(\boldsymbol{r},t)$, which interfere. Amplitude $c_1\mathcal{E}_1(\boldsymbol{r},t)$ is the amplitude for detecting at point $\boldsymbol{r}$ and time t the photon in mode $|\psi_1\rangle$; it is equal to the amplitude c_1 of finding the field in state $|\psi_1\rangle$, multiplied by the amplitude $\mathcal{E}_1(\boldsymbol{r},t)$ for detecting the photon at point $\boldsymbol{r}$ and time t when the field is in state $|\psi_1\rangle$. The amplitude $c_2\mathcal{E}_2(\boldsymbol{r},t)$ is interpreted in a similar way. During the detection process, the field goes from state $|\psi\rangle$ written in (11) to the vacuum state $|0\rangle$ following two possible paths: the photon is absorbed either while in mode $|\psi_1\rangle$, or while in mode $|\psi_2\rangle$. As nothing allows deciding which path the system followed, the two corresponding amplitudes interfere. This confirms what we stated above: in the quantum theory of radiation, the interference fringes observed on a photodetector signal are associated with the interference, not between two classical electromagnetic waves, but rather between two transition amplitudes corresponding to different paths (leading the system from the same initial state to the same final state).

2-c. Interference signals for a product of two one-photon wave packets

Let us generalize this type of interpretation, in terms of transition amplitudes, to interference experiments involving two photons and where one measures correlations between signals coming from two photodetectors.

α. State vector for the two photons

We now assume the field contains two photons, and can be described as the product of two wave packets such as the one written in (1):

$$|\psi_{12}\rangle = \int \mathrm{d}^3k\ c_1(\boldsymbol{k}) \int \mathrm{d}^3k'\ c_2(\boldsymbol{k}')\, a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'}^\dagger|0\rangle \tag{19}$$

Is it possible to observe spatial and temporal modulations on the signals w_I and w_{II} coming from one or two detectors placed in that field? We are going to show that the

answer to that question is "no" if one considers only single photodetections, but "yes" if one takes into account the correlations between double photon detections.

We assume the two wave packets in (19) to be well separated: there exists no overlap between the domain D_1 where the function $c_1(\boldsymbol{k})$ is different from zero and the domain D_2 where the function $c_2(\boldsymbol{k})$ is non-zero. Let us introduce the matrix element that generalizes relation (8) to the two-photon case:

$$\langle 0|E^{(+)}(\boldsymbol{r},t)\ E^{(+)}(\boldsymbol{r}',t')|\psi_{12}\rangle \tag{20}$$

We now insert in that expression the plane mode expansion (14) for both electric field operators. To yield a non-zero result, the annihilation operators appearing in these fields must act on a mode that, according to (19), contains one photon. This means that either the mode selected in $E^{(+)}(\boldsymbol{r},t)$ belongs to the D_1 domain and the one selected in $E^{(+)}(\boldsymbol{r}',t')$ to the D_2 domain, or the inverse. In the first case, the scalar product of the vacuum bra and the modes that came into play yields the detection amplitude $\mathcal{E}_2(\boldsymbol{r}',t')$ associated with the second wave packet, multiplied by the detection amplitude $\mathcal{E}_1(\boldsymbol{r},t)$ associated with the first one. In the second case, the wave packets are inverted. The final result is:

$$\langle 0|E^{(+)}(\boldsymbol{r},t)\ E^{(+)}(\boldsymbol{r}',t')|\psi_{12}\rangle = \mathcal{E}_1(\boldsymbol{r},t)\ \mathcal{E}_2(\boldsymbol{r}',t') + \mathcal{E}_2(\boldsymbol{r},t)\ \mathcal{E}_1(\boldsymbol{r}',t') \tag{21}$$

where the functions $\mathcal{E}_1$ and $\mathcal{E}_2$ are the detection amplitudes associated, as defined in (8), with the two individual wave packets included in $|\psi_{12}\rangle$.

β. Single photodetection signal $w_I(\boldsymbol{r},t)$

To get the single photodetection signal, we first compute the result of the action on state (19) of the field positive frequency component $E^{(+)}(\boldsymbol{r}',t')$. As we argued above, to yield a non-zero result, this operator must destroy a photon, either in a mode for which $c_1(\mathbf{k})$ is non-zero, or in a mode for which $c_2(\mathbf{k})$ is non-zero. In the first case, the summation over all the modes involved reconstructs the function $\mathcal{E}_1(\boldsymbol{r}',t')$ multiplied by the vacuum ket associated with these modes; the modes of the other wave packet remain unchanged. In the second case, the two wave packets exchange roles and it is now the function $\mathcal{E}_2(\boldsymbol{r}',t')$ that is reconstructed. This leads to:

$$E^{(+)}(\boldsymbol{r}',t')|\psi_{12}\rangle = \left[\mathcal{E}_1(\boldsymbol{r}',t')\int \mathrm{d}^3k'\ c_2(\boldsymbol{k}')\,a^\dagger_{\boldsymbol{k}'}|0\rangle + \mathcal{E}_2(\boldsymbol{r}',t')\int \mathrm{d}^3k\ c_1(\boldsymbol{k})\,a^\dagger_{\boldsymbol{k}}|0\rangle\right] \tag{22}$$

The probability per unit time to detect a photon at point $\boldsymbol{r}'$ and time t' is the square of this ket's norm. The terms in $|\mathcal{E}_1(\boldsymbol{r}',t')|^2$ and $|\mathcal{E}_2(\boldsymbol{r}',t')|^2$ contain the square of the two wave packets' norms, each equal to one; these terms do not oscillate, neither in space, nor in time. The cross terms are the only ones that could yield spatial and temporal modulations; they contain, however, the scalar product of the two wave packets, which is zero since we assumed the wave packets were orthogonal (there is no overlap between the two D_1 and D_2 domains). This means that, when the field is described by state (19), no interference fringes are observable in the signal of a single photodetector.

The interpretation of this result is similar to the one we gave before. The system can follow two paths: either an absorption of a photon from the first wave packet, or an

absorption of a photon from the second. However, as opposed to what happens when the system started from the initial state (11), the final state of the field is not the same for these two paths: if a photon from the first wave packet has been absorbed, the final state includes a photon from the second wave packet. Consequently, the two final states associated with the two paths are orthogonal, and observing the field's final state one could (in principle) determine which path the system has followed; this is why the two amplitudes cannot interfere.

Comment:

One could consider other states for the two modes, each containing several photons, as for example states $|\alpha_1\rangle \otimes |\alpha_2\rangle$ where each mode is in a coherent state, characterized by a classical normal variable, α_1 for mode $\{\boldsymbol{k}_1\}$, α_2 for mode $\{\boldsymbol{k}_2\}$. We then know that state $|\alpha_1\rangle$ is an eigenstate of operator $E^{(+)}(\boldsymbol{r},t)$ with an eigenvalue value $\mathcal{E}_{\rm cl}^{(+)}(\alpha_1,\boldsymbol{r},t)$ equal to the positive frequency component of the classical field in mode $\{\boldsymbol{k}_1\}$, corresponding to the classical normal variable α_1 (Chapter XVIII, § B-2). A similar result is valid for state $|\alpha_2\rangle$:

$$E^{(+)}(\boldsymbol{r},t)|\alpha_i\rangle = \mathcal{E}_{\rm cl}^{(+)}(\alpha_i,\boldsymbol{r},t)|\alpha_i\rangle \qquad i=1,2 \tag{23}$$

This leads to:

$$E^{(+)}(\boldsymbol{r},t)|\alpha_1,\alpha_2\rangle = \left[\mathcal{E}_{\rm cl}^{(+)}(\alpha_1,\boldsymbol{r},t)+\mathcal{E}_{\rm cl}^{(+)}(\alpha_2,\boldsymbol{r},t)\right]|\alpha_1,\alpha_2\rangle \tag{24}$$

The probability of detecting a photon at point $\boldsymbol{r}$ and time t is equal to the square of the norm of ket (24). It is proportional to:

$$\left|\mathcal{E}_{\rm cl}^{(+)}(\alpha_1,\boldsymbol{r},t)+\mathcal{E}_{\rm cl}^{(+)}(\alpha_2,\boldsymbol{r},t)\right|^2$$

As this is the squared modulus of the sum of two classical fields, it is the usual interference signal of classical fields. As opposed to what we found before for the radiation state (19), when the two modes are in coherent states, the one-photon detection signal exhibit interference. This is an illustration of the quasi-classical character of coherent states.

γ. *Double photodetection signal* $w_{II}(\boldsymbol{r}_B,t_B,\boldsymbol{r}_A,t_A)$

Assuming, as above, the field initial state is given by (19), we now focus on the probability $w_{II}(\boldsymbol{r}_B,t_B,\boldsymbol{r}_A,t_A)$ (per double unit time) that a detector, placed at $\boldsymbol{r}_A$, detects a photon at time t_A and that another detector, placed at $\boldsymbol{r}_B$, detects a photon at time t_B. This probability is proportional to the correlation function (Complement B_{XX}, § 2-d):

$$\langle\psi_{12}|E^{(-)}(\boldsymbol{r}_A,t_A)E^{(-)}(\boldsymbol{r}_B,t_B)E^{(+)}(\boldsymbol{r}_B,t_B)E^{(+)}(\boldsymbol{r}_A,t_A)|\psi_{12}\rangle \tag{25}$$

Since $|\psi_{12}\rangle$ contains only two photons, we can insert in the middle of this expression the projector onto the vacuum state, which leads to the squared modulus of expression (21). We obtain:

$$w_{II}(\boldsymbol{r}_B,t_B;\boldsymbol{r}_A,t_A) \propto \left|\mathcal{E}_2(\boldsymbol{r}_B,t_B)\mathcal{E}_1(\boldsymbol{r}_A,t_A)+\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_2(\boldsymbol{r}_A,t_A)\right|^2 \tag{26}$$

In addition to the square terms $|\mathcal{E}_2(\boldsymbol{r}_B,t_B)\mathcal{E}_1(\boldsymbol{r}_A,t_A)|^2$ and $|\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_2(\boldsymbol{r}_A,t_A)|^2$, which have a slow variation with $\boldsymbol{r}_A$, $\boldsymbol{r}_B$, t_A, t_B, we get cross terms:

$$\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_1^*(\boldsymbol{r}_A,t_A)\mathcal{E}_2(\boldsymbol{r}_A,t_A)\mathcal{E}_2^*(\boldsymbol{r}_B,t_B)+\text{c.c.}$$

which do have spatial and temporal modulations. If, for example, the first wave packet is centered around the values $\boldsymbol{k}_1$ and ω_1, and the second one around the values $\boldsymbol{k}_2$ and ω_2, these modulations are of the form:

$$\exp\Big\{i\big[(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot(\boldsymbol{r}_A-\boldsymbol{r}_B)-(\omega_1-\omega_2)(t_A-t_B)\big]\Big\}+\text{c.c.} \tag{27}$$

This result is not in contradiction with the fact that the probability of detecting a photon at $\mathbf{r}_A$, t_A, or at $\mathbf{r}_B$, t_B varies slowly with these variables: once a first photon has been detected at $\boldsymbol{r}_A,t_A$, the probability to detect another one at $\boldsymbol{r}_B,t_B$ varies sinusoidally with $\boldsymbol{r}_A-\boldsymbol{r}_B$ and t_A-t_B.

(i) ***Discussion***

The amplitude whose squared modulus appears on the right-hand side of (26) is the sum of two amplitudes associated with two possible paths leading the system from the initial state $|\psi_{12}\rangle$ (containing two photons) to the same final state $|0\rangle$ (where all the modes are empty). Along the first path, with amplitude $\mathcal{E}_2(\boldsymbol{r}_B,t_B)\mathcal{E}_1(\boldsymbol{r}_A,t_A)$, the k_1 mode photon is absorbed at $\boldsymbol{r}_A,t_A$ and the k_2 mode photon is absorbed at $\boldsymbol{r}_B,t_B$. Along the second path, with amplitude $\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_2(\boldsymbol{r}_A,t_A)$, the opposite happens: the k_2 mode photon is absorbed at $\boldsymbol{r}_A,t_A$ and the k_1 mode photon is absorbed at $\boldsymbol{r}_B,t_B$. As explained before (§ 2-b), interference occurs between the different transition amplitudes associated with two possible paths leading the system from the same initial state to the same final state, as long as there is no way one can determine which path is actually followed.

(ii) ***Another interpretation***

The photodetection signal (25) can also be written in the form:

$$\langle\psi_A|E^{(+)}(\boldsymbol{r}_B,t_B)E^{(+)}(\boldsymbol{r}_B,t_B)|\psi_A\rangle \tag{28}$$

with:

$$\begin{aligned}|\psi_A\rangle &= E^{(+)}(\boldsymbol{r}_A,t_A)|\psi_{12}\rangle\\ &= \Big[\mathcal{E}_1(\boldsymbol{r}_A,t_A)\int \mathrm{d}^3k'\ c_2(\boldsymbol{k}')\,a^\dagger_{\boldsymbol{k}'}|0\rangle+\mathcal{E}_2(\boldsymbol{r}_A,t_A)\int \mathrm{d}^3k\ c_1(\boldsymbol{k})\,a^\dagger_{\boldsymbol{k}}|0\rangle\Big]\end{aligned} \tag{29}$$

where, in the second line, we used relation (22). Signal (28) can be interpreted as the probability of detecting a photon when the field is that state $|\psi_A\rangle$ where the photon has a probability amplitude $\mathcal{E}_1(\boldsymbol{r}_A,t_A)$ to be in the wave packet with amplitude $c_2(\boldsymbol{k}')$, and a probability amplitude $\mathcal{E}_2(\boldsymbol{r}_A,t_A)$ to be in the other wave packet with amplitude $c_1(\boldsymbol{k})$. This situation is quite similar to that encountered in § 2-b-α, where we showed that the photodetection probability of a photon in state (11) exhibits modulations.

In other words, we started from a state $|\psi_{12}\rangle$ with no coherence. It is the detection of a first photon that introduces the state (29) where a second photon is now in a coherent superposition, the coherence arising from the fact that the detected photon can come either from the first wave packet, or from the second. The coefficients of the

superposition (29) depend on the point $\boldsymbol{r}_A$ and the instant t_A where the detection of the first photon occurred. In this description of the phenomena, it is the first detection that introduces quantum correlations between the two modes, and the dependence of these correlations on $\boldsymbol{r}_A$ and t_A explains why the probability of the second detection oscillates as a function of $\boldsymbol{r}_B - \boldsymbol{r}_A$ and $t_B - t_A$.

3. Absorption amplitude of a photon by an atom

We now replace the broadband photodetector by an atom with two discrete levels, a ground level a and an excited level b. This atom is placed at $\boldsymbol{r} = \boldsymbol{0}$, and interacts with the same wave packet $|\psi\rangle$ as that written in (1). We propose to compute the probability amplitude for the atom, initially in state a, to absorb the incident photon and be found in state b at time t.

3-a. Computation of the amplitude

The initial and final states of the process under study are:

$$|\psi_{\text{in}}\rangle = |a;\psi\rangle \qquad |\psi_{\text{fin}}\rangle = |b;0\rangle \tag{30}$$

since the absorption of the photon transfers the radiation from state $|\psi\rangle$ to the vacuum $|0\rangle$. According to relation (B-4) in Chapter XX, the amplitude we are looking for is, to first order in H_I:

$$\langle\psi_{\text{fin}}|\bar{U}(t,-\infty)|\psi_{\text{in}}\rangle = \frac{1}{i\hbar}\int_{-\infty}^{t} \mathrm{d}t' \ \langle\psi_{\text{fin}}|\bar{H}_I(t')|\psi_{\text{in}}\rangle \tag{31}$$

where the bar above the operators indicates they are expressed in the interaction picture, with respect to the non-perturbed Hamiltonian.

The interaction Hamiltonian H_I is given by[6]:

$$H_I = -DE^{(+)}(\boldsymbol{r} = \boldsymbol{0}) \tag{32}$$

The matrix element of $\bar{H}_I(t')$ appearing in (31) equals:

$$\langle\psi_{\text{fin}}|\bar{H}_I(t')|\psi_{\text{in}}\rangle = -\mathcal{D}e^{i\omega_0 t'}\langle 0|E^{(+)}(\boldsymbol{r} = \boldsymbol{0}, t')|\psi\rangle \tag{33}$$

In this equality, $\mathcal{D} = \langle b|D|a\rangle$, $\omega_0 = (E_b - E_a)/\hbar$ is the frequency of the atomic transition, and $E^{(+)}(\boldsymbol{r} = \boldsymbol{0}, t')$ the electric field positive frequency component in the interaction representation. Using notation (8) for the matrix element on the right-hand side of (33), we get:

$$\langle\psi_{\text{fin}}|\bar{H}_I(t')|\psi_{\text{in}}\rangle = -\mathcal{D}\ e^{i\omega_0 t'}\ \mathcal{E}(\boldsymbol{r} = \boldsymbol{0}, t') \tag{34}$$

which allows rewriting the absorption amplitude (31) as:

$$\langle\psi_{\text{fin}}|\bar{U}(t,-\infty)|\psi_{\text{in}}\rangle = -\frac{\mathcal{D}}{i\hbar}\int_{-\infty}^{t} \mathrm{d}t' \ e^{i\omega_0 t'}\ \mathcal{E}(\boldsymbol{r} = \boldsymbol{0}, t') \tag{35}$$

[6] As we have ignored the radiation polarization degrees of freedom, we also ignore here the vector character of the atomic dipole $\boldsymbol{D}$. Operator D appearing in (32) is actually the projection of $\boldsymbol{D}$ onto the radiation polarization vector.

The quantity $-\mathcal{D}\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t')$ that appears in this expression can be considered to be the interaction Hamiltonian between the atomic dipole $\mathcal{D}$ and a classical field $\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t')$. The field $\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t')$ thus appears as the classical field that would yield the same transition amplitude between the two atomic states a and b as a quantum field, when the radiation is in a one-photon wave packet described by $c(\boldsymbol{k})$.

3-b. Properties of that amplitude

Let us return to the wave packet in infinite space (1) and use relation (8) to get the probability amplitude for the detection of the photon at point $\mathbf{r}=\mathbf{0}$. Describing the field operator (4) with an integral over d^3k' instead of d^3k, and using it in (1), the commutation of operators $a(\boldsymbol{k}')$ and $a^\dagger(\boldsymbol{k})$ leads to a $\delta(\boldsymbol{k}-\boldsymbol{k}')$ function. This leads to:

$$\langle 0|E^{(+)}(\boldsymbol{r}=\boldsymbol{0},t')|\psi\rangle = \mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t') = \frac{i}{(2\pi)^{3/2}}\int \mathrm{d}^3k\,\sqrt{\frac{\hbar\omega_k}{2\varepsilon_0}}\;c(k)e^{-i\omega t'} \tag{36}$$

where $\omega_k = ck$. As $c(k)$ is centered around an average wave vector k_m, assumed very large compared to the width Δk of $c(k)$, amplitude (36) can be written as:

$$\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t') = e^{-i\omega_m t'}\;\mathcal{F}(t') \tag{37}$$

which is the product of a carrier wave of frequency $\omega_m = ck_m$ by an envelope $\mathcal{F}(t')$. This latter function has a t' variation that is slower than the preceding exponential; it can, for example, follow a bell-shaped curve, centered at $t=0$ and with width $\Delta t \simeq 1/c\Delta k$. Inserting (37) into (35), we get:

$$\langle\psi_{\mathrm{fin}}|\bar{U}(t,-\infty)|\psi_{\mathrm{in}}\rangle = -i\int_{-\infty}^{t}\mathrm{d}t'\;e^{i(\omega_0-\omega_m)t'}\;\Omega_1(t') \tag{38}$$

where $\Omega_1(t')$ is the instantaneous Rabi frequency defined as:

$$\hbar\Omega_1(t') = -\mathcal{D}\mathcal{F}(t') \tag{39}$$

Equation (38) allows understanding the behavior of the absorption amplitude of the photon when t increases from $-\infty$ to $+\infty$. As long as $t \ll -\Delta t$, both functions $\mathcal{E}(t')$ and $\Omega_1(t')$ are zero; the incident wave packet has not yet reached the atom's vicinity and no photon absorption can occur. As t increases from $-\Delta t/2$ to $+\Delta t/2$, the wave packet crosses the atom, and the integral in (38) becomes larger. When $t \gg +\Delta t$, the wave packet has left the atom; the absorption amplitude remains constant and equal to:

$$\langle\psi_{\mathrm{fin}}|\bar{U}(+\infty,-\infty)|\psi_{\mathrm{in}}\rangle = -i\int_{-\infty}^{+\infty}\mathrm{d}t'\;e^{i(\omega_0-\omega_m)t'}\;\Omega_1(t') \tag{40}$$

This expression yields the probability amplitude for a photon to have been absorbed once the wave packet has crossed the atom. This confirms the results of Chapter XX, but in the present approach we did not have to artificially introduce any initial or final times for the process.

Let us evaluate an order of magnitude for the amplitude (40). Assume first that $\omega_m = \omega_0$ (resonant wave packet). The integral in (40) is then of the order of $\Omega_1^{\max}\Delta t$,

where $\Omega_1^{\max}$ is the maximum value reached by the Rabi frequency when the atom is at the center of the wave packet, and the envelope $\mathcal{F}(t')$ takes on its largest value. When $\omega_m \neq \omega_0$ (off-resonance wave packet), the absorption amplitude is weaker. According to (40), this amplitude is actually the Fourier transform of $\Omega_1(t')$ at frequency $\omega_0 - \omega_m$. This result simply expresses energy conservation: for the incident photon to be absorbed, its frequency must be equal to the atomic transition frequency. However, as the field envelope varies over time intervals of the order of Δt, the photon average frequency does not have to be strictly equal to the atomic frequency; the two frequencies must be equal to within $\Delta\omega \simeq 1/\Delta t$.

4. Scattering of a wave packet

We now study a process involving two atoms: a wave packet impinges on an atom A placed at $\boldsymbol{r}_A$ on the z axis; after interacting with it, the wave packet is scattered in all directions, and then interacts with a second atom B placed at $\boldsymbol{r}_B$. The incident wave packet, propagating along the Oz direction, is described by the function $c(\boldsymbol{k})e^{-i\omega t}$. As before, we have two main goals. The first one is, while assimilating atom B with a device for measuring the photon scattered by A, to confirm the interpretation of $\mathcal{E}(\mathbf{r}, t)$ as a detection amplitude of a photon at point $\mathbf{r}_B$. The second goal is to study the time dependence of the scattering process itself.

We shall first study the spatial and temporal dependence of the scattered wave packet, in particular when the central frequency ω_m of the incident wave packet is close to the resonant frequency ω_0 of the scattering atom. We shall then compute the probability amplitude for the scattered wave packet to have excited at time t the atom B from its ground state a to its excited state b. As in § 1, we will associate with this amplitude a spatial wave packet describing the passage of the scattered wave packet by point $\boldsymbol{r}_B$.

4-a. Absorption amplitude by atom B of the photon scattered by atom A

We first consider a photon with a wave vector $\boldsymbol{k}_i$ parallel to the z axis, and a frequency $\omega_i = ck_i$. We are looking for the probability amplitude $\langle \boldsymbol{k}_f | S | \boldsymbol{k}_i \rangle$ for this photon to be scattered by atom A located at $\boldsymbol{r}_A$ from state $\boldsymbol{k}_i$ to state $\boldsymbol{k}_f$. This amplitude is given by relation (E-3) of Chapter XX, where we only take into account the resonant α processes (we assume the incident photon frequency to be fairly close to the atomic resonant frequency):

$$\begin{aligned}\langle \boldsymbol{k}_f | S | \boldsymbol{k}_i \rangle &= \langle \psi_{\text{fin}} | \bar{U}(\Delta t, 0) | \psi_{\text{in}} \rangle \\ &= -2\pi i \; A_{fi}^{\alpha}(E_{\text{in}}) \; \delta^{(\Delta t)}(E_{\text{fin}} - E_{\text{in}}) \equiv \mathcal{B}\, \delta^{(\Delta t)}(\omega_f - \omega_i)\end{aligned} \tag{41}$$

In this relation, $A_{fi}^{\alpha}(E_{\text{in}})$ is obtained from relation (E-4) in Chapter XX (here again we assume that the radiation is contained in a box of volume L^3):

$$A_{fi}^{\alpha}(E_{\text{in}}) = -\frac{\hbar\sqrt{\omega_i \omega_f}}{2\varepsilon_0 L^3} \frac{\langle a | \boldsymbol{\varepsilon}_f^* \cdot \boldsymbol{D} | c \rangle \langle c | \boldsymbol{\varepsilon}_i \cdot \boldsymbol{D} | a \rangle}{E_a + \hbar\omega_i - E_c} \; e^{i(\boldsymbol{k}_i - \boldsymbol{k}_f)\cdot \boldsymbol{r}_A} \tag{42}$$

where we have assumed that only one level c contributes, which explains why the sum over c has been suppressed; this is correct if the frequency of the radiation is close to the resonance frequency of one $a \leftrightarrow c$ transition, but far from all the other resonances. Note

that we have added to the right-hand side an exponential factor that comes from the spatial dependence of the electric field: in this complement, as in Chapter XX, we treat the atom's position $\boldsymbol{r}_A$ classically, but we no longer assume the atom to be placed at the coordinate origin. Expression (42) is a product of two matrix elements of the interaction Hamiltonian, one for the absorption of the $\boldsymbol{k}_i$ photon, the other for the emission of the $\boldsymbol{k}_f$ photon, divided by a common energy denominator. The function $\delta^{(\Delta t)}(\omega_f - \omega_i)$ simply expresses energy conservation, within $\hbar/\Delta t$, for the elastic scattering process, as was the case, for example, in §§ B-1-b and E-1-b of Chapter XX. We assume that the interaction time Δt is sufficiently long for this function to be assimilated to a real delta function $\delta(\omega_f - \omega_i)$.

The coefficient $\mathcal{B}$ introduced in the second equality (41) is proportional to expression (42); it contains the product of two matrix elements, which depends on the polar angles θ and φ of the vector $\boldsymbol{k}_f$ with respect to the direction of $\boldsymbol{k}_i$. We characterize this dependence by a function $f(k, \theta, \varphi)$, with:

$$k = |\boldsymbol{k}_i| = |\boldsymbol{k}_f| \tag{43}$$

As we assumed the frequency ω_i of the incident photon to be close to resonance, we can use the results of § C-2 in Chapter XX concerning resonant scattering. As in relation (E-11) of that chapter, we write the energy denominator in the form $\omega_i - \omega_0 + i\Gamma/2$, where Γ is the natural width of the excited state of the scattering atom A. Amplitude (41) then becomes:

$$\langle \boldsymbol{k}_f | S | \boldsymbol{k}_i \rangle = C_k \, \frac{f(k, \theta, \varphi)}{\omega_i - \omega_0 + i\Gamma/2} \, e^{i[(\boldsymbol{k}_i - \boldsymbol{k}_f)\cdot \boldsymbol{r}_A]} \, \delta(\omega_f - \omega_i) \tag{44}$$

where C_k is a coefficient proportional to k.

We now move to the next stage, the interaction of atom B with the $\boldsymbol{k}_f$ photon. As in (33), it is described by a matrix element (here again, we must add an exponential factor to account for the fact that atom B is not at the coordinate origin, but at point $\mathbf{r}_B$):

$$-\mathcal{D} e^{i\omega_0 t'} \langle 0 | E^{(+)}(\boldsymbol{r}_B, t') | a; \boldsymbol{k}_f \rangle \propto e^{i\boldsymbol{k}_f \cdot \boldsymbol{r}_B} e^{i(\omega_0 - \omega_f)t'} \tag{45}$$

We are now looking for the amplitude at time t of the complete process, scattering by A of the $\boldsymbol{k}_i$ photon, with an amplitude given by (44), and absorption by B, with an amplitude given by (45). Consequently, we multiply these two amplitudes and sum the product over all the possible $\boldsymbol{k}_f$ vectors for the scattered photon, and over the linear combination of states $\boldsymbol{k}_i$ forming the incident wave packet.

4-b. Wave packet scattered by atom A

To study the properties of the wave packet scattered by atom A, we successively carry out the two summations.

α. *Summation over all possible directions of the scattered photon*

Let us start with the summation over $\boldsymbol{k}_f$. Taking into account the function $\delta(\omega_f - \omega_i)$ appearing in (44), the summation over the modulus of $\boldsymbol{k}_f$ leads to:

$$k_f = k_i = k \tag{46}$$

Regrouping the $\boldsymbol{k}_f$ dependent terms in (44) and (45), we find that the summation over the directions of $\boldsymbol{k}_f$ introduces the angular integral:

$$\int \mathrm{d}\Omega \; f(k,\theta,\varphi) \; e^{i\boldsymbol{k}_f\cdot(\boldsymbol{r}_B-\boldsymbol{r}_A)} \tag{47}$$

The summation over the polar angles θ, φ of the exponentials describing the phase shift between $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$ of all the plane waves $\boldsymbol{k}_f$ yields a spherical wave centered at $\boldsymbol{r}_A$:

$$\int \mathrm{d}\Omega \; f(k,\theta,\varphi) \; e^{i\boldsymbol{k}_f\cdot(\boldsymbol{r}_B-\boldsymbol{r}_A)} \propto f(k,\theta_B,\varphi_B)\frac{e^{ikr}}{r} \qquad \text{with} \qquad r=|\boldsymbol{r}_B-\boldsymbol{r}_A| \tag{48}$$

where θ_B and φ_B are the polar angles of vector $\boldsymbol{r}_B - \boldsymbol{r}_A$ with respect to the direction $\boldsymbol{k}_i$ of the incident photon. The right-hand side of (48) is reminiscent of a classical result in collision theory – see for example relation (B-12) of Chapter VIII. This means that the sum of all the plane waves scattered by atom A located at $\boldsymbol{r}_A$ has the structure of an outgoing spherical wave with the same wave number k as the $\boldsymbol{k}_f$ waves it is composed of. The amplitude of this spherical wave varies as $1/r$, which ensures that the outgoing energy across a sphere of radius r and surface $4\pi r^2$ does not depend on r.

The fact that the polar angles θ_B and φ_B appearing on the right-hand side of (48) are those of vector $\boldsymbol{r}_B - \boldsymbol{r}_A$ can be understood by stationary phase arguments. The phase factor $e^{i\boldsymbol{k}_f\cdot(\boldsymbol{r}_B-\boldsymbol{r}_A)}$ associated with the scattered wave $\boldsymbol{k}_f$ is equal to $e^{ikr\cos\alpha_f}$, where α_f is the angle between $\boldsymbol{k}_f$ and $\boldsymbol{r}_B - \boldsymbol{r}_A$. Since $kr \gg 1$, this phase factor has a very rapid variation with α_f, except in the vicinity of points where $\cos\alpha_f$ is stationary with respect to α_f, i.e. when $\alpha_f = 0$ for the outgoing wave. The angular integral (47) therefore gets most of its contribution from values of the angles that are close to the polar angles θ_B and φ_B of vector $\boldsymbol{r}_B - \boldsymbol{r}_A$.

Taking (48) into account, we deduce from (44) and(45):

$$\sum_{\boldsymbol{k}_f}\langle b;0|\bar{H}_I(t')|a;\boldsymbol{k}_f\rangle\langle\boldsymbol{k}_f|S|\boldsymbol{k}_i\rangle \propto f(k,\theta_B,\varphi_B)\frac{e^{ikr}}{r}\,\frac{1}{\omega-\omega_0+i\Gamma/2}\,e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}_A}\,e^{i(\omega_0-\omega)t'} \tag{49}$$

where we have replaced ω_i and ω_f by ω.

β. Summation over the energies

The initial state of the scattering process is a superposition of states $|\boldsymbol{k}_i\rangle$ multiplied by $c(\boldsymbol{k}_i)$. We consider a one-dimensional wave packet, propagating along the Oz axis. With a proper choice of the coordinate origin, we can assume atom A is located at $\boldsymbol{r} = 0$, which amounts to replacing $\boldsymbol{r}_A$ by $\boldsymbol{0}$. As before, we assume $c(k)$ is real, so that at $t = 0$, the wave packet is centered at the position $\boldsymbol{r} = 0$ of the scattering atom A.

We now multiply (49) by $c(k)$, integrate over k, and over the time t' from $-\infty$ to t. The amplitude of the absorption by atom B, at time t, of the photon scattered by atom A is therefore proportional to:

$$\int_{-\infty}^{t}\mathrm{d}t' \; e^{i\omega_0 t'}\int \mathrm{d}\omega \; c(\omega)\omega \; f(\omega,\theta_B,\varphi_B)\frac{e^{i\omega r/c}}{r}\,\frac{1}{\omega-\omega_0+i\Gamma/2}\,e^{-i\omega t'} \tag{50}$$

where we have replaced the integral variable k by $\omega = ck$.

Let us compare (50) and (35). The integral over ω appearing in the integral over t' in (50) can be interpreted as the classical field scattered at time t' and at a distance r from point O along an axis with polar angles θ_B, φ_B:

$$\mathcal{E}_{\text{diff}}(r, \theta_B, \varphi_B, t') \propto \int \mathrm{d}\omega \; c(\omega)\omega \; f(\omega, \theta_B, \varphi_B) \frac{e^{i\omega r/c}}{r} \; \frac{1}{\omega - \omega_0 + i\Gamma/2} \, e^{-i\omega t'} \tag{51}$$

It will be useful for what follows to regroup the two exponentials of (50) and to set:

$$\tilde{t} = t' - r/c \tag{52}$$

We then get:

$$\mathcal{E}_{\text{diff}}(r, \theta_B, \varphi_B, t') = \int \mathrm{d}\omega \; c(\omega)\omega \; f(\omega, \theta_B, \varphi_B) \frac{1}{r} \; \frac{1}{\omega - \omega_0 + i\Gamma/2} \, e^{-i\omega \tilde{t}} \tag{53}$$

which means that the wave packet scattered along the direction θ_B, φ_B moves at velocity c, and that its amplitude decreases as $1/r$.

γ. *Spatial and temporal dependence of the scattered wave packet*

We now assume that the frequency width $\Delta\omega$ of the incident wave packet is much smaller than its average frequency ω_m:

$$\Delta\omega \ll \omega_m \tag{54}$$

but we do not make any hypothesis concerning the relative values of $\Delta\omega$ and Γ. The factor $\sqrt{\omega}$ appearing in the previous relations can now be replaced by $\sqrt{\omega_m}$, and comes out of the integral. We can also neglect the variations with ω of $f(\omega, \theta_B, \varphi_B)$ over the Γ interval where the function $(\omega - \omega_0 + i\Gamma/2)^{-1}$ varies significantly. The scattered field $\mathcal{E}_{\text{diff}}(r, \tilde{t})$ can then been seen as the temporal Fourier transform of the product of two functions $c(\omega)$ and $(\omega - \omega_0 + i\Gamma/2)^{-1}$. This field is the convolution of the Fourier transforms of these two functions of ω. Taking into account (36) and (37), the Fourier transform of the first function is:

$$c(\omega) \quad \underset{FT}{\longleftrightarrow} \quad \mathcal{E}(\tilde{t}) = e^{-i\omega_m \tilde{t}} \; \mathcal{F}(\tilde{t}) \tag{55}$$

For the second function, we get:

$$\frac{1}{\omega - \omega_0 + i\Gamma/2} \quad \underset{FT}{\longleftrightarrow} \quad e^{-i\omega_0 \tilde{t}} \; \theta(\tilde{t}) \; e^{-\Gamma \tilde{t}/2} \tag{56}$$

where $\theta(\tilde{t})$ is the Heaviside function, equal to 1 for $\tilde{t} > 0$, and to 0 for $\tilde{t} < 0$. This leads to:

$$\mathcal{E}_{\text{diff}}(r, \tilde{t}) \propto \frac{1}{r} \int_{-\infty}^{+\infty} \mathrm{d}\tau \; e^{-i\omega_m \tau} \; \mathcal{F}(\tau) \; e^{-i\omega_0 (\tilde{t} - \tau)} \; \theta(\tilde{t} - \tau) \; e^{-\Gamma(\tilde{t} - \tau)/2} \tag{57}$$

δ. Study of two limiting cases

Two interesting cases occur when the width $\Delta\omega$ of the incident wave packet is either very large or very small compared to the natural width Γ of the excited state of atom A.

$\Delta\omega \gg \Gamma$ ***limit***

The incident wave packet passes through a given point in a time $1/\Delta\omega$ that is very short compared to the radiative lifetime $1/\Gamma$ of the excited state. The envelope $\mathcal{F}(\tau)$ of the incident wave packet is different from zero only during a time interval much shorter than the characteristic times of the Fourier transform of $(\omega - \omega_0 + i\Gamma/2)^{-1}$. We can thus set $\tau = 0$ in the last two terms of (57), which yields:

$$\mathcal{E}_{\text{diff}}(r,\tilde{t}) \propto \frac{1}{r}\left[\int_{-\infty}^{+\infty} \mathrm{d}\tau\ e^{i(\omega_0-\omega_m)\tau}\ \mathcal{F}(\tau)\right]\ \left[e^{-i\omega_0\tilde{t}}\ \theta(\tilde{t})\ e^{-\Gamma\tilde{t}/2)}\right] \tag{58}$$

The term in the first bracket is proportional to the excitation amplitude of the scattering atom by the incident wave packet. The second bracket describes a free oscillation at the atomic frequency ω_0, starting at time $\tilde{t} = 0$ and damped over a time $2/\Gamma$.

The physical meaning of this result is as follows. The incident wave packet spends a very short time near atom A, and hence excites it in a percussive manner before moving away with velocity c. Once the incident wave packet is gone, the atomic dipole thus excited oscillates freely at frequency ω_0, until it is damped by spontaneous emission. This situation is the analog of the percussive excitation of an oscillator in classical mechanics.

$\Delta\omega \ll \Gamma$ ***limit***

We can now replace in (57) the function $\mathcal{F}(\tau)$ by $\mathcal{F}(\tilde{t})$ as $\tilde{t} - \tau$ cannot be larger, in modulus, than $1/\Gamma$. This is because of the presence of the last exponential term in (57) and the fact that, when $\Delta\omega \ll \Gamma$, the envelope of $\tilde{\mathcal{E}}(\tilde{t})$ varies very slowly over that time interval. One can then rewrite (57) in the form:

$$\mathcal{E}_{\text{diff}}(r,\tilde{t}) \propto \mathcal{F}(\tilde{t})e^{-i\omega_m\tilde{t}}\int_{-\infty}^{+\infty} \mathrm{d}\tau\ e^{i\omega_m(\tilde{t}-\tau)}e^{-i\omega_0(\tilde{t}-\tau)}\ \theta(\tilde{t}-\tau)\ e^{-\Gamma(\tilde{t}-\tau)/2)} \tag{59}$$

Let us make the change of variable $\tau' = \tilde{t} - \tau$ in the integral over τ. Taking (56) into account, we see that this integral is actually the Fourier transform of $e^{-i\omega_0\tau'}\theta(\tau')e^{-\Gamma\tau'/2}$ calculated at ω_m, which is $(\omega_m - \omega_0 + i\Gamma/2)^{-1}$. This leads to:

$$\mathcal{E}_{\text{diff}}(r,\tilde{t})\ \propto\ \mathcal{F}(\tilde{t})\, e^{-i\omega_m\tilde{t}}\ \frac{1}{\omega_m - \omega_0 + i\Gamma/2} \tag{60}$$

The physical meaning of this result is as follows. When $\Delta\omega \ll \Gamma$, the wave packet takes a long time passing atom A, whose dipole undergoes forced oscillation at the frequency ω_m. It thus emits radiation at the same frequency, with an amplitude that follows adiabatically the slow variation of the envelope $\mathcal{F}(\tilde{t})$ of the incident wave packet; this explains the first term in (60). The second term describes the linear response of the dipole with eigenfrequency ω_0 and damping time $2/\Gamma$ to an excitation of frequency ω_m. In this case, the oscillator's amplitude follows adiabatically the excitation's amplitude.

5. Example of wave packets with two entangled photons

In § 2-c, we considered two-photon states that were tensor products of two one-photon wave packets; the two photons described by these states were not entangled. There obviously exist a number of two-photon states that cannot be described in the form of a product of two one-photon states, and which thus describe entangled photons. In this last section, we shall focus on such an example where the entangled photons appear in an optical nonlinear process, called *parametric down-conversion*. This process has the advantage of producing pairs of photons bunched in time. Detecting one photon of the pair at a given time t allows predicting the second photon will be detected to within a very short time interval.

5-a. Parametric down-conversion

Computations involved in parametric down-conversion are similar to computations we already discussed. We shall simply outline the general ideas allowing a physical understanding of the process, without going into details that would unnecessarily lengthen the present complement.

α. Description of the process

In § E-1 of Chapter XX, we studied the elastic scattering of a photon by an atom. Figures 2 and 3 of Chapter XX show two possible diagrams representing such a process: an incident photon with angular frequency ω_i is absorbed and an ω_f photon emitted while the atom goes back to its initial level. Energy conservation of the total system atom + photon requires that $\omega_i = \omega_f$. In the present complement, we study a nonlinear scattering process during which, as before, an incident photon of angular frequency ω_0 is absorbed by an atom in state a, but where there are now two photons with angular frequencies ω_1 and ω_2 that are emitted. At the end of the scattering process the atom returns to state a; energy conservation now requires that $\omega_0 = \omega_1 + \omega_2$. Figure 1 gives two possible representations of such a process, analogous to those of Figure 2.α and 3.α in Chapter XX.

Several temporal orders are possible for the absorption and emission processes. For example, Figures 3.α and 3.β of Chapter XX do not have the same temporal order for the absorption and emission occurring in the scattering of one photon. For the three-photon process considered here, including one absorption and two emissions, $3! = 6$ possible temporal orders should, a priori, be considered; Figure 1 represents only one of these six possible orders.

β. Scattering amplitude

The principle for calculating the scattering amplitude of parametric down-conversion is similar to the one that led us to formulas (E-3) to (E-5) of Chapter XX, but involves now three, instead of two, interactions with the field; two relay states (instead of one) come into play. As an example, for the process represented in Figure 1, we must consider the following states:

- initial state: $|a;\omega_0\rangle$, with energy $E_{\rm in} = E_a + \hbar\omega_0$
- first relay state: $|b;0\rangle$, with energy $E_{\rm rel\ 1} = E_b$

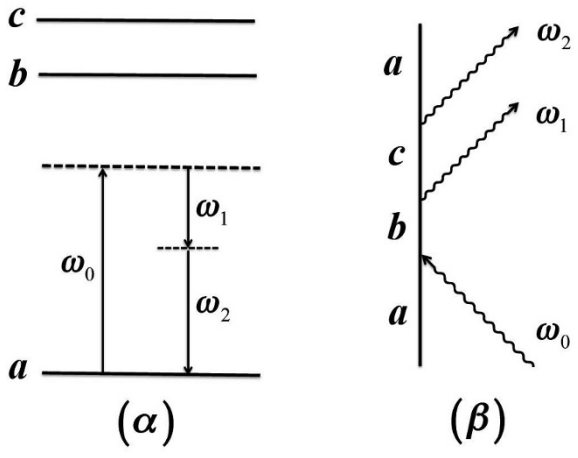


Figure 1: An incident photon, with angular frequency ω_0 is scattered by an atomic system in the initial state a. At the end of the scattering process, the atomic system has returned to state a, while two new photons have appeared with angular frequencies ω_1 and ω_2. Energy conservation requires that $\omega_0 = \omega_1 + \omega_2$. In the left hand side of the figure, the absorption (emission) processes are represented with upwards (downwards) arrows; in the right-hand side, these processes are shown with incoming (outgoing) wiggly arrows that also symbolize the photon propagation.

- second relay state: $|c;\omega_1\rangle$, with energy $E_{\text{rel }2} = E_c + \hbar\omega_1$
- final state: $|a;\omega_1,\omega_2\rangle$, with energy $E_{\text{fin}} = E_a + \hbar\omega_1 + \hbar\omega_2$

The probability amplitude associated with this process is obtained by generalizing relation (E-4) of Chapter XX. Within a constant and non-significant factor, it is the product of a function $\delta(\omega_1+\omega_2-\omega_0)$, imposing energy conservation[7], by the following expression:

$$-\left(\frac{\hbar}{2\varepsilon_0 L^3}\right)^{3/2}\sqrt{\omega_0\omega_1\omega_2}\sum_{b,c}\frac{\langle a|\boldsymbol{\varepsilon}_2^*\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_1^*\cdot\boldsymbol{D}|b\rangle\langle b|\boldsymbol{\varepsilon}_0\cdot\boldsymbol{D}|a\rangle}{(E_a+\hbar\omega_2-E_c)(E_a+\hbar\omega_0-E_b)} \tag{61}$$

where $\boldsymbol{\varepsilon}_0$, $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are, respectively, the polarizations of the photons of frequencies ω_0, ω_1 and ω_2. Compared to relation (E-4) of Chapter XX, expression (61) now contains three (instead of two) matrix elements in the numerator, and two (instead of one) energy denominators containing the differences in energy between the initial state and either the relay 1 or the relay 2 state.

Six similar amplitudes can be written, generalizing equations (E-3) to (E-5) of Chapter XX and corresponding to different temporal orders for the absorption and emission processes. Once they have been added, one must also sum these amplitudes over all

[7] As in § 4-a, we assume the total interaction time Δt to be sufficiently long for the function $\delta^{(\Delta t)}$ to be assimilated to a real delta function.

the atomic relay states b and c. All the contributions to the total amplitude contain the same function $\delta(\omega_1 + \omega_2 - \omega_0)$.

The final state of the system "atom + radiation" at the end of the scattering process is the sum over ω_1 and ω_2 of the components thus obtained, with the condition $\omega_1 + \omega_2 = \omega_0$. It can be written:

$$|\Psi\rangle = |a\rangle \otimes |\psi_R\rangle \tag{62}$$

where:

$$|\psi_R\rangle = \sum_{\mathbf{k}_1,\mathbf{k}_2} \delta(\omega_1 + \omega_2 - \omega_0)\ g(\mathbf{k}_1, \mathbf{k}_2)|\mathbf{k}_1, \mathbf{k}_2\rangle \tag{63}$$

with $\omega_{1,2} = c\,|\mathbf{k}_{1,2}|$. This state cannot be written as the product of two field states; it is therefore entangled (see Chapter XXI).

The function $g(\mathbf{k}_1, \mathbf{k}_2)$ characterizing the field state is the result of the ω_1 and ω_2 dependence of the scattering amplitudes, as well as of the density of final states appearing in the summations over the continuums[8] ω_1 and ω_2 (summation over the moduli of the two vectors $\mathbf{k}_1$ and $\mathbf{k}_2$). We assume here that the energies of all the relay states are far away from any resonance, so that the $g(\mathbf{k}_1, \mathbf{k}_2)$ dependence on $|\mathbf{k}_1|$ and $|\mathbf{k}_2|$ does not present any kind of narrow structure. In other words, all the energy differences $\Delta E = E_{\text{in}} - E_{\text{rel}}$ in the denominators of the scattering amplitudes are of the order of a fraction of an optical frequency. We mentioned in comment (i) at the end of § 1 in Complement A_{XX} that the time spent in a relay state during the scattering process is, according to the time-energy uncertainty relation, of the order of $\hbar/|\Delta E|$. This means that the times separating the emission of the two photons ω_1 and ω_2 cannot differ by more than a few optical periods, i.e. a few tens of femtoseconds. This qualitative argument shows that the two photons ω_1 and ω_2 are emitted quasi-simultaneously.

Comment:

If the interaction Hamiltonian appearing in the three matrix elements of the scattering amplitude is the electric dipole Hamiltonian, and if the atomic states have a well defined parity, this atomic parity changes with each interaction. After the three interactions, the parity is therefore changed, which forbids the final atomic state to be the same as the initial state. Consequently, the parametric down-conversion process we just studied can occur only when the atomic states do not have a well defined parity. Such a situation is encountered when the atomic Hamiltonian is not invariant upon reflection. This happens for example when the atom is inserted in a crystal where the local crystalline field, which has the symmetry of an external electric field, is not invariant upon spatial reflection.

5-b. Temporal correlations between the two photons generated in parametric down-conversion

We now compute the double detection signal from the two photons generated in parametric down-conversion. The experiment we analyze is schematized in Figure 2. An incident pump beam, with frequency ω_0, propagates along a direction with unit vector

[8] Two continuums of final states come into play in this problem, but the condition $\omega_1 + \omega_2 = \omega_0$ reduces it to one.

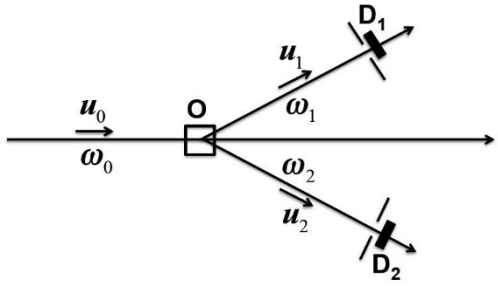


Figure 2: A pump beam with angular frequency ω_0, propagating along a direction with unit vector $\boldsymbol{u}_0$, impinges on a nonlinear crystal placed in O. The parametric down-conversion process generates two beams of frequencies ω_1 and ω_2, with $\omega_1 + \omega_2 = \omega_0$. Diaphragms allow fixing the directions $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ of these two beams. The two detectors D_1 and D_2 register the arrivals of the photons and permit studying their temporal correlations.

$\boldsymbol{u}_0$, and impinges onto a nonlinear crystal O containing atoms performing the conversion. Two diaphragms placed in front of the two detectors D_1 and D_2 allow selecting two directions, with unit vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, for the two beams generated by parametric down-conversion.

We focus on the temporal, rather than spatial, aspect of the phenomenon. For the sake of simplicity, we assume that the three field states appearing in Figure 2 are plane waves, infinite in the two transverse directions. The only variable characterizing the modes involved is therefore the longitudinal component of the vector $\mathbf{k}$ or, equivalently, the frequency ω. The incident photon is described by a wave packet, characterized in frequency space by a real function $c(\omega_0)$, centered at ω_m^0 and of width $\Delta\omega_0 \ll \omega_m^0$. The center of the incident wave packet arrives at crystal O at time $t = 0$, and passes through it in a time of the order of:

$$\Delta t \simeq 1/c\Delta k = 1/\Delta\omega_0 \tag{64}$$

The two-photon wave packet generated by parametric down-conversion is described by an expression similar to (63), in which we now use the variables $\omega_{1,2} = c\,|\mathbf{k}_{1,2}|$; this wave packet depends, to a certain extent, on ω_1 and ω_2 via the function $g(\omega_1, \omega_2)$.

α. *Double photodetection signal $w_{II}(\boldsymbol{r}_1, t; \boldsymbol{r}_2, t+\tau)$*

We first compute the absorption amplitude of the two photons[9], one at time t by detector D_1 located at $\boldsymbol{r}_1$, the other at time $t+\tau$ by detector D_2 located at $\boldsymbol{r}_2$:

$$\langle 0|E^{(+)}(\boldsymbol{r}_2, t+\tau)E^{(+)}(\boldsymbol{r}_1, t)|\psi_R\rangle = \frac{\hbar}{2\varepsilon_0 L^3}\sum_{\omega_0}\sum_{\omega_1,\omega_2} c(\omega_0)\ \sqrt{\omega_1\omega_2}\ g(\omega_1,\omega_2)$$
$$\times\, e^{i[\boldsymbol{k}_2\cdot\boldsymbol{r}_2-\omega_2(t+\tau)]}\ e^{i[\boldsymbol{k}_1\cdot\boldsymbol{r}_1-\omega_1 t]}\ \delta(\omega_1+\omega_2-\omega_0) \tag{65}$$

[9]The signal w_{II} is the squared modulus of this amplitude.

In this equation, $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are the wave vectors of the two photons propagating freely along $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. We note L_1 and L_2 the distances between O and D_1, O and D_2, and define $T_1 = L_1/c$ and $T_2 = L_2/c$ as the times taken by the photons to travel these two distances. We have:

$$\begin{aligned} \boldsymbol{k}_1 \cdot \boldsymbol{r}_1 &= \frac{\omega_1}{c} L_1 = \omega_1 T_1 \\ \boldsymbol{k}_2 \cdot \boldsymbol{r}_2 &= \frac{\omega_2}{c} L_2 = \omega_2 T_2 \end{aligned} \tag{66}$$

so that (65) can be rewritten in the form:

$$\begin{aligned} \langle 0|E^{(+)}(\boldsymbol{r}_2, t+\tau)E^{(+)}(\boldsymbol{r}_1, t)|\psi_R\rangle = \frac{\hbar}{2\varepsilon_0 L^3} \sum_{\omega_0} \sum_{\omega_1,\omega_2} c(\omega_0)\ \sqrt{\omega_1\omega_2}\ g(\omega_1,\omega_2) \\ \times\, e^{i\omega_2[T_2-t-\tau]} e^{i\omega_1[T_1-t]}\ \delta(\omega_1+\omega_2-\omega_0) \end{aligned} \tag{67}$$

We now replace the two variables ω_1 and ω_2 by a single variable $\delta\omega$ by setting:

$$\begin{aligned} \omega_1 &= \frac{\omega_0}{2} + \delta\omega \\ \omega_2 &= \frac{\omega_0}{2} - \delta\omega \end{aligned} \tag{68}$$

Condition $\omega_1 + \omega_2 = \omega_0$ is then automatically satisfied, so that the delta function appearing on the second line of (67) is no longer necessary. The summation over ω_1 and ω_2 becomes a summation over $\delta\omega$, and the function $g(\omega_1, \omega_2)$ is replaced by a function $\widetilde{g}(\delta\omega)$. If we assume, for the sake of simplicity, that $T_1 = T_2 = T$, we finally obtain:

$$\begin{aligned} \langle 0|E^{(+)}(\boldsymbol{r}_2, t+\tau)E^{(+)}(\boldsymbol{r}_1, t)|\psi\rangle \propto \frac{\hbar}{2\varepsilon_0 L^3} \sum_{\omega_0} c(\omega_0)\ e^{i\omega_0[T-t-\tau]/2} e^{i\omega_0[T-t]/2} \\ \times \sum_{\delta\omega} \sqrt{(\frac{\omega_0}{2}+\delta\omega)(\frac{\omega_0}{2}-\delta\omega)}\ \widetilde{g}(\delta\omega)\ e^{i\delta\omega\ \tau} \end{aligned} \tag{69}$$

β. Discussion

The τ dependence of the double photodetection signal is given by the summation over $\delta\omega$ on the second line of (69). Going to the continuous limit, it is therefore the Fourier transform of $\sqrt{(\frac{\omega_0}{2}+\delta\omega)(\frac{\omega_0}{2}-\delta\omega)} \times \widetilde{g}(\delta\omega)$. Nevertheless, the product $\sqrt{(\frac{\omega_0}{2}+\delta\omega)(\frac{\omega_0}{2}-\delta\omega)}$ varies very slowly with $\delta\omega$ and can be taken as constant, as can be the state densities introduced when replacing the discrete summation by an integral. We also saw above (§ 5-a-β) that the variation of g as a function of ω_1 and ω_2, hence the variation of $\widetilde{g}(\delta\omega)$, is very slow as long as no resonant (or quasi-resonant) relay states are involved in the scattering process. We must thus take the Fourier transform of a function of $\delta\omega$ that has a very large width, of the order of a fraction of the optical frequency. This means that the double photodetection signal is different from zero only if the two photodetections are separated by a time interval of the order of a few optical periods. In other words, the two detections are always quasi-simultaneous.

Consider now the summation over ω_0 in the first line of (69). We are going to see that the t dependence of the signal involves time scales much longer than those

characterizing the variation with τ of the signal w_{II}. To show this, we replace τ by 0 on the right-hand side; after going to the continuous limit, we get:

$$\int \mathrm{d}\omega_0 \; c(\omega_0) \; e^{i\omega_0(T-t)} \tag{70}$$

which is the Fourier transform of the incident wave packet. This packet arrives at point O at $t = 0$, and for the entire packet to pass that point, it takes a certain time Δt; this time interval Δt is much longer, in general, than the time characterizing the τ dependence of signal w_{II}. Relation (70) thus indicates that both detectors yield (almost simultaneously) a signal at any time t within a time interval Δt centered around $t = T$; this time corresponds to the arrivals at D_1 and D_2 of the photons generated at O by the incident wave packet. But once a photon is detected by one of the detectors, the other photon is detected practically at the same instant by the other detector. Such a temporal correlation could not be predicted by a semiclassical treatment.

These results remain valid when the parametric down-conversion process is produced, not by a single incident photon described by a wave packet, but rather by a continuous laser excitation. The two beams generated by the parametric down-conversion process then contain a series of pairs of photons, that are detected at the same instant; they are referred to as *twin photons*.

Such *twin beams* can excite two-photon transitions in a much more efficient way than ordinary beams. This is because, in the absence of resonant relay states in the two-photon absorption process, an argument similar to that presented above shows that the two absorptions must be separated by a very short time interval (the two photons must interact quasi-simultaneously with the absorbing atom). The two incident photons must impinge on the atom at exactly the same time, which can be the case for twin beams (with ordinary beams, one can only observe two-photon absorptions due to accidental coincidences)

In practice, radiation parametric down-conversion is often performed, not on an isolated atom, but rather on atoms or molecules in a solid. It is then imperative to take into account the interference between beams generated in different parts of the solid, and identify the conditions for getting a constructive interference. The refractive optical index of the medium in which the beams propagate then plays an important role, which leads to the so called *phase matching condition*. This discussion, outside the scope of the present complement, is treated in detail in quantum optics books [65] [66].

Chapter XXI

Quantum entanglement, measurements, Bell's inequalities

A Introducing entanglement, goals of this chapter 2188

B Entangled states of two spin-1/2 systems 2190

B-1 Singlet state, reduced density matrices 2191

B-2 Correlations 2191

C Entanglement between more general systems 2193

C-1 Pure entangled states, notation 2193

C-2 Presence (or absence) of entanglement: Schmidt decomposition 2193

C-3 Characterization of entanglement: Schmidt rank 2196

D Ideal measurement and entangled states 2196

D-1 Ideal measurement scheme (von Neumann) 2196

D-2 Coupling with the environment, decoherence; "pointer states" 2199

D-3 Uniqueness of the measurement result 2201

E "Which path" experiment: can one determine the path followed by the photon in Young's double slit experiment? 2202

E-1 Entanglement between the photon states and the plate states 2203

E-2 Prediction of measurements performed on the photon 2204

F Entanglement, non-locality, Bell's theorem 2204

F-1 The EPR argument 2205

F-2 Bohr's reply, non-separability 2207

F-3 Bell's inequality 2208

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

We discuss in this last chapter an essential concept of quantum mechanics, entanglement. This will highlight a number of aspects of quantum mechanics that have no equivalent in classical physics.

A. Introducing entanglement, goals of this chapter

Consider two physical systems A and B, each having a state space $\mathcal{E}_A$ and $\mathcal{E}_B$; they can be grouped into a total system $A+B$, whose state space is the tensor product $\mathcal{E}_A \otimes \mathcal{E}_B$. If we assume system A is described by a normalized quantum state $|\varphi_A\rangle$ belonging to $\mathcal{E}_A$, system B by a normalized quantum state $|\chi_B\rangle$ belonging to its space state $\mathcal{E}_B$, the ket $|\Phi\rangle$ describing the total system is the tensor product:

$$|\Phi\rangle = |\varphi_A\rangle \otimes |\chi_B\rangle \tag{A-1}$$

In this case, each of the three physical systems A, B, and $A + B$ is described by a state vector, which is the most precise possible description in quantum mechanics.

The situation is different when the state vector of the global system is no longer a simple product. Let us note $|\varphi_A\rangle$, $|\zeta_A\rangle$, .., orthonormalized quantum states belonging to the state space $\mathcal{E}_A$ of the first system, and $|\chi_B\rangle$, $|\xi_B\rangle$,..., orthonormalized quantum states belonging to the state space $\mathcal{E}_B$ of the second. We can then build products different from (A-1) for the state of the global system, for example:

$$|\Phi'\rangle = |\zeta_A\rangle \otimes |\xi_B\rangle \tag{A-2}$$

Now, we can also, in view of the superposition principle, form any linear combination $|\Psi\rangle$ of $|\Phi\rangle$ and $|\Phi'\rangle$, which will no longer be a simple product:

$$|\Psi\rangle = \alpha\,|\varphi_A\rangle \otimes |\chi_B\rangle + \beta\,|\zeta_A\rangle \otimes |\xi_B\rangle \tag{A-3}$$

In this relation, the complex coefficients α and β can take on any value, as long as they obey the normalization condition:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{A-4}$$

We shall assume, however, that neither of these coefficients is zero so that (A-3) is not reduced to a simple product:

$$\alpha\beta \neq 0 \tag{A-5}$$

A state such as (A-3), which contains a coherent superposition of two (or more) components, each component being a product, is called an "entangled state". The general property associated with these states is called "quantum entanglement". It expresses the fact that the quantum state of each subsystem is, in a way, conditioned by the state of the other.

In Complement $\mathrm{E_{III}}$, we introduced the concept of a density operator, which provides a more general description of a physical system than a state vector. The density operator of the total physical system $A + B$, whose state vector is known, is simply the projector onto $|\Psi\rangle$:

$$\rho_{A+B} = |\Psi\rangle\,\langle\Psi| \tag{A-6}$$

whose trace is equal to one:

$$\mathrm{Tr}\left\{\rho_{A+B}\right\} = \langle\Psi\,|\Psi\rangle = 1 \tag{A-7}$$

When a physical system can be described by a state vector, it is said to be in a "pure state". Its density operator obeys the relation:

$$\left[\rho_{A+B}\right]^2 = \rho_{A+B} \tag{A-8}$$

and thus:

$$\mathrm{Tr}\left\{\left[\rho_{A+B}\right]^2\right\} = 1 \tag{A-9}$$

Under such conditions, we can choose to describe the total physical system either by its state vector $|\Psi\rangle$, or by the density operator ρ_{A+B}. We are going to show that this is no longer the case for the two subsystems A and B, for which only the density operator can be used.

Imagine, for example, that we are only interested in measurements performed on subsystem A. We saw in § 5-b of Complement $\mathrm{E_{III}}$ that, when the total system is entangled as in (A-3), there generally does not exist any state vector belonging to $\mathcal{E}_A$ that allows computing the probabilities of measurements performed solely on A. Instead, one must necessarily use a density operator ρ_A obtained by taking a partial trace (taken over the state space $\mathcal{E}_B$ of the non-observed system – we recall in § B-1 below how to compute the matrix elements of a partial trace):

$$\rho_A = \mathrm{Tr}_B\left\{\rho_{A+B}\right\} \tag{A-10}$$

Like operator (A-6), this operator is Hermitian, non-negative, and its trace is equal to one; it is however not the projector onto a single state vector. When the system is described by the entangled state (A-3), this density operator is given by:

$$\rho_A = |\alpha|^2\,|\varphi_A\rangle\,\langle\varphi_A| + |\beta|^2\,|\zeta_A\rangle\,\langle\zeta_A| \tag{A-11}$$

which is actually the sum of two projectors. Consequently, subsystem A is in state $|\varphi_A\rangle$ with a probability $|\alpha|^2$, and in state $|\zeta_A\rangle$ with a probability $|\beta|^2$: as opposed to the state of $A+B$, the quantum state of A is not known with certainty, but only with a certain probability. The density operator ρ_A can be called a "statistical mixture", underlying the fact that the results of measurements performed on A are predicted by computing averages on the (non-observed) properties of B. We then get the inequality:

$$\mathrm{Tr}\left\{\left[\rho_A\right]^2\right\} \leq 1 \tag{A-12}$$

where the equality occurs only if one of the two coefficients α or β is zero; the equivalent of relation (A-9) is, in general, not satisfied by ρ_A. Inequality (A-12) expresses the fact that, as the quantum state of A is known only in a statistical way, the quantum description of A is less precise than the description of the total system $A+B$. This discussion can be easily generalized to the case where $|\Psi\rangle$ is the superposition of, not only two components as in (A-3), but of three or more.

We find ourselves in a situation that might look rather surprising, as it does not have any equivalent in classical physics. We do know that a perfect classical description

of the total system $A + B$ automatically implies that each of its two subsystems is also perfectly described. This is because a complete description of the state of the total system is simply the collection of all the complete descriptions of its subsystems. As an example, a perfect classical description of the solar system is simply the knowledge of all the positions and velocities of the planets, satellites, and all their constituent particles. In quantum mechanics, things are drastically different: the most precise description of the total system by a state vector (pure state) does not imply, in general, that its subsystems can be described with the same precision. This difference radically changes the usual relation between the parts and the whole of a physical system. Schrödinger, who first introduced in 1935 the words "quantum entanglement" commented on this new concept [67]: "As far as I am concerned, I would not call this property *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction, the two representatives [the quantum states] have become entangled... Another way of expressing the peculiar situation is: the best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts, even though they may be entirely separate and therefore virtually capable of being 'best possibly known' i.e., of possessing, each of them, a representative (state vector) of its own."

In a general way, when the state vector of a global system is not a simple product, and quantum entanglement occurs, the quantum predictions for observations on part of the system can become rather unexpected. This chapter will discuss a number of special physical effects related to quantum entanglement. As a general introduction, § B studies the simple case of two spin-1/2 systems entangled in a singlet state. This example is generalized, in § C, to any physical system, and we present some of the properties of entangled quantum states. The relations between entanglement and quantum measurements is discussed in § D, using in particular the ideal measurement scheme proposed by von Neumann. In § E, we describe an experiment where one tries to observe interference fringes of a particle going through a two-slit plate while determining at the same time which slit the particle went through; if this were possible, one would face a contradiction. However, a partial trace operation on an entangled state of the particle and the plate allows proving the coherence of the quantum formalism and illustrates an aspect of complementarity. Finally, § F discusses the relations between entanglement and quantum non-locality, in the framework of the general Einstein, Podolsky and Rosen argument, and of Bell's theorem.

B. Entangled states of two spin-1/2 systems

We first discuss a very simple case, which will prove useful for the rest of the chapter: each of the two systems A and B is a spin 1/2; each state space $\mathcal{E}_{A,B}$ is then spanned by the two eigenstates $|\pm\rangle$ of the spin component on the Oz axis. We assume these two states are entangled in a singlet state, such as the one written in relation (B-22) of Chapter X:

$$\begin{aligned} |\Psi\rangle &= \frac{1}{\sqrt{2}}\Big[\,|A:+\rangle \otimes |B:-\rangle - |A:-\rangle \otimes |B:+\rangle\Big] \\ &= \frac{1}{\sqrt{2}}\Big[\,|+,-\rangle - |-,+\rangle\Big] \end{aligned} \tag{B-1}$$

(on the second line, we have simplified the notation, assuming that the first index in the ket refers to spin A, and the second to spin B).

B-1. Singlet state, reduced density matrices

In the basis of the 4 kets $|+,+\rangle$, $|+,-\rangle$, $|-,+\rangle$, $|-,-\rangle$ taken in that order, the matrix representing the density operator ρ_{A+B} is written:

$$(\rho_{A+B}) = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{B-2}$$

It is easy to check, performing a matrix product, that $(\rho_{A+B})^2 = (\rho_{A+B})$, hence that relation (A-9) is verified: this means the total system is in a pure state.

As indicated in § 5-b of Complement E_{III}, the matrix representing the density operator ρ_A is obtained by taking a partial trace, i.e. by adding the matrix elements of (ρ_{A+B}) that are diagonal with respect to the quantum numbers of the second spin (this amounts to summing over the states of the non-observed spin):

$$\langle m_A|\,\rho_A\,|m'_A\rangle = \sum_{m_B} \langle m_A, m_B|\,\rho_{A+B}\,|m'_A, m_B\rangle \tag{B-3}$$

This leads to:

$$(\rho_A) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{B-4}$$

We then get:

$$(\rho_A)^2 = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{B-5}$$

and thus $\mathrm{Tr}\left\{(\rho_A)^2\right\} = 1/2$; this means that spin A is not in a pure state. By symmetry, the same result would obviously be found for (ρ_B). Note that after taking the partial trace, all the non-diagonal elements (coherences) of (B-2) have completely disappeared. When they are considered as an isolated system, each of the two spins is in a "completely depolarized" state, and measuring its Oz spin component (or any of its spin component, for that matter) will yield the results + or − with the same probability 1/2. At the level of each individual spin, the minus sign that characterizes the entanglement of the state vector (B-1) becomes irrelevant; on the other hand, we are going to show that this entanglement yields very strong correlations between the results of measurements pertaining to both spins.

B-2. Correlations

Imagine now that we perform simultaneously measurements on both spins, the first one along a direction in the plane xOz, making an angle θ_A with the Oz axis, the second along a direction in that same plane, making an angle θ_B with Oz. The results

we are about to obtain will be important for the discussion of Bell's theorem in § F-3-a. Relations (A-22) of Chapter IV (taking the angle φ equal to zero and $\theta = \theta_{A,B}$) yield the expressions for the eigenvectors of the measurements in the state spaces $\mathcal{E}_A$ and $\mathcal{E}_B$:

$$\begin{aligned} |+\rangle_\theta &= \cos\frac{\theta}{2}\,|+\rangle + \sin\frac{\theta}{2}\,|-\rangle \\ |-\rangle_\theta &= -\sin\frac{\theta}{2}\,|+\rangle + \cos\frac{\theta}{2}\,|-\rangle \end{aligned} \tag{B-6}$$

In the space of the two-spin states, the ket corresponding to a double + result of the measurement is written:

$$\begin{aligned} |+\rangle_{\theta_A} \otimes |+\rangle_{\theta_B} = \cos\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\,|+,+\rangle + \cos\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\,|+,-\rangle \\ + \sin\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\,|-,+\rangle + \sin\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\,|-,-\rangle \end{aligned} \tag{B-7}$$

This means that, when the system is in the singlet state (B-1), the probability amplitude for obtaining this double result is:

$$\begin{aligned} \left[\langle+|_{\theta_A} \otimes \langle+|_{\theta_B}\right]|\Psi\rangle &= \frac{1}{\sqrt{2}}\left[\cos\frac{\theta_A}{2}\sin\frac{\theta_B}{2} - \sin\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\right] \\ &= -\frac{1}{\sqrt{2}}\sin\frac{\theta_A-\theta_B}{2} \end{aligned} \tag{B-8}$$

The probability of the result $(+,+)$ when measuring the components of both spins along the θ_A and θ_B directions is therefore:

$$\mathcal{P}_{++}(\theta_A,\theta_B) = \frac{1}{2}\sin^2\frac{\theta_A-\theta_B}{2} \tag{B-9}$$

One can redo the same calculations for the three other possible pairs of results, $(+,-)$, $(-,+)$ and $(-,-)$. This does not present any difficulty but it is easier to note that changing θ_A into $\theta_A+\pi$ exchanges the two eigenkets of (B-6), and hence the results + and − for the first spin A; the same operation can be done with the second spin B. We then make these changes in (B-9) and get the probabilities for the 4 possible results in the form:

$$\begin{aligned} \mathcal{P}_{++}(\theta_A,\theta_B) &= \mathcal{P}_{--}(\theta_A,\theta_B) = \frac{1}{2}\sin^2\frac{\theta_A-\theta_B}{2} \\ \mathcal{P}_{+-}(\theta_A,\theta_B) &= \mathcal{P}_{-+}(\theta_A,\theta_B) = \frac{1}{2}\cos^2\frac{\theta_A-\theta_B}{2} \end{aligned} \tag{B-10}$$

When both spins are measured, strong correlations between the results appear[1]. These correlations are the direct consequence of the entanglement present in the singlet state vector (B-1).

[1]The probabilities cannot, in general, be factored. As an example, relation (B-10) shows that $\mathcal{P}_{++}/\mathcal{P}_{-+}$ is different from $\mathcal{P}_{+-}/\mathcal{P}_{--}$. This means that the ratio of the probabilities of obtaining for the first spin the result + or the result − depends on the state of the second spin, clearly showing correlations.

C. Entanglement between more general systems

The concept of entanglement is obviously not limited to the singlet state of two spin-1/2 particles. We now study how to characterize the presence of entanglement when the total system is in a pure state.

C-1. Pure entangled states, notation

We consider two quantum systems A and B belonging, respectively, to state spaces $\mathcal{E}_A$ (with dimension P) and $\mathcal{E}_B$ (with dimension Q). The normalized state vector $|\Psi\rangle$ describing the total system $A+B$ belongs to the tensor product space $\mathcal{E}_A \otimes \mathcal{E}_B$, with dimension PQ. Some of the states $|\Psi\rangle$ can be written as a simple product:

$$|\Psi\rangle = |\varphi_A\rangle \otimes |\varphi_B\rangle \tag{C-1}$$

where $|\varphi_A\rangle$ and $|\varphi_B\rangle$ are any normalized kets of $\mathcal{E}_A$ and $\mathcal{E}_B$, respectively. In such a case, the two physical subsystems A and B are not entangled; each of them, as well as the total state, can be described by a state vector (pure state). On the other hand, the majority of the states $|\Psi\rangle$ cannot be factored this way, and must necessarily be written as a sum of products (the singlet state studied above is such an example); the two subsystems A and B are then entangled.

It is not always obvious to guess from the expression of any given state vector $|\Psi'\rangle$ if it can actually be written as a simple tensor product. This ket has, in general, $P \times Q$ components, and is expressed as:

$$|\Psi'\rangle = \sum_{p=1}^{P} \sum_{q=1}^{Q} c_{p,q} \ |\chi_A^p\rangle \otimes |\xi_B^q\rangle \tag{C-2}$$

where the $|\chi_A^p\rangle$ as well as the $|\xi_B^q\rangle$ are orthonormalized kets. Now if we expand the kets $|\varphi_A\rangle$ and $|\varphi_B\rangle$, appearing in the tensor product (C-1), onto the kets $|\chi_A^p\rangle$ and $|\xi_B^q\rangle$ as

$$|\varphi_A\rangle = \sum_{p=1}^{P} x_p \ |\chi_A^p\rangle \qquad \text{and} \qquad |\varphi_B\rangle = \sum_{q=1}^{Q} y_q \ |\xi_B^q\rangle \tag{C-3}$$

we obtain for a ket $|\Psi\rangle$ of the type (C-1):

$$|\Psi\rangle = \sum_{p=1}^{P} \sum_{q=1}^{Q} x_p \ y_q \ |\chi_A^p\rangle \otimes |\xi_B^q\rangle \tag{C-4}$$

It is not obvious at all, just from the knowledge of the coefficients $c_{p,q}$ of $|\Psi'\rangle$, to know if they can be factored into an expression of this type, leading to a product as in (C-1). We present in the next section a systematic method for asserting if this factorization is possible and actually performing it.

C-2. Presence (or absence) of entanglement: Schmidt decomposition

It can be shown (see the demonstration below) that any pure state $|\Psi\rangle$ describing the ensemble of the two physical systems A and B can be written in the form:

$$|\Psi\rangle = \sum_i \sqrt{q_i} \, |u_i, w_i\rangle \tag{C-5}$$

where the $|u_i\rangle$ are a set of orthonormal vectors in the state space of the first system, and the $|w_i\rangle$ another set of orthonormal vectors in the second state space. This expression is the Schmidt decomposition of a pure state, also called the "biorthonormal decomposition".

Whether the state of the total system $|\Psi\rangle$ is entangled or not, it always has a corresponding density operator, written in (A-6). By taking partial traces, each of the subsystems can be described by the density operators:

$$\rho_A = \mathrm{Tr}_B \left\{\rho_{A+B}\right\} \quad ; \quad \rho_B = \mathrm{Tr}_A \left\{\rho_{A+B}\right\} \tag{C-6}$$

Performing these partial traces using (C-5), we get two symmetric expressions:

$$\rho_A = \sum_i q_i \, |u_i\rangle \, \langle u_i| \tag{C-7}$$

and:

$$\rho_B = \sum_i q_i \;\; |w_i\rangle \, \langle w_i| \tag{C-8}$$

This means that, when the total system is in a pure state, the two partial density operators always have the same eigenvalues[2]. In the particular case where they are all zero except one, each of the two subsystems is in a pure state and state $|\Psi\rangle$ can be factored: no entanglement exists in the total system. Most of the time, however, several eigenvalues are non-zero, in which case $(\rho_A)^2$ is obviously not equal to ρ_A, and the same is true for ρ_B; entanglement is then present in the pure state $|\Psi\rangle$.

Demonstration of relation (C-5):

As the two operators ρ_A and ρ_B are Hermitian, non-negative and have a trace equal to unity, their corresponding matrices can be diagonalized to yield real eigenvalues included between 0 and 1. We call $|u_i\rangle$ the normalized eigenvectors of ρ_A (the index i takes on P different values, where P is the dimension of subsystem A) and q_i the corresponding eigenvalues, all positive or zero (but not necessarily different. Similarly, the eigenvectors of ρ_B are noted $|v_l\rangle$ (where l takes on Q different values, Q being the dimension of the second subsystem B), and r_l the corresponding eigenvalues. The two partial density operators can then be expanded as:

$$\rho_A = \sum_{i=1}^{P} q_i \;\; |u_i\rangle \, \langle u_i| \quad \text{and} \quad \rho_B = \sum_{l=1}^{Q} r_l \;\; |v_l\rangle \, \langle v_l| \tag{C-9}$$

with $0 \leq q_i,\, r_l \leq 1$.

State $|\Psi\rangle$ can then be expanded on the basis of the tensor products $\{|A : u_i\rangle \otimes |B : v_l\rangle\}$, that we shall simply note $\{|u_i, v_l\rangle\}$ assuming the first ket represents a state of A and the second a state of B; we call $c_{i,l}$ the components of $|\Psi\rangle$ in this basis and get:

$$|\Psi\rangle = \sum_{i=1}^{P} \sum_{l=1}^{Q} c_{i,l} \, |u_i, v_l\rangle \tag{C-10}$$

[2]Note that this is not necessarily the case if the total system is described by a statistical mixture rather than a pure state. As an example, we can assume that ρ_{A+B} equals a tensor product $\rho_A \otimes \rho_B$, where ρ_A and ρ_B can be chosen arbitrarily, and hence have different eigenvalues.

We now introduce the ket $|\overline{w}_i\rangle$, belonging to the state space $\mathcal{E}_B$ (this ket is not necessarily normalized), as:

$$|\overline{w}_i\rangle = \sum_{l=1}^{Q} c_{i,l} \, |v_l\rangle \tag{C-11}$$

Expansion (C-10) for $|\Psi\rangle$ now simply becomes:

$$|\Psi\rangle = \sum_{i=1}^{P} |u_i, \overline{w}_i\rangle \tag{C-12}$$

We also know that the matrix elements of the partial trace ρ_A are:

$$\langle u_i| \, \rho_A \, |u_j\rangle = \sum_m \langle u_i, v_m \, |\Psi\rangle \, \langle\Psi| \, u_j, v_m\rangle \tag{C-13}$$

Now expression (C-12) for $|\Psi\rangle$ leads to:

$$|\Psi\rangle \, \langle\Psi| = \sum_{i',j'} |u_{i'}, \overline{w}_{i'}\rangle \, \langle u_{j'}, \overline{w}_{j'}| \tag{C-14}$$

Inserting this result into (C-13), we are only left with terms for which $i' = i$ and $j' = j$, which yields:

$$\langle u_i| \, \rho_A \, |u_j\rangle = \sum_m \langle v_m| \, \overline{w}_i\rangle \, \langle \overline{w}_j| \, v_m\rangle = \langle \overline{w}_j| \, \overline{w}_i\rangle \tag{C-15}$$

This means that:

$$\rho_A = \sum_{i,j} |u_i\rangle \, \langle u_i| \, \rho_A \, |u_j\rangle \, \langle u_j| = \sum_{i,j} \, |u_i\rangle \, \langle u_j| \times \langle \overline{w}_j \, |\overline{w}_i\rangle \tag{C-16}$$

Now in the basis $\{|u_i\rangle\}$ we have used, we know that ρ_A is diagonal and given by expression (C-9); the comparison with (C-16) shows that we must necessarily have:

$$\langle \overline{w}_j \, |\overline{w}_i\rangle = \delta_{i,j} \ \times q_i \tag{C-17}$$

For non-zero eigenvalues q_i, this relation shows that one can define a set of orthonormal vectors $|w_i\rangle$ belonging to the state space of system B as:

$$|w_i\rangle = \frac{1}{\sqrt{q_i}} \, |\overline{w}_i\rangle \tag{C-18}$$

For all the values of the index i associated with eigenvalues q_i equal to zero, that same relation shows that the kets $|\overline{w}_i\rangle$ are zero.

Replacing in (C-12) the $|\overline{w}_i\rangle$ by $\sqrt{q_i} \, |w_i\rangle$, we complete the demonstration of equality (C-5), and of relations (C-7) and (C-8) which follow directly.

C-3. Characterization of entanglement: Schmidt rank

The number of non-zero q_i eigenvalues, i.e. the number of non-zero terms in (C-5), is called the "Schmidt rank" of $|\Psi\rangle$ and noted R. When $R = 1$, the state of the total system is not entangled, the two subsystems each being in a pure state. When $R = 2$, we are in the case studied in § B where the two subsystems are entangled, when $R = 3$ we get a more complex entanglement, etc. This gives us a criterion for deciding whether a general ket (C-2) is entangled or not. We just have to compute a partial density operator for one of the subsystems and the number of its non-zero eigenvalues (for the sake of simplicity, we shall obviously choose the subsystem whose state space's dimension is the smallest). If that number equals 1, the eigenvector associated with that non-zero eigenvalue becomes one of the factors of the decomposition, which makes it easy to find the other. If that number is greater than 1, however, the decomposition into a single tensor product is no longer possible.

In a way, entanglement is symmetrically shared between A and B. For example, it is not possible for one of the two subsystems to be in a pure state and the other in a statistical mixture. The rank R must be lower than the smallest of the dimensions P and Q of the state spaces of A and B; to allow a high rank entanglement, the two subsystems must thus have state spaces with high enough dimensions.

Comment:

When all the q_i eigenvalues of ρ_A (and of ρ_B) are distinct, the Schmidt decomposition is unique. This is because decompositions (C-9) and (C-8) of ρ_B on the projectors onto its eigenvectors must then necessarily coincide; the set of eigenvectors $|w_i\rangle$ is identical to that of the $|v_l\rangle$. In this case, the eigenvectors of the partial density operators directly yield the unique Schmidt decomposition. On the other hand, when certain q_i eigenvalues are degenerate, this is no longer true. As an example, we saw that for the singlet state (B-1), the two partial density matrices have two eigenvalues, both equal to 1/2; this singlet state, decomposed in (B-1) into products of eigenvectors of the Oz spin components, can equally well be decomposed into products of eigenvectors of the spin components on any spatial direction. There are an infinity of possible Schmidt decompositions for that state.

D. Ideal measurement and entangled states

Entanglement also plays an essential role in any quantum measurement process, as it generally appears while the measured system S and the measuring apparatus M interact. Furthermore, we shall see that it even propagates further and brings the environment of the measuring apparatus into play.

D-1. Ideal measurement scheme (von Neumann)

Von Neumann's quantum measurement scheme proposes a general framework that allows characterizing the quantum measurement process in terms of entanglement appearing (or disappearing) in the state vector describing the total system $S + M$. The two systems S and M are initially described by a factored state $|\Psi_0\rangle$; however, as they

interact during a certain time, they reach an entangled state $|\Psi'\rangle$. After the measurement, we assume they no longer interact, imagining for example they have moved far away from each other.

In the state space of S, with dimension N_S, the physical quantity measured on S is described by an operator O_A whose normalized eigenvectors are the kets $|a_k\rangle$ with eigenvalues a_k (that we shall assume non-degenerate, to simplify the notation):

$$O_A\,|a_k\rangle = a_k\,|a_k\rangle \tag{D-1}$$

Initially, state $|\varphi_0\rangle$ of S is any linear combination of the $|a_k\rangle$:

$$|\varphi_0\rangle = \sum_{k=1}^{N_S} c_k\,|a_k\rangle \tag{D-2}$$

with complex coefficients c_k, having only the constraint that the sum of their squared moduli be equal to 1 (normalization condition). As for the measuring apparatus M, we assume it is, initially, always in the same normalized quantum state $|\Phi_0\rangle$. The initial state of the total system is then:

$$|\Psi_0\rangle = |\varphi_0\rangle \otimes |\Phi_0\rangle \tag{D-3}$$

D-1-a. Basic process

We start with the particular case where the measurement result is certain and where the system S is initially in one of the eigenstates associated with the measurement: $|\varphi_0\rangle = |a_k\rangle$. In that case:

$$|\Psi_0\rangle = |a_k\rangle \otimes |\Phi_0\rangle \tag{D-4}$$

Once the measurement is done, S stays in the same state $|a_k\rangle$, but the measuring apparatus reaches a state $|\Phi_k\rangle$ different from $|\Phi_0\rangle$ and which depends on k: this is a necessary condition for the result to be experimentally accessible. This is because the position of the "pointer" used for the reading of the result (a needle in a macroscopic apparatus, the recording of the result in a memory, etc.) must necessarily depend on k to allow for the acquisition of the data. It is also natural to assume that the different states $|\Phi_k\rangle$ are orthogonal to each other, since the pointer necessarily involves a large number of atoms whose different states will allow a macroscopic observer to read the result. The measurement process for the total system can be summed up, in this simple case, as follows:

$$|\Psi_0\rangle = |a_k\rangle \otimes |\Phi_0\rangle \implies |\Psi'\rangle = |a_k\rangle \otimes |\Phi'_k\rangle \tag{D-5}$$

where $|\Phi'_k\rangle$ is a normalized state of M. At this stage, no correlation or entanglement has appeared between the measuring apparatus and the measured system. This is what happens in the simple case where the measurement result is certain.

In the general case, the initial state of system S is a superposition (D-2) of eigenstates $|a_k\rangle$. In this case, state (D-4) must be replaced by the linear combination, with the same coefficients:

$$|\Psi_0\rangle = \sum_k c_k\ |a_k\rangle \otimes |\Phi_0\rangle \tag{D-6}$$

As Schrödinger's equation is linear, we get:

$$|\Psi_0\rangle \Longrightarrow |\Psi'\rangle = \sum_k c_k \ |a_k\rangle \otimes |\Phi'_k\rangle \tag{D-7}$$

which now becomes a state in which the measuring apparatus is entangled with the measured system S. The states of S and M are therefore strongly correlated: when the "pointer's position" is associated with one of the state vectors $|\Phi'_k\rangle$, the state of the system S must be described by the ket $|a_k\rangle$ associated with a definite eigenvalue of O_A.

After the measurement, one can no longer attribute a state vector (pure state) to the system S, which can only be described by a partial density operator. As the states $|\Phi'_k\rangle$ are normalized and orthogonal to each other, this density operator is given by:

$$\rho'_S = \mathrm{Tr}_M \left\{ |\Psi'\rangle \langle\Psi'| \right\} = \sum_k |c_k|^2 \ |a_k\rangle \langle a_k| \tag{D-8}$$

This relation was to be expected: it simply states that the system has a probability $|c_k|^2$ of being in the state $|a_k\rangle$ associated with the measurement result a_k, which is in agreement with the Born rule for the usual probabilities. This useful formula sums up in a simple way a number of characteristics of the quantum measurement postulate. Note, however, that at this stage, all the possible results are still present in the partial trace, as they are considered possible, even after the measurement. Nothing at this point tells us that only one result is actually measured when the experiment is performed, nor that the squares of the coefficients $|c_k|$ can be interpreted as classical probabilities associated with mutually exclusive observations. The evolution predicted by Schrödinger's equation cannot, by itself alone, explain the uniqueness of the results observed at the macroscopic level. This is why von Neumann introduced the postulate of the state vector's reduction (also called the "wave packet reduction" or "wave packet collapse", *cf.* Chapter III, § B-3-c); more detail on this point will be given in § D-3.

D-1-b. Dynamics of the entanglement process

A simple interaction Hamiltonian between systems S and M can explain the appearance of entanglement between these systems, and lead to relations (D-5) or (D-7). As an example, imagine this interaction Hamiltonian H_{int} can be written as:

$$H_{int} = g \ O_A \otimes P_M \tag{D-9}$$

where O_A is the operator (acting only on S) already introduced above, P_M an operator acting only on M, and g a coupling constant. We shall also assume that, in the state space of M, there exists a Hermitian conjugate operator X_M of the operator P_M:

$$[X_M, P_M] = i\hbar \tag{D-10}$$

This commutation relation means that P_M generates the translation operators with respect to X_M. In other words, the action of $e^{-i\Delta x P_M/\hbar}$ on any eigenvector $|x_M\rangle$ of X_M:

$$X_M \, |x_M\rangle = x_M \, |x_M\rangle \tag{D-11}$$

leads to a translation by Δx of the eigenvalue x_M:

$$e^{-i\Delta x P_M/\hbar} \, |x_M\rangle = |x_M + \Delta x\rangle \tag{D-12}$$

where Δx is any real number – see relation (13) of Complement $\mathrm{E_{II}}$.

We assume that $|\Phi_0\rangle$ (state of the measuring apparatus before the measurement) is a normalized eigenstate of X_M with eigenvalue x_0, and we ignore[3] any other source for the evolution of the total system other than the interaction between S and M. The evolution operator between the time $t = 0$ before the measurement, and the time $t = \tau$ when the interaction is over, is:

$$U(\tau, 0) = e^{-ig\tau O_A P_M/\hbar} \tag{D-13}$$

Its action on the ket (D-4) yields:

$$U(\tau, 0)\,|a_k\rangle \otimes |\Phi_0(x_0)\rangle = |a_k\rangle \otimes |\Phi_0(x_0 + g\tau a_k)\rangle \tag{D-14}$$

where the variables in parentheses in the states of the measuring apparatus[4] refer to the eigenvalues of X_M. This means that the states $|\Phi'_k\rangle$ introduced in (D-5) are the kets:

$$|\Phi'_k\rangle = |\Phi(x_0 + g\tau a_k)\rangle \tag{D-15}$$

These relations show that, as far as the measuring apparatus M is concerned, the eigenvalue of X has been shifted by a quantity $g\tau a_k$ that depends on the eigenvalue a_k of S. The observable X_M therefore plays the role of a "pointer's position" in the measuring apparatus (measuring needle), which yields the measurement result once the two systems have interacted. As for the observable O_A, it pertains to the system S being measured by the pointer's position.

If now the initial state is in a coherent superposition as in (D-6), the state after the interaction (D-7) is written:

$$|\Psi'\rangle = \sum_k c_k \;\; |a_k\rangle \otimes |\Phi(x_0 + g\tau a_k)\rangle \tag{D-16}$$

which is a biorthonormal decomposition such as the one obtained in § C-2. If, initially, the system S is not in an eigenstate of O_A, the interaction with the measuring apparatus changes its state into a statistical mixture (D-8). On the other hand, if the system is initially in an eigenstate of O_A, it will stay in the same eigenstate after the measurement: the measurement process does not change its state. The measurement is then said to be a "quantum non-demolition" measurement, or QND measurement.

D-2. Coupling with the environment, decoherence; "pointer states"

We now examine under which conditions the interaction and entanglement process we have considered constitutes a good measurement. A first obvious condition to be satisfied is for the states $|\Phi'_n\rangle$ of the measuring apparatus to store the information about

[3] To avoid this hypothesis, the computation could be performed within the interaction point of view (exercise 15 in Complement $\mathrm{L_{III}}$) with respect to free evolutions of both S and M; this would somewhat complicate the results. However, as we focus here on the dynamics induced by their mutual interaction, we shall keep the computations simple and assume that these free evolutions have a negligible effect during the duration τ of the interaction.

[4] Needless to say, a measuring apparatus is macroscopic and has many other degrees of freedom apart from the pointer's position. For the sake of simplicity, these other degrees of freedom have not been introduced in the notation.

the measurement result in a robust way, and prevent it from being destroyed as M continues to evolve on its own. This condition will be fulfilled if X_M is a constant of the motion of M or, in other words, if X_M commutes with the Hamiltonian H_M of M. In addition, the measuring apparatus cannot remain isolated from its environment, even at a microscopic level. In view of its function, the apparatus must be able to interact and be correlated with a measurement recording device, and even with the observer collecting the measurement result; it is, by definition, an "open" apparatus, which can interact with the outside world. In any case, for this apparatus to be completely isolated, would require that none of its atoms, electrons, etc., be in interaction and correlated with any of the environment particles, which is obviously impossible to achieve with a macroscopic apparatus.

This means that, as far as the coupling between M and the environment E is concerned, an entanglement phenomenon occurs, reminiscent of the one discussed for S and M. We must determine which basis, in the state space of M, will lead for the entangled state $M + E$ to a biorthonormal decomposition similar to (D-16). The computation of the partial density operator of M is similar to the computation that yielded (D-8) for the entanglement between S and M: it is in the basis of this biorthonormal decomposition that the partial density operator of M (which plays the role of S) remains diagonal; with another basis, the density matrix will have in general non-diagonal elements. As, furthermore, the entanglement continues to propagate further and further into the environment, it is necessary that the relevant basis of M remains constant in time. It is thus important to find this privileged basis.

Depending on the circumstances, the coupling between a measuring apparatus and its environment can take on various forms, in general complex due to the large number of degrees of freedom involved; several time constants come into play. Different models have been proposed to account for this coupling and the dynamics it produces. Without going into any details we shall make a few general remarks. The measurement process involves a whole chain of amplification between S and the macroscopic pointer, which can be composed of mesoscopic or macroscopic objects sensitive to the environment. Entanglement propagates along that chain via local interactions: the interaction potentials are diagonal in the position representation, and have a microscopic range. Consequently, they cannot couple quantum states corresponding to macroscopically different positions of the objects concerned; the branches of the state vector corresponding to different spatial positions propagate independently. This means that the coupling with the environment tends to favor the basis of states where the positions of the different elements of the measuring apparatus, including in particular its pointer, occupy well defined positions in space. The corresponding preferred basis in the state space of the measuring apparatus M, in which its density matrix remains diagonal over time, is called the basis of the "pointer states". In this basis, and only in this one, defined by pointer localization criteria, the entanglement with E is prone to destroy the coherences (non-diagonal elements of the density matrix), without changing the diagonal elements (meaning the positions of the pointer's particles).

To sum up, several conditions are necessary for a device to be considered as an acceptable measuring apparatus M for a physical quantity of S. In the first place, the coupling between S and M must be capable of transferring the right information from one to the other. The transferred information must then be conserved over time, while M continues its own evolution, and is coupled with the environment E. Obviously, these are

necessary conditions. In practice, an effective measuring apparatus must be conceived taking into account many other imperatives, such as high sensitivity, or strong protection against unavoidable external perturbation.

D-3. Uniqueness of the measurement result

As mentioned above, nothing in the dynamics associated with Schrödinger's equation can explain the uniqueness of the results observed at the macroscopic level. This is not surprising as (D-8) is a direct consequence of Schrödinger's equation, which is incapable of stopping on its own the endless propagation of the "von Neumann chain", as we shall now discuss.

D-3-a. The infinite von Neumann chain

Let us go back to the ideal measurement scheme of § D-1. After the measurement, the state of $S+M$ is the entangled state (D-7), a superposition of components associated with all the possible measurement results. One may wonder if, using a second measuring apparatus M_2 to observe M, one might be able to resolve this superposition and obtain a unique result. In fact, the same entanglement process that occurred between S and M will occur again, leading to a final state:

$$|\Psi''\rangle = \sum_k c_k \ |a_k\rangle \otimes |\Phi'_k\rangle \otimes |\Xi'_k\rangle \tag{D-17}$$

where the kets $|\Xi'_k\rangle$ represent the states of the second measuring apparatus M_2, orthogonal to each other for different values of k. Adding a third measuring apparatus M_3 will, obviously, only continue further the entanglement's progression, each additional apparatus playing the role of an environment for the previous one. This chain of measuring apparatus may continue all the way to infinity without permitting at any stage the resolution of the superposition, and the demonstration of the uniqueness of the measurement result. This is called the von Neumann chain (and the logical problem it poses is called the "von Neumann's infinite regress").

The well-known "Schrödinger's cat paradox" involves a similar situation. The system S is supposed to be a radioactive nucleus in a superposition of two states, $|a_1\rangle$ where the nucleus is still in the excited state, and $|a_2\rangle$ where it has disintegrated, emitting a particle. The kets $|\Phi_k\rangle$, $|\Xi_k\rangle$, etc. represent the states of the measuring apparatus that can detect this particle, and then trigger a mechanical system killing the cat in the case of positive detection. The last of these kets $|Z\rangle$ characterizes the cat, which can therefore be in state $|Z_1\rangle$ where it is still alive, or in state $|Z_2\rangle$ where it is dead. Schrödinger points out the absurdity of a physical description involving a cat that can be at the same time both in an alive and a dead state.

As we just discussed, the uniqueness of the measurement results cannot be proven with Schrödinger's equation; this equation merely predicts that the pointer of a measuring apparatus, and any other macroscopic object, can become superpositions of states located at points very far away in space. Because of the linearity of Schrödinger's equation, nothing prevents the different components of the state vector from propagating further and further away, without this infinite chain of entanglements ever reducing into a single one of its components. It is precisely to solve this problem that von Neumann introduced a specific postulate: the postulate of the reduction of the state vector (Chapter III (§ B-3-c) which "forces" the uniqueness of the measurement result.

D-3-b. Postulate of reduction of the state vector

The postulate of reduction (or collapse) of the state vector is also called the "projection postulate", or the "postulate of collapse of the wave packet". As we saw in Chapter III (§ B-3-c, this postulate states that, once the measurement has been performed, one must suppress the summations appearing in (D-7), (D-8), and (D-17): one only keeps, among all the terms, the component $k = m$ corresponding to the measurement result actually observed. After the measurement, the state vector becomes again a simple product from which entanglement has disappeared; S is once again in a pure state. This means that the entanglement, initially created by the measurement, disappears once the result has been recorded.

This postulate, as efficient as it may be, is somewhat difficult to interpret. Using this postulate amounts to considering that the state vector can evolve under the influence of two different processes: a "normal" continuous evolution, obeying Schrödinger's differential equation, and a sudden discontinuous evolution upon measurement, governed by the von Neumann reduction postulate. Obviously, this duality immediately introduces the question of the limit between these two evolutions: from which time on, exactly, should we consider that the measurement has been performed? In other words, how far does the coherent superposition (D-17) propagate? Which physical processes constitute a measurement, as opposed to those leading to a continuous Schrödinger evolution? These difficult questions were the motivation for introducing other interpretations of quantum mechanics. As an example, there are "non-standard" interpretations where Schrödinger's equation is modified by the adjunction of a small stochastic term. This term is chosen so as to be totally negligible at the microscopic level, while coming into play at a certain macroscopic level; its role is to suppress all the macroscopically different components of the state vector, except one. Both Schrödinger and von Neumann dynamics are then unified into a single equation for the evolution of the state vector. Many other interpretations have been proposed: additional variables, modal interpretation, Everett, all suggesting different solutions for the problem. The interested reader may consult reference [68].

E. "Which path" experiment: can one determine the path followed by the photon in Young's double slit experiment?

Let us now return to a question already discussed in Complement D_I. In Young's double slit experiment where the photon may follow two different paths to reach the detection screen, is it possible to observe interference fringes between these paths and simultaneously obtain information as to which path the photon followed? Figure 1 of Complement D_I, reproduced here in the above Figure 1, shows an interference experimental set-up using a plate pierced with two slits F_1 and F_2; this plate is mobile in a direction perpendicular to the incident photon. As it receives momentum transfers Δp_1 and Δp_2 that will be different, depending on whether the photon goes through F_1 or F_2, one could naively imagine observing interference while knowing through which slit the particle went through. However, using the momentum-position uncertainty relations applied to this mobile plate, we showed that the interference fringes were blurred as soon as the momentum transfers Δp_1 and Δp_2 were sufficiently different to provide this information. The reason is that if we want to be able to distinguish these two momentum transfers, the momentum uncertainty of the mobile plate must be less than the modulus of $\Delta p_1 - \Delta p_2$. A simple calculation then shows that when this condition is met, the uncertainty in the

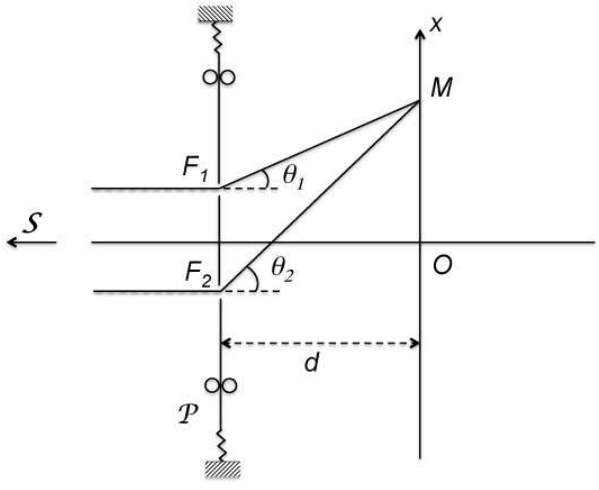


Figure 1: Young's double slit experiment using a plate $\mathcal{P}$, mobile along the x axis, and pierced with two slits F_1 and F_2. A photon, emitted by a source $\mathcal{S}$ assumed to be far away at infinity, reaches the detection screen at point M. The x component of the momentum transferred by the photon to the plate $\mathcal{P}$ depends on whether it goes through slit F_1 or slit F_2.

position of the plate must necessarily be larger than the fringe spacing, which blurs out the fringes. It is impossible to know which of the slits the photon went through without destroying at the same time the interference pattern.

We shall take the analysis a step further and consider the entanglement between the plate and the paths followed by the photon. This should allow us to envisage intermediary situations where partial information on the particle's path can be obtained.

E-1. Entanglement between the photon states and the plate states

Consider the path F_1M followed by the photon if it goes through F_1 and arrives at M on the detection screen (Figure 1). We call $|\psi_1\rangle$ the photon state when it follows that path and transfers a momentum Δp_1 to the mobile plate as it goes through F_1. In that case, after the photon's transit, the state of the plate is:

$$|E_1\rangle = \exp(i\Delta p_1 X/\hbar)|E_0\rangle \tag{E-1}$$

where $|E_0\rangle$ is the initial state of the plate and $\exp(i\Delta p_1 X)$ the momentum space translation operator, by a quantity Δp_1. The state of the global system photon + plate, along the path F_1M, is therefore $|\psi_1\rangle \otimes |E_1\rangle$. A similar reasoning would yield the result $|\psi_2\rangle \otimes |E_2\rangle$ along the path F_2M. As Schrödinger's equation is linear, the state of the global system after the photon has crossed the plate is:

$$|\Psi\rangle = |\psi_1\rangle \otimes |E_1\rangle + |\psi_2\rangle \otimes |E_2\rangle \tag{E-2}$$

which clearly shows the entanglement between the photon[5] and the plate.

E-2. Prediction of measurements performed on the photon

Measurements performed only on the photon after its crossing the plate can be predicted from the reduced density operator ρ, which is the partial trace over the plate variables of the global system density operator $|\Psi\rangle\langle\Psi|$. The matrix elements of ρ are obtained via the standard calculation of a partial trace (Complement $\mathrm{E_{III}}$, § 5-b) and lead to the operator expression:

$$\rho = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| + |\psi_1\rangle\langle\psi_2|\langle E_2|E_1\rangle + |\psi_2\rangle\langle\psi_1|\langle E_1|E_2\rangle \tag{E-3}$$

(in this equation we did not write the factors of $|\psi_1\rangle\langle\psi_1|$ and $|\psi_2\rangle\langle\psi_2|$, which are the scalar products $\langle E_1|E_1\rangle$ and $\langle E_2|E_2\rangle$, both equal to 1 if the state $|E_0\rangle$ is normalized). The interference between the two paths is described by the terms in $|\psi_1\rangle\langle\psi_2|$ and $|\psi_2\rangle\langle\psi_1|$, which are multiplied by the scalar products $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$.

Two extreme cases then appear. If the two states $|E_1\rangle$ and $|E_2\rangle$ are very close to each other, the two scalar products are practically equal to 1 and the interference terms in (E-3) are barely modified by the presence of the factors $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$: the interference is then quite visible on the detection screen. In that case, however, the states $|E_1\rangle$ and $|E_2\rangle$ are too close to give any information as to whether the photon went through F_1 or F_2. In the other extreme case where $|E_1\rangle$ and $|E_2\rangle$ are very different from each other, their scalar product is practically zero: the interference terms disappear from (E-3), but one can, in principle, determine which of the two slits the photon went through. The present calculation allows studying intermediate situations where the scalar products $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$ take on values included between 0 and 1. They describe how the contrast of the fringes diminishes when $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$ continuously decrease from 1 to 0.

Actually, these scalar products can easily be computed from (E-1) and the equivalent relation for $|E_2\rangle$. This leads to:

$$\langle E_2|E_1\rangle = \langle E_0|\exp\left[i(\Delta p_1 - \Delta p_2)x\right]|E_0\rangle \tag{E-4}$$

Using the expression of $\Delta p_1 - \Delta p_2$ (noted $p_1 - p_2$ in Complement $\mathrm{D_I}$) and equations (6) and (7) of that complement, we can show that $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$ are equal to the overlap integrals between the plate initial wave function, and that same wave function translated in momentum space by the amount h/a where a is the fringe spacing.

F. Entanglement, non-locality, Bell's theorem

We now present two important theorems, the EPR (for Einstein, Podolsky and Rosen) theorem, and Bell's theorem, which are related, the second actually being a logical continuation of the first. The EPR theorem was presented in an article published in 1935 by these three authors [69], and is one of the episodes of the famous discussion between Einstein and Bohr concerning the foundations of quantum mechanics (in particular during the Solvay conferences). Einstein's position was that the entire physical world had

[5] All the conclusions of this section remain valid for Young's interference type experiments performed with a massive particle instead of a photon.

to be expressed in the general framework of relativity, where the concept of space-time events is fundamental. Bohr had a different point of view, and considered that quantum theory demanded abandoning a description of microscopic events in space-time terms, while of course conforming to the actual predictions of relativity.

F-1. The EPR argument

The EPR theorem can be stated as follows: "*If all the predictions of quantum mechanics are correct (even for systems made of several remote particles) and if physical reality can be described in a local (or separable) way, then quantum mechanics is necessarily incomplete: some 'elements of reality' exist in Nature that are ignored by this theory* ".

To demonstrate their theorem, Einstein, Podolsky and Rosen imagined an experiment where two physical systems, originating for example from a common source S and described by an entangled quantum state, are then measured in remote regions of space. Historically, EPR developed their argument for correlated particles whose position and momentum are measured. It is however simpler to present an equivalent version of the argument concerning spins and discrete results, a version initially proposed by Bohm (and often called for that reason EPRB).

F-1-a. Exposing the argument

Imagine that two spin 1/2 particles are emitted by a source S in a singlet state (B-1), which is an entangled state where the spins are strongly correlated. The particles then move towards two remote regions of space, without their spins interacting with the outside world; the initial spin entanglement remains unchanged. In these remote regions of space, the particle spin components are measured along a direction defined by angle a for the region on the left, and by angle b for the region on the right (Fig. 2). One often calls Alice and Bob the two observers who perform the measurements in the two different laboratories, which can be very far away from each other. Alice chooses the direction a freely, which defines her "measurement type". With a spin 1/2, she can only obtain two results, that we will note $+1$ or -1, whatever measurement type was chosen. In a similar way, Bob chooses the direction b arbitrarily and obtains one of the two results $+1$ or -1. In the thought EPRB experiment, one assumes for simplicity that the two spins, once they have been emitted by the source, will only interact with the measuring apparatus (without having any free evolution, as was the case above). Standard quantum mechanics then predicts (§ B) that the distances and instants at which the measurements are performed do not play any role in the probability of obtaining the different possible double results.

To keep things simple, let us assume Alice and Bob limit their choices to a finite number of directions a and b for their respective measurements. It may then happen, by chance, that their chosen directions are parallel. Now if the angles a and b are chosen to be equal (parallel measurement directions), relations (B-10) indicate that the results will be always opposite for the two measurements: each time Alice observes ± 1, Bob observes the opposite value ∓ 1. This remains valid even if the measurements are performed at points greatly separated in space, whatever the choice $a = b$, and even if the two observers operate in totally independent ways, in their own regions of space; for example, they could make their choice at the last moment, even after the emission and

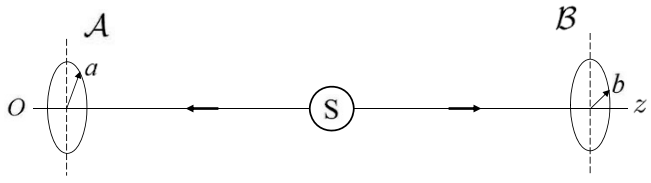


Figure 2: In an EPRB experiment, a source S emits pairs of particles in a singlet spin state (entangled quantum state). These particles then propagate along the Oz axis, towards two remote regions of space $\mathcal{A}$ and $\mathcal{B}$, where Stern-Gerlach apparatus are used by observers Alice and Bob to measure the components of their spin along directions perpendicular to Oz. For the first particle, the measurement direction is defined by angle a, and for the second, by angle b. Each measurement yields a result $+1$ or -1, and one looks for correlations between these results when the experiment is repeated a great number of times.

propagation of the pair of particles.

Let us assume Alice performs her measurement along a direction a before Bob starts doing his own measurement. When Alice finishes her measurement, it becomes certain that, should Bob decide to chose a direction b parallel to a, he will observe the opposite result; the result is certain in that particular case. Such a certainty can only come from the fact that the particle measured by Bob possesses a physical property that determines this certain result; this property (called "element of reality" by EPR) will influence the way that particle interacts with the measuring apparatus in B and will determine the result. On the other hand, the particle propagating towards Bob cannot be influenced by events occurring in Alice's laboratory. This means that this physical property we are discussing existed before the measurement performed by Alice.

The reasoning is obviously symmetrical and establishes that, before any measurement, the particles already possessed physical properties that determined the outcome of the future measurements. As the direction a chosen by Alice was random, it means that the particles must possess enough properties to determine the results for any analysis directions chosen by the observers. Now quantum mechanics does not predict the existence of such properties, as it only gives a description of the particles via a singlet state vector, which always predicts a totally random result for the first measurement. Furthermore, there exists no quantum state for which all the spin components on arbitrary directions can be simultaneously determined (the corresponding operators do not commute with each other). This means that quantum mechanics accounts only partially for the physical properties of the system; it is therefore incomplete.

F-1-b. Assumptions and conclusions

Let us discuss in more detail the logical structure of the EPR argument.

(i) It starts by assuming that the predictions of quantum mechanics for the proba-

bilities of measurement results are correct. The argument thus assumes that the perfect correlations predicted by this theory are always observed, whatever the distance between the two measuring apparatus.

(ii) Another essential ingredient of the EPR argument is the concept of "elements of reality" defined with the following criterion [69]: "if, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity". In other words, a certainty cannot be built on nothing: an experimental result known beforehand can only be the consequence of a preexisting physical quantity.

(iii) And last, but not least, the EPR argument brings in the notion of space-time and locality: the elements of reality they discuss are attached to regions of space where the experiments take place, and they cannot suddenly vary (and certainly not appear) under the influence of events occurring in a very distant region of space. Einstein wrote in 1948 [70]: "Physical objects are thought of as arranged in a space-time continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects are situated in different parts of space". To sum up, one can say that the basic conviction of EPR is that regions of space contain their own elements of reality (attributing distinct elements of reality to separated regions of space is sometimes called "separability"), and that their time evolution is local – one often refers to "local realism" in the literature to qualify the ensemble of the EPR hypotheses.

Basing their argument on these hypotheses, EPR show that, for any chosen values of a and b, the measurement results are functions:

(i) of the individual properties of the spins the particles carry with them (the EPR elements of reality);

(ii) and of the orientations a, b of the Stern and Gerlach analyzers (which is obvious).

It follows that the results are given by well defined functions of these variables, meaning that no non-deterministic process occurs: a particle with spin brings along all the necessary information to yield the result of a future measurement, whatever the choice of the orientation a (for the first particle) or b (for the second). This implies that all the components of each spin must have simultaneously well determined values.

F-2. Bohr's reply, non-separability

Bohr rapidly replied [71] to the EPR article presenting their argument. In Bohr's view, the only physical system to be considered is the entire experimental set-up, including the measured quantum system and all the measuring apparatus, which are treated classically. It is thus meaningless to try and select among this ensemble subsystems having individual physical properties. The physical system Bohr considers is a whole that one should not attempt to separate into parts. This is often called the "non-separability" rule. In other words, Bohr considers that spatial separation does not lead to separability.

It is not the EPR reasoning that Bohr criticizes, but he considers that their starting assumptions are not relevant in the framework of quantum physics. From Bohr's point of view, the EPR criterion for elements of reality "contains an essential ambiguity when applied to quantum phenomena". Along the same line, more than ten years later (in 1948), Bohr made his point of view explicit [72]: "Recapitulating, the impossibility of subdividing the individual quantum effects and of separating a behavior of the objects

from their interactions with the measuring instrument serving to define the conditions under which the phenomena appear implies an ambiguity in assigning conventional attributes to atomic objects, which calls for a reconsideration of our attitude towards the problem of physical explanation ". It is thus the very need for a physical explanation involving such a subdivision that is questioned by Bohr.

Bohr refutes Einstein's basic idea, namely that one can attribute distinct physical properties to two objects located in very remote space-time regions. He believes that "quantum non-separability" applies, even in such a situation. It is understandable that Einstein was unwilling to abandon concepts that are the pillars of special and general relativity (gravitation).

F-3. Bell's inequality

In 1964, more than thirty years after the publishing of the EPR argument, an article by Bell shed an entirely new light on the question [73]. This article, in a way, took up the EPR argument from the point at which its authors had left it. Taking at their face value the existence of the EPR elements of reality, and using the same local realism considerations, Bell showed that there is actually no way to complete quantum mechanics without changing its predictions, at least in some cases. This means that one must either accept that certain predictions of quantum mechanics are sometimes incorrect, or abandon certain EPR hypotheses, however natural they may seem.

F-3-a. Bell's theorem

Following Bell's idea, let us assume that λ represents the "elements of reality" associated with the spins; λ is, actually, just a concise notation that could represent a multiple component vector, so that the number of elements of reality contained in λ is totally arbitrary. One can even include in λ components that play no particular role in the problem; the only important hypothesis is that λ must contain enough information to yield the results of all the possible spin measurements. For each pair of spins emitted in the course of the experiment, λ is fixed.

Another commonly used notation for the two measurement results is A and B, not to be confused with the small letters a and b used for the parameters of the two measuring apparatus. As expected, A and B are functions not only of λ, but also of the measurement parameters a and b. However, locality requires that b has no influence on result A (since the distance between the two measurements' locations is arbitrarily large); conversely, a has no influence on result B. We shall note $A(a,\lambda)$ and $B(b,\lambda)$ the corresponding functions, which can take on two values, $+1$ or -1. Figure 3 schematizes the experiment we are discussing.

To establish Bell's theorem, it is sufficient to take into account only two directions for each individual measurement; we shall then use the simpler notation:

$$A \equiv A(a,\lambda) \qquad A' \equiv A(a',\lambda) \tag{F-1}$$

and:

$$B \equiv B(b,\lambda) \qquad B' \equiv B(b',\lambda) \tag{F-2}$$

For each emitted pair of particles, as λ is fixed, the four above numbers have well defined values, which can only be ± 1.

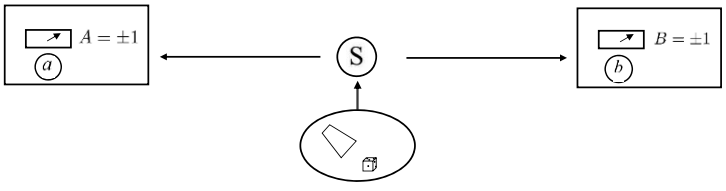


Figure 3: Source S emits particles toward two measuring apparatus located far away from each other, each being set up with its own measurement parameter, respectively a and b; each apparatus yields a ±1 result. The oval under the source symbolizes a fluctuating random process, which controls the particles' emission process, and hence their properties. Correlations between the measured results are observed; they are due to the common random properties the particles have acquired upon their emission by the fluctuating process.

Consider then the sum of products:

$$M(\lambda) = AB - AB' + A'B + A'B' \tag{F-3}$$

that can also be written as:

$$M(\lambda) = A(B - B') + A'(B + B') \tag{F-4}$$

If $B = B'$, the above expression reduces to $2A'B = \pm 2$; if $B = -B'$, it reduces to $2AB = \pm 2$. In both cases, we see that $M = \pm 2$.

If we now take the average value $\overline{M}$ of $M(\lambda)$ over a large number of emitted pairs (average over λ), we get:

$$\overline{M} = \overline{AB}_\lambda - \overline{AB'}_\lambda + \overline{A'B}_\lambda + \overline{A'B'}_\lambda \tag{F-5}$$

where $\overline{AB}_\lambda$ denotes the average value over λ of the product $AB \equiv A(a,\lambda)B(b,\lambda)$, and a similar notation has been used for the 3 other terms. As each $M(\lambda)$ value can only be ± 2, we necessarily have:

$$-2 \leq \overline{M} \leq +2 \tag{F-6}$$

This result is the so called BCHSH (Bell, Clauser, Horne, Shimony et Holt) form of Bell's theorem. This inequality must be satisfied by all sorts of measurements yielding random results, whatever mechanism creates the correlations, as long as the locality condition is obeyed: A does not depend on the measurement parameter b, and B does not depend on a.

This means that any theory that fits in the framework of "local realism" must lead to predictions satisfying relation (F-6). Realism is necessary since we used in the demonstration the concept of EPR elements of reality to introduce the functions A and

B; locality is also essential as it forbids A to depend on b and, conversely, B to depend on a.

The simplicity of this demonstration is such that the inequality may be expected to remain valid in many situations. This actually happens any time the observed correlations can be explained by fluctuations having a past common cause; they are then referred to as "classical correlations". In such cases, each time the experiment is performed, this common cause is fixed, and the 4 numbers A, B, A' and B' take on well-defined values (even though they are, a priori, unknown) all equal to ± 1; the number M is therefore also well defined, equal to -2 or $+2$. Whatever the M values found in a random series of N measurements, it is mathematically impossible for the sum of these M values to be greater than $2N$ or smaller than $-2N$. Consequently, the average value obtained by dividing this sum by N necessarily obeys (F-6): the mere existence of these 4 numbers is sufficient to obtain the inequality.

F-3-b. Contradictions

It would seem natural for any reasonable physical theory to automatically lead to predictions satisfying this inequality. Now, surprisingly enough, this is not the case for quantum mechanics and, furthermore, this contradiction has been experimentally confirmed.

α. *Contradictions with quantum mechanics predictions*

Relations (B-10) allow computing the average value of the product of the ± 1 results obtained in the measurements of the two spins along directions making an angle θ_{ab} with each other. This average value is given by (we write $\hat{A}$ and $\hat{B}$ to emphasize that these letters now denote operators, not numbers):

$$\left\langle \hat{A}(a)\hat{B}(b) \right\rangle = \mathcal{P}_{+,+} \; + \; \mathcal{P}_{-,-} \; - \; \mathcal{P}_{+,-} \; - \; \mathcal{P}_{-,+} = -\cos\theta_{ab} \tag{F-7}$$

This expression is the quantum equivalent of the average value over the λ variable of the product of results $A(a,\lambda)B(b,\lambda)$ in a theory with local realism. To get the quantum equivalent $\langle Q \rangle$ of the combination of the four products of results as they appear in (F-3), we must compute the same combination of average values of these products of results, which yields:

$$\begin{aligned} \langle Q \rangle &= \left\langle \hat{A}(a)\hat{B}(b) \right\rangle - \left\langle \hat{A}(a)\hat{B}(b') \right\rangle + \left\langle \hat{A}(a')\hat{B}(b) \right\rangle + \left\langle \hat{A}(a')\hat{B}(b') \right\rangle \\ &= -\cos\theta_{ab} + \cos\theta_{ab'} - \cos\theta_{a'b} - \cos\theta_{a'b'} \end{aligned} \tag{F-8}$$

Imagine now that the four directions are in the same plane, and that the vectors, arranged in the order $\mathbf{a}$, $\mathbf{b}$, $\mathbf{a}'$ and $\mathbf{b}'$, all make a 45° angle with the preceding vector (Fig. 4); all the cosines are then equal to $1/\sqrt{2}$, except for $\cos\theta_{ab'}$ that is equal to $-1/\sqrt{2}$. We then get $\langle Q \rangle = -2\sqrt{2}$; exchanging the directions of $\mathbf{b}$ and $\mathbf{b}'$, we get $\langle Q \rangle = 2\sqrt{2}$. In both cases, BCHSH inequality (F-6) is violated by a factor $\sqrt{2}$, i.e. by more than 40 %. In spite of the seemingly simple cosine variation of expression (F-7), we just showed that no theory with local reality is able to account for it, as this would violate inequality (F-6). This means that the EPR-Bell argument leads to an important quantitative contradiction with quantum mechanics, proving it to be a theory that does not comply with local realism in the sense of EPR.

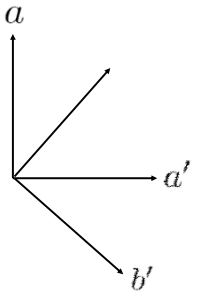


Figure 4: Position of the four vecteurs a, b, a' and b' leading to a maximum violation of BCHSH inequalities for two spin-1/2 particles in a singlet state. These vectors define the spin components to be measured, along a or a' for the spin on the left, and along b or b' for the spin on the right. This means that the entire experiment needs four different set-ups. The only pair of vectors leading to a negative correlation between the two measurement results is (a, b'), as the angle between the corresponding directions is larger than 90°.

How is it possible to encounter such a contradiction and how come such an apparently flawless argument does not apply to quantum mechanics? Several answers can be given:

(i) Had Bohr been aware of Bell's theorem, he would very likely have rejected the existence of the 4 preexisting numbers A, A', B, B'. If these numbers do not exist, the argument of § F-3-a is no longer possible and the BCHSH inequality disappears. Bohr would have considered Bell's theorem as mathematically correct in probability theory, but totally irrelevant in quantum mechanics, as being improper for the quantum description of the experiment under study.

Even if he had accepted to reason about these numbers, as unknown variables to be determined later as is often the case in algebra, would the inequality have survived? The answer is no, still reasoning with Bohr's logic. As already mentioned in § F-2, Bohr's point of view is that the entire experiment must be considered as a whole. One cannot distinguish two separate measurements that would be performed, each on one of the particles: the only true measurement process concerns the ensemble of both particles together. A fundamentally indeterminist and delocalized process occurs in the whole region of space containing the entire experiment.

The functions A and B then both depend on both the measurement parameters, and must be written as $A(a, b)$ and $B(a, b)$; this immediately forces an abandon of locality. Instead of the 2 numbers A and A', we now have 4 numbers, $A = A(a, b)$, $A' = A(a', b)$, as well as $A'' = A(a, b')$ and $A''' = A(a', b')$; the same is true for B and B', which must be replaced by 4 numbers. We now must deal with a total of 8 numbers instead of 4. The demonstration of the BCHSH inequality is then no longer possible and the contradiction

disappears.

(ii) One may prefer a more local point of view for the measurement process, which allows keeping the concept of a measurement on a single particle. To avoid the contradiction with quantum mechanical predictions, one must then consider that it is meaningless to attribute four well defined values A, $A^{'}$, B, $B^{'}$ to each pair. Since only a maximum of two of them can be measured in a given experiment, we should not be able to talk about these four numbers or argue about them even as unknown quantities. A well know phrase of Péres [75] very clearly sums up this point of view: "unperformed experiments have no results". Wheeler [76] expresses the same idea as he writes: "no phenomenon is a real phenomenon until it is an observed phenomenon.".

β. Contradictions with experimental results

The question was: does Bell's theorem allows pointing out very particular situations where quantum mechanics is no longer valid? Or, on the contrary, are the predictions of quantum mechanics always valid, which immediately entails that certain hypotheses leading to the inequalities must be abandoned? A great number of experiments have been performed from 1972 on; they all confirmed the predictions of quantum mechanics, measuring, sometimes with great precision [77], the violation of Bell's inequalities.

After a moment of doubt, it now seems well established that quantum mechanics yields perfectly correct predictions, even in situations where it implies a violation of Bell's inequalities. However plausible they might look, one must abandon at least one of the hypotheses that led to these inequalities.

Conclusion

The concept of quantum entanglement is quite essential; it leads to situations where certain types of correlations, totally impossible in classical physics, can be produced and observed. These situations can occur even when the observations are performed in regions of space arbitrarily remote from one another. A fundamental idea of quantum mechanics, without any classical counterpart, is that the most precise description of a whole does not necessarily entail an equivalently precise description of its parts. This means that there exists no theory both local and realistic, for describing a system containing two remote and entangled particles (it would contradict quantum mechanics).

Entanglement also plays an essential role in the measurement processes, and comes into play at different levels: entanglement between the measured system S and the measuring apparatus M, between M and the environment E, between two environments E and E', and so forth. We also discussed how entanglement determines the contrast of the fringes observed in an interference experiment where a particle has to cross a plate pierced with two holes.

In addition to these important aspects, entanglement also plays an essential role in quantum computing: one seeks to take advantage of the parallel evolution of the various entangled branches of the state vector to perform computations. This domain of research has undergone intense development in recent years, but is too extensive to be treated in the present volume. The reader may want to consult specialized books on the subject, as for example that of D.Mermin [78]. Entanglement also plays a central role in quantum cryptography, whose aim is to fabricate devices for secure quantum key distribution that

cannot be spied on, as any eavesdropping is detectable; a review on this subject can be found in the article by N.Gisin, G.Ribordy, W.Tittel and H.Zbinden [79].

There still remains the fact that, in the presence of entanglement, and in particular during a measurement process, the standard interpretation of quantum mechanics may present some difficulties. Schrödinger's evolution equation does not predict the uniqueness of the measurement result observed in the macroscopic world. To obtain this uniqueness in the framework of the theory, one can introduce an ad hoc postulate, such as the von Neumann postulate of reduction of the state vector. It then raises the question of where to set the border: when exactly should one stop using the continuous evolution of Schrödinger's equation and impose the reduction of the state vector? How can one reconcile the intrinsic irreversibility of this ad hoc postulate with the reversibility of Schrödinger's equation?

Another open question concerns the status of the state vector. We have used it throughout this book as a mathematical tool, good for computing probabilities, but what does it really represent? Does it directly describe physical reality? Or does it simply give information about physical reality? A number of quantum mechanics interpretations have been proposed (see reference [68]) that discuss this fundamental difficulty.

A_{XXI} : DENSITY OPERATOR AND CORRELATIONS; SEPARABILITY	A_{XXI} : This complement introduces von Neumann statistical entropy associated with a density operator, discussing its properties and establishing some important inequalities it must satisfy. Also discussed are the differences between classical and quantum correlations (arising from quantum entanglement effects). The concept of "quantum non-separability" is introduced.
B_{XXI} : GHZ STATES; ENTANGLEMENT SWAPPING	GHZ states provide an example of conflict between quantum mechanics and the usual concept of local realism. The contradiction is even stronger than for Bell's inequalities, as it is expressed as an opposition in signs. Entanglement swapping allows entangling particles without them ever having to interact with each other.
C_{XXI} : MEASUREMENT INDUCED RELATIVE PHASE BETWEEN TWO CONDENSATES	When two Bose-Einstein condensates overlap, their relative phase is a priori totally undetermined. However, such a phase may appear, induced by a measurement process sensitive to that phase. As measurements proceed, this relative phase will progressively acquire a more precise value.
D_{XXI} : EMERGENCE OF A RELATIVE PHASE WITH SPIN CONDENSATES; MACROSCOPIC NONLOCALITY AND THE EPR ARGUMENT	This complement is an extension of the previous one, studying the case where the two condensates are formed of particles with spins. The same phenomenon occurs: the emergence of a relative phase, but in a context where the EPR argument is harder to refute because of the macroscopic character of the measured quantities. Furthermore, situations may arise where Bell's inequalities are violated, which proves that the measurement induced phase between the two condensates is of a non-classical nature.

Complement A_{XXI}

Density operator and correlations; separability

1 Von Neumann statistical entropy 2217
1-a General definition 2217
1-b Physical system composed of two subsystems 2219
2 Differences between classical and quantum correlations . . 2221
2-a Two levels of correlations 2221
2-b Quantum monogamy 2221
3 Separability 2223
3-a Separable density operator 2223
3-b Two spins in a singlet state 2225

In Chapter XXI, we mainly considered global systems $A + B$ described by a state vector (pure state). This complement will examine what happens when these global systems are described by a density operator (statistical mixture); we shall study the correlations quantum mechanics predicts in that case between the two subsystems A and B. We start by introducing, in § 1, the concept of statistical entropy, which yields a useful measure of their degree of correlation. We then analyze, in § 2, the differences between classical correlations (introduced at the probability level) and the quantum correlations (which can arise from the coherent superposition of state vectors). Finally, in § 3, we will come back to the important concept of separability already introduced in § F-2 of Chapter XXI.

1. Von Neumann statistical entropy

The statistical entropy introduced by von Neumann permits, in a straightforward way, to distinguish between a pure state and a statistical mixture; in the latter case, it also yields a measurement of the statistical character of the information known about the physical system. It is also a useful tool for studying in a quantitative way the amount of correlation between two physical systems.

1-a. General definition

With any density operator ρ, we associate a statistical entropy S by the relation:

$$S = -k_B \mathrm{Tr}\left\{\rho \ln \rho\right\} \tag{1}$$

where k_B is the Boltzmann constant. As ρ is Hermitian, this operator can be diagonalized. Noting n_k its eigenvalues, we get:

$$S = -k_B \sum_k n_k \ln n_k \tag{2}$$

Since all the n_k are included between 0 and 1, we necessarily have:

$$S \geq 0 \tag{3}$$

where the equality occurs only if ρ has one eigenvalue equal to 1, all the others being equal to zero. The entropy associated with ρ is therefore equal to zero only if this operator is a projector, and hence corresponds to a pure state. On the other hand, whenever ρ describes a statistical mixture, S is different from zero. It takes on its maximum value when the density operator ρ has equal populations in all the system's accessible states, i.e. if it is proportional to the identity operator in the state space. To prove this, let us vary each n_k by an amount $\mathrm{d}n_k$, and impose a zero variation for the sum over k of (2), while maintaining constant the sum of all n_k using a Lagrangian multiplier λ. We then get:

$$\mathrm{d}S - \lambda k_B \sum_k \mathrm{d}n_k = -k_B \sum_k \left[1 + \mathrm{Log}\ n_k + \lambda\right] \mathrm{d}n_k = 0 \tag{4}$$

For this expression to be zero for any $\mathrm{d}n_k$ means that all the $\ln\ n_k$, and hence all the n_k themselves, must be equal.

One can associate a concept of information, or rather a lack of information, with the entropy S. When the physical system is in a pure state, that state provides the maximum information on the system, compatible with quantum mechanics. In this situation, there is no lack of information and $S = 0$. On the other hand, when the system is spread over several pure states with comparable probabilities, a large value of S means that a lot of information about the system is lacking.

Comment:

The statistical entropy S characterizes the populations of the density matrix n_k (§ 4-c of Complement $\mathrm{E_{III}}$), but not the corresponding eigenvectors. Moreover, the same density operator can in general be obtained from several different statistical mixtures of pure states (*cf.* comment at the end of §4-a of Complement $\mathrm{E_{III}}$); the value of the entropy does not distinguish between these different mixtures.

A statistical mixture of several density operators can only increase the entropy of the system. Imagine, for example, that the density operator ρ is actually the combination of several density operators ρ_n with probabilities p_n (all positive, and whose sum over n is equal to 1), written as:

$$\rho = \sum_{\mathbf{n}} p_n\ \rho_n \tag{5}$$

Noting S_n the entropies associated with ρ_n:

$$S_n = -k_B \mathrm{Tr}\left\{\rho_n\ \ln \rho_n\right\} \tag{6}$$

we can write[1]:

$$S \geq \sum_n p_n\ S_n \tag{7}$$

[1]This properties is often called "entropy concavity".

Demonstration:

In § 1-b of Complement G_{XV}, we showed that, when ρ and ρ' are two density operators with traces equal to 1, one always has the following inequality:

$$\mathrm{Tr}\,\{\rho\,\ln\rho\} \geq \mathrm{Tr}\left\{\rho\,\ln\rho'\right\} \tag{8}$$

(the equality occurring only if $\rho' = \rho$). We can then write:

$$\begin{aligned} -\frac{S}{k_B} = \mathrm{Tr}\,\{\rho\,\ln\rho\} &= \sum_n p_n\ \mathrm{Tr}\,\{\rho_n\,\ln\rho\} \\ &\leq \sum_n p_n\ \mathrm{Tr}\,\{\rho_n\,\ln\rho_n\} = -\frac{1}{k_B}\sum_n p_n\ S_n \end{aligned} \tag{9}$$

which establishes relation (7).

1-b. Physical system composed of two subsystems

We now compare the entropies S_{A+B}, S_A and S_B associated, respectively, with the total density operator ρ_{A+B} and with the partial density operators ρ_A and ρ_B. We are going to show that $S_A + S_B = S_{A+B}$ when the systems A and B are not entangled, and that $S_A + S_B \geq S_{A+B}$ otherwise.

α. *Pure state*

Imagine first that the total system is in a pure entangled state. We have seen that, in that case, the two subsystems A and B are not described by pure but by statistical mixtures of states, so that:

$$\begin{aligned} S_A &= -k_B \mathrm{Tr}\,\{\rho_A\,\ln\rho_A\} \geq 0 \\ S_B &= -k_B \mathrm{Tr}\,\{\rho_B\,\ln\rho_B\} \geq 0 \end{aligned} \tag{10}$$

As the entropy S_{A+B} associated with a pure state is zero, it follows that:

$$S_A + S_B \geq S_{A+B} \tag{11}$$

(the equality corresponds to the special case where the pure state $|\Psi\rangle$ is a product, without entanglement, and where the Schmidt rank is equal to 1; see Chapter refch21, § C-3).

We can also use the Schmidt decomposition for $|\Psi\rangle$, which yields relations (C-7) and (C-8) of Chapter XXI, to get:

$$S_A = -k_B \sum_i q_i\,\ln\,q_i = S_B \tag{12}$$

Both entropies of the two subsystems are thus always equal whenever the total system is in a pure state.

β. Statistical mixture

When the total system is described by a density operator ρ_{A+B} not necessarily corresponding to a pure state, its entropy S_{A+B} may not be equal to zero. We are going to show, however, that this entropy S_{A+B} always remains lower or equal to the sum of the entropies of each subsystem, meaning that relation (11) remains valid in this more general case; this property is referred to as the "entropy subadditivity". The equality in (11) is obtained solely in the case where ρ_{A+B} is a product:

$$\rho_{A+B} = \rho_A \otimes \rho_B \tag{13}$$

which corresponds to the case of two subsystems, separately described by statistical mixtures, while remaining uncorrelated. The difference $S_A + S_B - S_{A+B}$ yields an estimate of the loss of precision between the quantum description of the total system, and the separate quantum descriptions of the two subsystems.

Demonstration:

According to inequality (8), we can write:

$$\mathrm{Tr}\left\{\rho \ln\rho\right\} \geq \mathrm{Tr}\left\{\rho \ln\left(\rho_A \otimes \rho_B\right)\right\} \tag{14}$$

We note $|u_i\rangle$ the eigenvectors of ρ_A with eigenvalues q_i, and $|v_l\rangle$ the eigenvectors of ρ_B with eigenvalues s_l. Let us now compute the trace of the right-hand side of (14) in the basis $\{|u_i, v_l\rangle\}$ of the tensor products of the eigenvectors of the two operators, with respective eigenvalues q_i and s_l; we get:

$$\begin{aligned}
\mathrm{Tr}\Big\{\rho\ \mathrm{Log}\left(\rho_A \otimes \rho_B\right)\Big\} &= \sum_{i,l} \langle u_i, v_l|\, \rho\ \mathrm{Log}\left(\rho_A \otimes \rho_B\right) |u_i, v_l\rangle \\
&= \sum_{i,l} \langle u_i, v_l|\, \rho\ \mathrm{Log}\left(q_i \times s_l\right) |u_i, v_l\rangle \\
&= \sum_{i,l} \langle u_i, v_l|\, \rho\, |u_i, v_l\rangle \mathrm{Log}\left(q_i\right) + \sum_{i,l} \langle u_i, v_l|\, \rho\, |u_i, v_l\rangle \mathrm{Log}\left(s_l\right)
\end{aligned} \tag{15}$$

Let us now choose $\rho = \rho_{A+B}$. The first term on the right-hand side can be written as:

$$\begin{aligned}
\sum_{i,l} \langle u_i, v_l|\, \rho_{A+B}\ \ |u_i, v_l\rangle \ln\left(q_i\right) &= \sum_i \langle u_i|\, \rho_A\ \ |u_i\rangle \ln\left(q_i\right) \\
&= \sum_i \langle u_i|\, \rho_A\ \ln\rho_A\, |u_i\rangle \\
&= \mathrm{Tr}\left\{\rho_A\ \ln\rho_A\right\}
\end{aligned} \tag{16}$$

The second term on the right-hand side of (15) yields a similar expression, where B replaces A. Finally, inequality (14) can be written as:

$$\mathrm{Tr}\left\{\rho_{A+B}\ \ln\rho_{A+B}\right\} \geq \mathrm{Tr}\left\{\rho_A\ \ln\rho_A\right\} + \mathrm{Tr}\left\{\rho_B\ \ln\rho_B\right\} \tag{17}$$

and leads to (11). The equality occurs if and only if (14) becomes an equality, i.e. if ρ_{A+B} is equal to the product (13).

2. Differences between classical and quantum correlations

Quantum mechanics offers more possibilities than classical physics for describing correlations between physical systems. We now briefly discuss such examples.

2-a. Two levels of correlations

The concept of correlation is not, intrinsically, a quantum notion, and it is well known in classical physics. It is then based on probabilistic calculations, which results in the linear weighting of a certain number of possibilities. In this classical context, one introduces a distribution yielding the probability for the first system to occupy a certain state, and the second system, another state; the two systems are correlated when this distribution is not a simple product. If, on the other hand, the distribution turns out to be a product, the two systems are not correlated; a measurement on one of the systems does not change the information about the other. In particular, this is what happens if the states of the two systems, and consequently that of the total system, are perfectly well defined, a case where the notion of correlation between the two systems becomes irrelevant. This means that the notion of correlation between two classical systems is closely linked to an imperfect definition of the state of the total system.

In quantum mechanics, things are totally different. To begin with, even if a physical system is perfectly well defined by a state vector, many of its physical properties are not so precisely defined: during several realizations of the experiment, their measurement can provide fluctuating results. These results can nevertheless be correlated: as an example, we saw in § B of Chapter XXI that the components of each of the two spin-1/2 particles are completely indeterminate but perfectly correlated. Such correlations appear directly at the level of the state vector itself, which can be written as a linear superposition of states where the spins have various orientations. The correlations are therefore related to the the quantum mechanical superposition principle; this is totally different from the combinations of probabilities, which are quadratic functions of that state vector. Letting correlations appear directly at the probability amplitude level, one has access to a level that is, in a way, "a step ahead" of the linear weighting of classical probabilities, and maintains the possibility of quantum interference effects. Note, however, that the existence of this level of combinations does not exclude classical probabilities from coming into play. One can also assume, in quantum mechanics, that the state of the total system is only known in a probabilistic way, so that the two probability levels may coexist. To sum up, it is clear that the concept of quantum correlations covers many more possibilities than correlations in classical physics[2].

2-b. Quantum monogamy

Another purely quantum property is that, if a physical system A is strongly entangled with a physical system B, it cannot be strongly entangled with another system C. Such a property does not have any equivalent in classical physics, where, obviously, nothing prevents correlating a third system C with two others A and B, all the while keeping

[2]We shall introduce, in § 3, a criterion (negativity of the coefficients of the total density operator expansion into a sum of products) for confirming the quantum nature of the correlations between two subsystems.

their initial correlation. This quantum property is often referred to as "entanglement monogamy".

Let us assume, for example, that two spins are in a state of the same type as the singlet state (B-1) of Chapter XXI:

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-\rangle + e^{i\xi}\,|A:-;B:+\rangle\right] \tag{18}$$

(the singlet state is obtained for $\xi = \pi$). How can we add an additional spin without destroying the correlation between the first two? One could imagine the three spin state to be written as:

$$|\Psi_{ABC}\rangle = |\Psi_{AB}\rangle \otimes |C:\theta\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-;C:\theta\rangle + e^{i\xi}\,|A:-;B:+;C:\theta\rangle\right] \tag{19}$$

where $|\theta\rangle$ is any normalized state for the third spin. This ket obviously conserves the same entanglement between spins A and B as in state (18), but the third spin is then totally uncorrelated with the first two.

Another possibility is to choose as a state vector:

$$|\Psi_{ABC}\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-;C:\theta_1\rangle + e^{i\xi}\,|A:-;B:+;C:\theta_2\rangle\right] \tag{20}$$

The density operator ρ_{AB} describing spins A and B is obtained by taking the partial trace (Complement E_{III}, § 5-b):

$$\rho_{AB} = \mathrm{Tr}_C\left\{|\Psi_{ABC}\rangle\langle\Psi_{ABC}|\right\} \tag{21}$$

Computing the matrix elements of this partial trace shows that:

$$\begin{aligned}\rho_{AB} = \frac{1}{2}\Big[& \langle\theta_1|\theta_1\rangle \;\; |A:+;B:-\rangle\langle A:+;B:-| \\ & + \;\; \langle\theta_2|\theta_2\rangle \;\; |A:-;B:+\rangle\langle A:-;B:+| \\ & + e^{i\xi}\langle\theta_1|\theta_2\rangle \;\; |A:-;B:+\rangle\langle A:+;B:-| \\ & + \;\; e^{-i\xi}\langle\theta_2|\theta_1\rangle \;\; |A:+;B:-\rangle\langle A:-;B:+|\Big]\end{aligned} \tag{22}$$

One can then distinguish several cases:

– If $|\theta_1\rangle = |\theta_2\rangle$, we find again (19), and the third spin is not entangled with the first two. The density operator ρ_{AB} is then written:

$$\begin{aligned}\rho_{AB} = {} & |A:+;B:-\rangle\langle A:+;B:-| \;\; + \;\; |A:-;B:+\rangle\langle A:-;B:+| \\ & + e^{i\xi}\,|A:-;B:+\rangle\langle A:+;B:-| \;\; + \;\; e^{-i\xi}\,|A:+;B:-\rangle\langle A:-;B:+|\end{aligned} \tag{23}$$

which is simply the projector onto state (18); it conserves all the entanglement of spins A and B.

– The opposite case is when $|\theta_1\rangle$ and $|\theta_2\rangle$ are orthogonal, so that $|\Psi_{ABC}\rangle$ becomes a so called GHZ state (Greenberger, Horne and Zeilinger; *cf.* Complement B_{XXI}:

$$|\Psi_{GHZ}\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-;C:+\rangle + e^{i\xi}\,|A:-;B:+;C:-\rangle\right] \tag{24}$$

where, when one goes from the first component to the second, the three spins switch from one state to the orthogonal one. The second line in (22) then cancels out and the partial trace ρ_{AB} becomes:

$$\rho_{AB} = \frac{1}{2}\left[\,|A:+;B:-\rangle\,\langle A:+;B:-| + |A:-;B:+\rangle\,\langle A:-;B:+|\,\right] \tag{25}$$

which is a statistical mixture of two possibilities, with probabilities 1/2: the two spins are either in the state $|A:+;B:-\rangle$, or in the state $|A:-;B:+\rangle$. The quantum coherences between these two states (terms dependent on the phase ξ) have totally disappeared. The correlation between the two spins A and B is then of a classical nature[3], and no entanglement comes into play.

– In the intermediate situation where $|\theta_1\rangle$ and $|\theta_2\rangle$ are neither parallel nor orthogonal, we see from (22) that a certain coherence remains (non-diagonal elements). The more parallel $|\theta_1\rangle$ and $|\theta_2\rangle$ are, the more the partial density operator resembles that of the two initial spins which remain entangled, whereas the third one becomes less and less entangled with the first two; conversely, the more orthogonal they are, the more the initial spins lose their correlation, which becomes entirely transmitted to the three spin level.

This is actually a general property: when two physical systems are maximally entangled, a principle of mutual exclusion makes it impossible to entangle them with a third system. Mathematically, this property is expressed by the Coffman-Kundu-Wooters inequality [80].

3. Separability

From Bohr's point of view (§ F-2 of Chapter XXI), one must give up the notion of separability. Even when two physical subsystems are well separated in space, it does not infer that they have, each of them separately, their own physical properties (as EPR assumed); only the total system, including the measuring apparatus, can have such properties. On the other hand, we saw in § 2 that in quantum mechanics, there are two ways for introducing correlations between two systems: a classical way (by assuming they have given probabilities to be found in such or such individual correlated states), and a quantum way (by assuming entanglement directly at the level of a common state vector). We also know that, even though there are situations where quantum mechanics predicts violations of Bell' inequalities and hence of local realism, there are many others where its predictions obey these inequalities; a violation is, in a way, the signature of an ultra-quantum situation. It is thus interesting to look for a criterion allowing a distinction between these two types of correlations.

3-a. Separable density operator

Consider a total system described by a density operator ρ_{A+B} and composed of two subsystems A and B. Let us assume that ρ_{A+B} can be expanded as a series of density operators ρ_A^n and ρ_B^n pertaining to each of the two subsystems, with real and

[3]The density operator is separable in a sense that will be defined in § 3, and therefore cannot lead to violations of Bell's inequalities.

positive π_n coefficients whose sum equals to one, and can be assimilated to probabilities:

$$\rho_{A+B} = \sum_n \pi_n \ \rho_A^n \otimes \rho_B^n \tag{26}$$

with:

$$0 \leq \pi_n \leq 1 \qquad \text{and} \qquad \sum_n \pi_n = 1 \tag{27}$$

Intuitively, one can guess that the correlations contained in ρ_{A+B} must then be of a classical nature. The total system is, with a probability π_n, described by a density operator that is a product, without correlations, of density operators each describing one of the subsystems. The correlations between these subsystems are therefore introduced in a classical way, even if nothing prevents each subsystem from exhibiting strongly quantum individual properties.

Any density operator that can be decomposed as in (26) with positive π_n coefficients is, by definition, said to be "separable" [81, 82]. On the other hand, if any decomposition of ρ_{A+B} such as (26) necessarily includes π_n coefficients that are not real and positive, the density operator ρ_{A+B} contains quantum entanglement and is said to be non-separable. When the total system $A + B$ is separable, correlation measurements between physical properties of the two subsystems A and B can never lead to violations of Bell's inequalities. These violations are thus a sure sign of the non-separability of the density operator.

Demonstration

To show this, let us assume we perform two simultaneous measurements on the systems A and B, the first one depending on the measurement parameter a, and the second, on the measurement parameter b. We note $P_a(R_A)$ the projector acting in the state space of A and corresponding to the measurement result R_A (this projector is the sum of the projectors onto the eigenvectors associated with that measurement). In a similar way, we note $P_b(R_B)$ the projector in the state space of B corresponding to the measurement result R_B. When the total system is described by the density operator (26), the joint probability of obtaining both results R_A and R_B is:

$$\mathcal{P}(R_A, R_B) = \sum_n \pi_n \ \mathcal{P}_a^n(R_A) \times \mathcal{P}_b^n(R_B) \tag{28}$$

with:

$$\begin{aligned} \mathcal{P}_a^n(R_A) &= \mathrm{Tr}_A\{\rho_A^n P_a(R_A)\} \\ \mathcal{P}_b^n(R_B) &= \mathrm{Tr}_B\{\rho_B^n P_b(R_B)\} \end{aligned} \tag{29}$$

As all the numbers appearing on the right-hand side of (28) are positive, this equality has a natural interpretation in classical physics, which is the framework of our present argument. We are dealing with two levels of probabilities. At one level, the total system is prepared, with probability π_n, in a state where the two subsystems are uncorrelated. At a second level, for each value of n, the individual states of the subsystems are only known in a statistical way via the probability $\mathcal{P}_a^n(R_A)$ of a result R_A, and the probability $\mathcal{P}_b^n(R_B)$ of a result R_B.

We now show that if a relation of the type (28) is verified for all the measurement parameters a and b, with any positive probabilities $\mathcal{P}_a^n(R_A)$ and $\mathcal{P}_b^n(R_B)$, and with positive values for all the π_n, Bell's inequalities are always satisfied. Let's assume both results R_A and R_B can take on the values ± 1.

To start with, we assume that the physical properties of each pair of systems A and B depend on a classical random variable μ; this variable takes on a series of values μ_n, corresponding to each term in the summation over n appearing in (28), each with the probability π_n. This means that this summation over n can be interpreted as an average value over the random variable μ.

We then assume that the properties of the classical system A depend on a different random variable ν_A, which determines the result of the measurement performed on A. As an example, one can imagine that ν_A is regularly distributed on the segment $[0,1]$; result R_A is a function of ν_A, and takes the value $+1$ on a fraction of the segment of length $\mathcal{P}_a^n(R_A = +1)$, and the value -1 on the rest of the segment. This function models the probability written on the first line of (29); the measurement result is thus a function of ν_A, of the measurement parameter a, and of μ (which replaces n). Finally, we introduce the random variable ν_B, which determines the result of the measurement performed on B, with a distribution modeling in a similar way the probability written on the second line of (29), for any value of n and any choice of the measurement parameter b.

If we now regroup the three variables μ, ν_A and ν_B, as being the three components of a single variable λ, we reproduce the exact same hypotheses stated at the beginning of § F-3-a in Chapter XXI: the measurement results are functions, the first one of λ and a and the other one of λ and b. The same reasoning then leads to Bell's inequalities. Note that, at no point in this classical reasoning, did we have to consider the ensemble $A+B$ as a whole; it was thus to be expected that Bell's inequality would be established in this case.

3-b. Two spins in a singlet state

Let's go back to the example of two spin-1/2 particles in a singlet state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big[\,|+,-\rangle - |-,+\rangle\Big] \tag{30}$$

In the basis of the 4 kets $|+,+\rangle$, $|+,-\rangle$, $|-,+\rangle$, $|-,-\rangle$ taken in that order, the matrix representing the density operator ρ_{A+B} is written:

$$(\rho_{A+B}) = |\Psi\rangle\langle\Psi| = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{31}$$

This matrix density (ρ_{A+B}) has non-diagonal elements between states $|+,-\rangle$ and $|-,+\rangle$, as, for example:

$$-1 = \langle +,-|\,\rho_{A+B}\,|-,+\rangle \tag{32}$$

To obtain such a non-diagonal term by a sum of products such as (26), will require:

$$-1 = \sum_n \pi_n\ \langle +|\,\rho_A^n\,|-\rangle\ \langle -|\,\rho_B^n\,|+\rangle \tag{33}$$

This demands introducing at least one term n that contains partial density operators ρ_A^n and ρ_B^n, both having non-diagonal elements. Now each of these two operators is a positive-definite operator. This means, for ρ_A^n for example, that it must have populations (diagonal matrix elements) $\langle +| \rho_A^n |+\rangle$ and $\langle -| \rho_A^n |-\rangle$ in the two individual spin states, and the same is true for ρ_B^n. The corresponding n term will necessarily introduce in (ρ_{A+B}) populations in the 4 states $|+,+\rangle$, $|+,-\rangle$, $|-,+\rangle$, $|-,-\rangle$; it will then be impossible to cancel those populations by adding other products of density operators (whose populations are positive) with positive π_n coefficients. Consequently, this density operator (ρ_{A+B}) is non-separable, and this is why it can lead to violations of Bell's inequalities.

Complement B_{XXI}

GHZ states, entanglement swapping

1 Sign contradiction in a GHZ state 2227
1-a Quantum calculation 2227
1-b Reasoning in the local realism framework 2230
1-c Discussion; contextuality 2231
2 Entanglement swapping 2232
2-a General scheme 2232
2-b Discussion 2235

Greenberger, Horne and Zeilinger (GHZ) showed in 1989 [83, 84] that violations of local realism even more spectacular than violations of Bell's inequalities could be observed on systems containing more than two correlated particles. These violations involve a contradiction in sign (and hence a violation of 100 %) for perfect correlations between measurement results, as opposed to inequalities violated by 40% for imperfect correlations. Observation of these violations requires creating an initial entanglement between three particles or more, as will be discussed in § 1. Another example, involving the entanglement of more than two particles, is the "entanglement swapping" method, explained in § 2. This method highlights a surprising property of entanglement: the possibility of entangling together two quantum systems, without them ever having to interact with each other.

1. Sign contradiction in a GHZ state

We consider a system composed of three spin-1/2 particles, as it is the simplest case for explaining how the GHZ contradiction can appear.

1-a. Quantum calculation

The three-spin system is described by the normalized quantum state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big[\,|+,+,+\rangle + \eta\,|-,-,-\rangle\Big] \tag{1}$$

In this equality, the states $|\pm\rangle$ symbolize the eigenstates of the spin components along the Oz axis in an $Oxyz$ reference frame; to simplify the notation of the three particle ket, the spins are not numbered: the first sign corresponds to the state of the first spin, the second to that of the second, and similarly for the third spin. The number η stands for either $+1$, or -1. We now look for the quantum probabilities of measurement results of the components of each of the spins $\sigma_{1,2,3}$ of the three particles along two possible directions: either along the Ox direction, or along the perpendicular Oy direction (Fig. 1).

We start with the measurement of the product $\sigma_{1y}\,\sigma_{2y}\,\sigma_{3x}$. As we now show, $|\Psi\rangle$ is an eigenvector of this operator product, with eigenvalue $-\eta$, which means that the

measurement result is certain. The action of the first operator[1] is written:

$$\begin{aligned}\sigma_{3x}|\Psi\rangle &= \frac{1}{2}\left[\sigma_+(3)+\sigma_-(3)\right]|\Psi\rangle \\ &= \frac{1}{2\sqrt{2}}\left[\eta\sigma_+(3)|-,-,-\rangle+\sigma_-(3)|+,+,+\rangle\right] \\ &= \frac{1}{\sqrt{2}}\left[\eta|-,-,+\rangle+|+,+,-\rangle\right] \end{aligned} \tag{2}$$

The second operator then yields:

$$\begin{aligned}\sigma_{2y}\sigma_{3x}|\Psi\rangle &= \frac{1}{2i}\left[\sigma_+(2)-\sigma_-(2)\right]\sigma_{3x}|\Psi\rangle \\ &= \frac{1}{i\sqrt{2}}\left[\eta|-,+,+\rangle-|+,-,-\rangle\right] \end{aligned} \tag{3}$$

Finally, the product of the three operators yields:

$$\begin{aligned}\sigma_{1y}\sigma_{2y}\sigma_{3x}|\Psi\rangle &= \frac{1}{2i}\left[\sigma_+(1)-\sigma_-(1)\right]\sigma_{2y}\sigma_{3x}|\Psi\rangle \\ &= -\frac{1}{\sqrt{2}}\left[\eta|+,+,+\rangle+|-,-,-\rangle\right] \\ &= -\eta|\Psi\rangle \end{aligned} \tag{4}$$

The probability of observing the result $-\eta$ is:

$$\mathcal{P}(\sigma_{1y}\,\sigma_{2y}\,\sigma_{3x} \Longrightarrow -\eta) = 1 \tag{5}$$

whereas the probability $\mathcal{P}(\sigma_{1y}\,\sigma_{2y}\,\sigma_{3x} \Longrightarrow +\eta)$ of observing the other result is zero.

As the three spins play the same role, it is clear that $|\Psi\rangle$ is also an eigenvector of the two operator products $\sigma_{1x}\,\sigma_{2y}\,\sigma_{3y}$ and $\sigma_{1y}\,\sigma_{2x}\,\sigma_{3y}$, with eigenvalues $-\eta$. The corresponding probabilities are therefore:

$$\begin{aligned}&\mathcal{P}(\sigma_{1x}\,\sigma_{2y}\,\sigma_{3y} \Longrightarrow -\eta) = 1 \\ &\mathcal{P}(\sigma_{1y}\,\sigma_{2x}\,\sigma_{3y} \Longrightarrow -\eta) = 1\end{aligned} \tag{6}$$

It is thus certain that the three products take on the value $-\eta$.

We now consider the measurement result of the product of the three spin components along the same Ox axis. We use again (2), but (3) is now replaced by:

$$\begin{aligned}\sigma_{2x}\sigma_{3x}|\Psi\rangle &= \frac{1}{2}\left[\sigma_+(2)+\sigma_-(2)\right]\sigma_{3x}|\Psi\rangle \\ &= \frac{1}{\sqrt{2}}\left[\eta|-,+,+\rangle+|+,-,-\rangle\right] \end{aligned} \tag{7}$$

and (4) by:

$$\begin{aligned}\sigma_{1x}\sigma_{2x}\sigma_{3x}|\Psi\rangle &= \frac{1}{2}\left[\sigma_+(1)+\sigma_-(1)\right]\sigma_{2x}\sigma_{3x}|\Psi\rangle \\ &= \frac{1}{\sqrt{2}}\left[\eta|+,+,+\rangle+|-,-,-\rangle\right] \end{aligned} \tag{8}$$

[1] The Pauli matrices are defined in Complement A_{IV}; the operators $\sigma_\pm = \sigma_x \pm i\sigma_y$ obey the relation $\sigma_\pm|\mp\rangle = 2|\pm\rangle$.

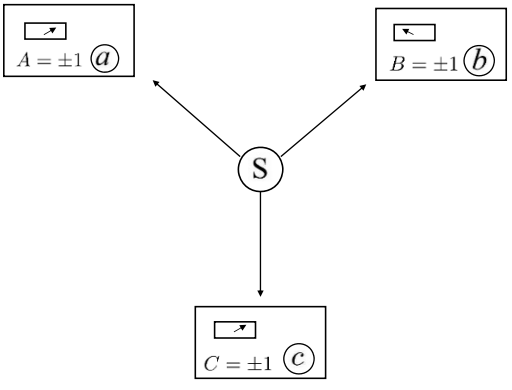


Figure 1: Schematic set-up for observing the GHZ contradictions. Three spins, initially in state $|\Psi\rangle$ written in (1), are measured in three different regions of space. In each of these regions a measuring apparatus is placed, with a setting enabling the local observer to choose between two possible spin component measurements, either along Ox, or along Oy. Whatever choices the three observers make, the results given by the three apparatus are $A = \pm 1$, $B = \pm 1$ and $C = \pm 1$.

This shows that $|\Psi\rangle$ is also an eigenvector of the operator product $\sigma_{1x}\,\sigma_{2x}\,\sigma_{3x}$, but this time with the eigenvalue $+\eta$. It follows that:

$$\mathcal{P}(\sigma_{1x}\,\sigma_{2x}\,\sigma_{3x} \Longrightarrow +\eta) = 1 \tag{9}$$

One can conclude, with certainty, that the measurement result of this product will be equal to $+\eta$.

Quantum mechanical measurement of a product of commuting operators:

Three operators such as σ_{1y}, σ_{2y} and σ_{3x}, acting on different spins, commute with each other; they form a CSCO (Complete Set of Commuting Observables) in the state space of the three spins. One can thus build a basis of eigenvectors $|\xi_1, \xi_2, \xi_3\rangle$ common to the three operators, labeled by the eigenvalue $\xi_1 = \pm 1$ of σ_{1y}, the eigenvalue $\xi_2 = \pm 1$ of σ_{2y} and the eigenvalue $\xi_3 = \pm 1$ of σ_{3x}. Any vector $|\Psi\rangle$ can be decomposed onto this basis as:

$$|\Psi\rangle = \sum_{\xi_1,\xi_2,\xi_3} c\,(\xi_1, \xi_2, \xi_3)\ \ |\xi_1, \xi_2, \xi_3\rangle \tag{10}$$

The action of the operator product $\sigma_{1y}\sigma_{2y}\sigma_{3x}$ on any ket is therefore to simply multiply each of its component $c\,(\xi_1, \xi_2, \xi_3)$ by the product $\xi_1\,\xi_2\,\xi_3$. Now the vector $|\Psi\rangle$ written in (1) is an eigenvector of that operator product, with the eigenvalue $-\eta$. The uniqueness of decomposition (10) then means that the only non-zero $c\,(\xi_1, \xi_2, \xi_3)$ coefficients are those

for which:

$$\xi_1\,\xi_2\,\xi_3 = -\eta \tag{11}$$

Suppose we measure, in a first experiment, the component σ_{1y} of the first spin. The result $\xi_1 = \pm 1$ is random. After the measurement, the projection postulate leads to a state that depends on this result, obtained by keeping in (1) only half of the components – those that correspond to the observed ξ_1 value. The components of the projected state vector still obey relation (11), where ξ_1 is now fixed. Similarly, if we continue the experiment and measure σ_{2y} for the second spin, the result $\xi_2 = \pm 1$ is also random, but the components of the new projected state vector still obey that same relation. As now ξ_1 and ξ_2 are both known, the same is true for ξ_3, whose value is determined by the first two measurements.

To sum up, the results observed for each spin component measurement fluctuate from one experiment to another, but these fluctuations are correlated and the product of the three results remain constant. One can obviously do the same analysis for the other sets of operators considered above, σ_{1x}, σ_{2y} and σ_{3y} for example.

1-b. Reasoning in the local realism framework

Let us leave, for a moment, standard quantum mechanics and examine what a local realistic theory (in the EPR sense of these words) would predict in such a situation. As we are in a particularly simple case where the initial quantum state is an eigenvector of all the observables coming into play (all the results are certain), one could expect nothing particular to happen. On the contrary, we now show that a complete contradiction appears between local realism and the predictions of quantum mechanics.

The local realism argument we present is a direct generalization of that used to obtain Bell's inequalities in § F-3-a of Chapter XXI. We first notice that the perfect correlations imply that the measurement result of a spin component along Ox (or Oy) of any particle can be deduced from the results of measurements performed on other particles, at arbitrarily large distances. The EPR argument then requires the existence of elements of reality corresponding to these two component directions, that we shall note $A_{x,y} = \pm 1$ for the first spin, $B_{x,y} = \pm 1$ for the second, and finally $C_{x,y} = \pm 1$ for the third. According to the EPR argument, for each experiment (i.e. for each emission of a group of three particles), these six numbers have well determined values, even though they are a priori unknown. These numbers are simply the results that shall be obtained, should measurements be performed later on. As an example, a measurement on the first spin will necessarily yield A_x if the chosen analysis direction is along Ox, or A_y if it is along Oy, independently of the type of measurements performed on the other two spins.

To have an agreement with the three equalities (5) and (6) imposes that:

$$\begin{aligned} A_y B_y C_x &= -\eta \\ A_x B_y C_y &= -\eta \\ A_y B_x C_y &= -\eta \end{aligned} \tag{12}$$

Now, in the logic of local realism, the same values of A, B and C can also be used for an experiment where the three spin components are measured along the same Ox direction: the result should simply be the product $A_x B_x C_x$. As the squares of the numbers A_y^2,

etc., are always equal to +1, we can obtain that product by multiplying the lines of (12), which yields:

$$A_x B_x C_x = -\eta \tag{13}$$

That is where the contradiction shows up: equality (9) predicts that the measurement of $\sigma_{1x}\,\sigma_{2x}\,\sigma_{3x}$ must always yield the result $+\eta$, which has the opposite sign! There could not be a greater contradiction between local realism and quantum mechanical predictions.

1-c. Discussion; contextuality

Compared with the violations of Bell's inequality, the GHZ contradiction seems far more spectacular, since a 100 % contradiction is obtained with 100 % certainty. From an experimental point of view, however, the necessity to bring into play three remote and entangled particles is a complex challenge.

To easily identify the three spins (deciding which measurement pertains to spin noted A, to spin noted B, and to spin noted C), and to be sure the three measurements are performed far from each other, let us assume the spins each occupy a different region of space. When the spatial variables are taken into account, the ket (1) can be rewritten more explicitly in the form:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\,|1:u_a\rangle\,|2:u_b\rangle\,|3:u_c\rangle \otimes \Big[\,|1:+;2:+;3:+\rangle + \eta\;\;|1:-;2:-;3:-\rangle\Big] \tag{14}$$

where $|u_{a,b,c}\rangle$ are three orbital states whose wave functions do not overlap. They can for example be entirely localized in separate boxes where the measurements are performed. One then assumes that none of the particles will be left unmeasured and that each of them is separately observed. The experimental procedure is to first choose, for each box, a component Ox or Oy, then perform the three corresponding measurements in each box, obtain the three results $A_{x,y}$, $B_{x,y}$ and $C_{x,y}$, and finally compute their product.

A first necessary verification is to perform a large number of experiments and measure successively the three products$A_y B_y C_x$, $A_x B_y C_y$ and $A_y B_x C_y$, to be sure that the perfect correlations predicted by quantum mechanics are indeed observed (it is an essential step for the EPR argument, which infers from it the existence of 6 separate elements of reality). One then measures the product $A_x B_x C_x$ and, if quantum mechanics is right, one will observe a sign opposite to the EPR prediction. This means that the value obtained in a measurement of σ_{1x} (for example) depends on the Ox or Oy components measured on the other spins; this remains true even if the corresponding operators commute with σ_{1x}. This leads us to the general concept of "quantum contextuality": in an experiment where several commuting observables are measured, one must take into account, according to Bohr's prescription, the ensemble of the experimental set-up (the whole context of the system to be measured); it would not be correct to reason as if these measurements were independent processes.

Experimental tests of GHZ equalities have been performed [85, 86]. These experiments require three particles to be placed in the quantum state (14), which is not an easy task. Nevertheless, using elaborate quantum optics techniques, the correctness of quantum mechanical predictions has been verified in such a case, with experiments involving 3 or 4 entangled photons, as well as with NMR (Nuclear Magnetic Resonance) techniques.

2. Entanglement swapping

We now describe the "entanglement swapping" method, which enables entangling particles coming from independent sources (i.e. having no common past) through a quantum measurement process.

2-a. General scheme

Consider two sources S_{12} and S_{34} each creating a pair of entangled photons (Fig. 2). The first one creates a photon with momentum $\mathbf{k}_1$ and another one with momentum $\mathbf{k}_2$, whose polarizations are entangled in states H (horizontal polarization, in the plane of the figure) and V (vertical polarization, perpendicular to the plane of the figure). In a similar way, the second source creates a photon with momentum $\mathbf{k}_3$ and another one with momentum $\mathbf{k}_4$, whose polarizations are entangled in the same way. The initial state describing the two pairs is the tensor product of two states, each describing two particles:

$$|\Psi\rangle = \frac{1}{2}\left[|\mathbf{k}_1, H; \mathbf{k}_2, V\rangle + |\mathbf{k}_1, V; \mathbf{k}_2, H\rangle\right] \otimes \left[|\mathbf{k}_3, H; \mathbf{k}_4, V\rangle + |\mathbf{k}_3, V; \mathbf{k}_4, H\rangle\right] \tag{15}$$

While the two photons emitted by a given source are strongly entangled, no entanglement exists between the two pairs of photons, emitted by each of the two sources. It is useful to introduce the four different states pertaining to the wave vectors $\mathbf{k}_i$, $\mathbf{k}_j$:

$$\begin{aligned}
\left|\Phi^B_{i,j}\right\rangle_{(\eta)} &= \frac{1}{\sqrt{2}}\left[|\mathbf{k}_i, H; \mathbf{k}_j, H\rangle + \eta\,|\mathbf{k}_i, V; \mathbf{k}_j, V\rangle\right] \\
\left|\Theta^B_{i,j}\right\rangle_{(\eta)} &= \frac{1}{\sqrt{2}}\left[|\mathbf{k}_i, H; \mathbf{k}_j, V\rangle + \eta\,|\mathbf{k}_i, V; \mathbf{k}_j, H\rangle\right]
\end{aligned} \tag{16}$$

with, here again, $\eta = \pm 1$. These states (often called "Bell states" in the literature, hence the superscript B) form an orthonormal basis of the state space associated with particles i and j. One can show that (the computation is straightforward but a bit tedious and will not be detailed here):

$$\left|\Phi^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(+1)} - \left|\Phi^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(-1)} = |H; V; V; H\rangle + |V; H; H; V\rangle \tag{17}$$

and that:

$$\left|\Theta^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(+1)} - \left|\Theta^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(-1)} = |H; V; H; V\rangle + |V; H; V; H\rangle \tag{18}$$

(to simplify the notation, it is implicitly assumed, on the right-hand side of both equations, that the order of the particle's momenta is always $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$ and $\mathbf{k}_4$). We can then write state (15) in the form:

$$\begin{aligned}
|\Psi\rangle = \frac{1}{2}\Big[&\left|\Phi^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(+1)} - \left|\Phi^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(-1)} + \\
&+ \left|\Theta^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(+1)} - \left|\Theta^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(-1)}\Big]
\end{aligned} \tag{19}$$

Figure 2 schematizes the experiment to be performed. After they are emitted, the particles with momenta $\mathbf{k}_2$ and $\mathbf{k}_3$ undergo a measurement in which they interfere. This is achieved by sending these two particles to a beam splitter BS, followed by two

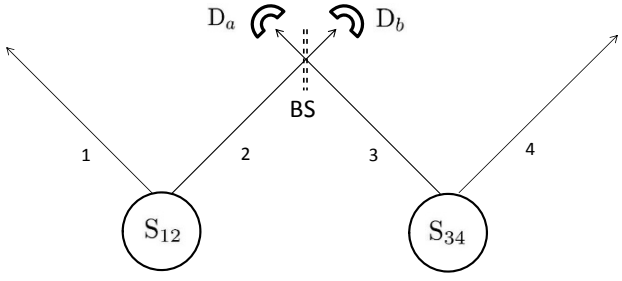


Figure 2: Schematic diagram of the "entanglement swapping" method. Two sources S_{12} and S_{34} each emit a pair of entangled particles, with wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ for the first one, $\mathbf{k}_3$ and $\mathbf{k}_4$ for the second. These sources are independent. A beam splitter BS is inserted in the path of particles $\mathbf{k}_2$ and $\mathbf{k}_3$; it is followed by two detectors D_a and D_b that measure the particle number in each of the exit channels a and b. This measurement has the effect of projecting the state vector, hence bringing the two particles $\mathbf{k}_1$ and $\mathbf{k}_4$ into a totally entangled state, even though these particles have never interacted.

detectors D$_a$ and D$_b$ measuring which exit channel were followed by the particles. If the two particles exit through two different channels, the corresponding eigenvector for this measurement result is the state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$; this is because, as we show below, the three other states $\left|\Theta^B_{2,3}\right\rangle_{(+1)}$, $\left|\Phi^B_{2,3}\right\rangle_{(+1)}$ and $\left|\Phi^B_{2,3}\right\rangle_{(-1)}$ correspond to situations where the two particles exit through the same channel. The measurement thus projects state (19) onto the last of its four components. The net result is that if the two particles with momenta $\mathbf{k}_2$ and $\mathbf{k}_3$ exit through different channels (which happens one out of four times), the two particles with momenta $\mathbf{k}_1$ and $\mathbf{k}_4$ reach the state $\left|\Theta^B_{1,4}\right\rangle_{(-1)}$. This means that the two non-observed particles reach a totally entangled state though they can be arbitrarily far from each other. It is worth noting that the initial entanglement concerns the two particles $\mathbf{k}_1$ and $\mathbf{k}_2$, and, separately, the two particles $\mathbf{k}_3$ and $\mathbf{k}_4$. Performing a suitable measurement on a particle of each pair, one projects the two remaining particles into a strongly entangled state, even though they never interacted at any stage of the process.

Demonstration:

Let us show that $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ is an initial state of two interfering particles that will lead to their exiting through different channels. We introduce for that purpose the two creation operators $a^\dagger_{\mathbf{k}_2,H}$ and $a^\dagger_{\mathbf{k}_2,V}$ in the individual state with wave vector $\mathbf{k}_2$ and polarization H or V, as well as the two operators $a^\dagger_{\mathbf{k}_3,H}$ and $a^\dagger_{\mathbf{k}_3,V}$ in the individual state with wave vector $\mathbf{k}_3$ and polarization H or V. The state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ can be written:

$$\left|\Theta^B_{2,3}\right\rangle_{(-1)} = \frac{1}{\sqrt{2}}\left[a^\dagger_{\mathbf{k}_2,H}a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,H}\right]|0\rangle \tag{20}$$

As the particles go through the beam splitter, their polarizations are not modified, but

their wave vectors are. In terms of creation operators, this leads to the unitary transformations:

$$\begin{aligned} a^\dagger_{\mathbf{k}_2,H} &\Rightarrow \frac{1}{\sqrt{2}}\left[a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right] \\ a^\dagger_{\mathbf{k}_3,H} &\Rightarrow \frac{1}{\sqrt{2}}\left[ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right] \end{aligned} \tag{21}$$

where the i factors come from the phase change in a light beam as it undergoes internal reflection. Similar equalities are obtained for the V polarization, so that:

$$\begin{aligned} &a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}\ a^\dagger_{\mathbf{k}_3,H} \\ &\Rightarrow \frac{1}{2}\Big[\left(a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right)\left(ia^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,V}\right) \\ &\qquad\qquad - \left(a^\dagger_{\mathbf{k}_2,V} + ia^\dagger_{\mathbf{k}_3,V}\right)\left(ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right)\Big] \end{aligned} \tag{22}$$

As creation operators in different modes commute with each other, this operator is equal to:

$$\left[a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}\ a^\dagger_{\mathbf{k}_3,H}\right] \tag{23}$$

so that state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ is transformed, after the beam splitter, into:

$$\left|\Theta^B_{2,3}\right\rangle_{(-1)} \Rightarrow \frac{1}{\sqrt{2}}\left[a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,H}\right]|0\rangle \tag{24}$$

This shows that if the state before crossing the beam splitter is $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$, the two photons are still in two different exit channels after the crossing.

If now the state before crossing the beam splitter is $\left|\Theta^B_{2,3}\right\rangle_{(+1)}$, we must replace (22) by:

$$\begin{aligned} a^\dagger_{\mathbf{k}_2,H}a^\dagger_{\mathbf{k}_3,V} + a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,H} &\Rightarrow \frac{1}{2}\left[\left(a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right)\left(ia^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,V}\right)\right. \\ &\qquad \left. + \left(a^\dagger_{\mathbf{k}_2,V} + ia^\dagger_{\mathbf{k}_3,V}\right)\left(ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right)\right] \\ &= i\left[a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,H}\ a^\dagger_{\mathbf{k}_3,V}\right] \end{aligned} \tag{25}$$

which means that the two photons always exit the beam splitter through the same channel. In the same way, for the state $\left|\Phi^B_{2,3}\right\rangle_{(\pm1)}$, we get the operator:

$$\begin{aligned} a^\dagger_{\mathbf{k}_2,H}a^\dagger_{\mathbf{k}_3,H} \pm a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,V} &\Rightarrow \frac{1}{2}\left[\left(a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right)\left(ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right)\right. \\ &\qquad \left. \pm \left(a^\dagger_{\mathbf{k}_2,V} + ia^\dagger_{\mathbf{k}_3,V}\right)\left(ia^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,V}\right)\right] \\ &= i\left[\left(a^\dagger_{\mathbf{k}_2,H}\right)^2 + \left(a^\dagger_{\mathbf{k}_3,H}\right)^2 \pm \left(a^\dagger_{\mathbf{k}_2,V}\right)^2 \pm \left(a^\dagger_{\mathbf{k}_3,V}\right)^2\right] \end{aligned} \tag{26}$$

It shows again that for each term the photons exit through the same channel. The state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ is therefore the only one that will lead to the photons exiting through different channels.

2-b. Discussion

In classical physics, it is also possible to obtain correlations between two objects initially totally independent, by sorting objects with which each of them is correlated. To underline the fundamental difference with entanglement swapping, we now discuss such a classical experiment. Imagine that two independent sources emit pairs of correlated objects, numbered 1 and 2 for the first source, 3 and 4 for the second, as in Figure 2. Each time the experiment is performed, each source emits two classical objects sharing a common property (such as, for example, the same color, or opposite angular momenta, etc.). The two sources are nevertheless totally uncorrelated (the objects emitted by two different sources present no correlations between their colors, their angular momenta, etc.). If, however, one selects particular experiments where particles 2 and 3 present a certain correlation (for example identical colors, or else parallel or antiparallel angular momenta, etc.), it is clear that the particles 1 and 4 will also be correlated, even if they never interacted in the past and if they are very far apart. It is a mere consequence of the selection performed in a classical probability distribution, and could be called "classical correlation swapping".

Note, however, that this selection remains purely classical; no entanglement can be produced by this method. Should a Bell experiment be performed on the objects 1 and 4, the correlations obtained will necessarily obey Bell's inequalities since we are in a classical physics context. The entanglement swapping method, however, allows creating by selection a true entanglement leading to strong violations of Bell's inequalities. This method is a way of producing stronger correlations than classical correlation swapping, and has been demonstrated in several experiments [87, 88].

Conclusion

The two examples we discussed illustrate the variety of situations where quantum entanglement produces significant physical effects, even when the entangled quantum systems are arbitrarily far from each other. In each situation, it is essential to follow the basic rules of quantum mechanics, and perform the computations with a global state vector, including all the physical systems under study. Any attempt to perform separate computations in different regions of space, and then add correlations using classical probability calculations, will necessarily lead to predictions ignoring numerous non-local quantum effects, in contradiction with experimental results.

Complement C_{XXI}

Measurement induced relative phase between two condensates

1 Probabilities of single, double, etc. position measurements 2239
- 1-a Single measurement (one particle) . . . 2239
- 1-b Double measurement (two particles) . . . 2240
- 1-c Generalization: measurement of any number of positions . . . 2241

2 Measurement induced enhancement of entanglement 2242
- 2-a Measuring the single density $\mathcal{P}(x_1)$ . . . 2243
- 2-b Entanglement between the two modes after the first detection 2243
- 2-c Measuring the double density $\mathcal{P}(x_2, x_1)$ . . . 2244
- 2-d Discussion . . . 2244

3 Detection of a large number Q of particles 2245
- 3-a Probability of a multiple detection sequence . . . 2245
- 3-b Discussion; emergence of a relative phase . . . 2247

Introduction

Referring to Bose-Einstein condensation (Complement C_{XV}, a system of identical bosons, all occupying the same individual state $|u_i\rangle$ is called a "condensate". It is described by a Fock state such as the one given by (A-17) of Chapter XV, where all the occupation numbers are zero, except for n_i, whose value N can be very large:

$$|u_i : N\rangle = \frac{1}{\sqrt{N!}} \left[a^\dagger_{u_i}\right]^N |0\rangle \tag{1}$$

The operator $a^\dagger_{u_i}$ is the creation operator of a particle in the individual state $|u_i\rangle$, and $|0\rangle$ is the vacuum state (for which all the occupation numbers are zero). In a similar way, a "double condensate" is described by a Fock state where N_1 particles are in the individual state $|u_i\rangle$ and N_2 particles in the individual state $|u_j\rangle$; its normalized state is written as:

$$|\Phi_0\rangle = |u_i : N_1; u_j : N_2\rangle = \frac{1}{\sqrt{N_1!N_2!}} \left[a^\dagger_{u_i}\right]^{N_1} \left[a^\dagger_{u_j}\right]^{N_2} |0\rangle \tag{2}$$

We shall focus on the case where the individual states u_i and u_j are states with well defined but opposite wave vectors $\pm k$:

$$|\Phi_0\rangle = |+k : N_1; -k : N_2\rangle = \frac{1}{\sqrt{N_1!N_2!}} \left[a^\dagger_{+k}\right]^{N_1} \left[a^\dagger_{-k}\right]^{N_2} |0\rangle \tag{3}$$

In such a state, while the occupation numbers are perfectly well defined, the relative phase between the two condensates is completely undetermined; we will confirm this

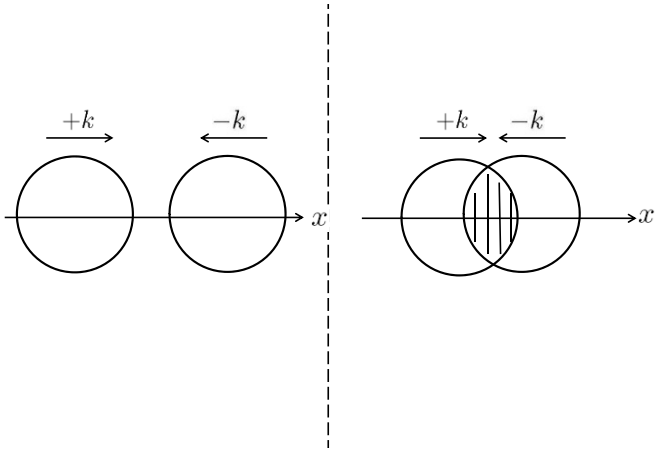


Figure 1: The left-hand side of the figure represents two groups of particles prepared independently. The first one is composed of a large number of particles, N_1, all in the same individual state with momentum $+k$ along the Ox axis, and propagating towards the right; the second group includes N_2 particles, in the other individual state with opposite $-k$ momentum, propagating towards the left. Each of these groups of independent particles is in a "condensate". The right-hand side of the figure shows that, after a certain time, the two condensates overlap in space; this allows measuring the positions of the particles in the overlap region. For clarity, the computations are limited to one dimension, taking into account only the x coordinate.
The first position measurement is totally random, but as measurements continue, there appear a periodic bunching of the observed positions, progressively forming a sharper fringe pattern. These fringes result from the emergence of a relative phase between the two condensates, which can only be a consequence of the position measurements, as it was totally absent at the beginning of the experiment.
If the whole process is repeated from the beginning, fringes appear with a position generally different from the first experiment: the phase appearing in each new experiment is totally independent of the one observed in previous experiments.

later (§ 2-a) by showing that measuring the position of a single particle with such a state does not lead to any observable interference fringes.

Now, recent experiments [89] have shown that when the positions of many particles are measured, interference fringes can indeed be observed in the region where the two condensates overlap (Figure 1). This remains true even if the condensates have been created in a totally independent way. This fringe pattern corresponds to a well defined value of the relative phase of the two condensates; one may then wonder about the origin of this observed phase. The object of this complement is to study the mechanism responsible for the emergence of this relative phase. We will show that it results from the successive detections of particles, which progressively modifies the initial state: as more and more particles are detected, it produces a progressively increasing entanglement between the two condensates, defining their relative phase in a more and more precise way.

During the course of one experiment, the position of the fringes is determined. However, should one repeat the experiment, preparing the condensates in exactly the same way, a new relative phase will progressively appear during the successive particle detections; its value is, in general, completely different from the one previously obtained. This means that if one averages observations over a large number of independent successive experiments, the fringes will be blurred and eventually completely disappear. The emergence of the phase is clearly observable only in the course of one specific experiment.

We first compute, in § 1, the probability of measurements concerning the positions of one, two, and more particles; we will show that these probabilities are proportional to spatio-temporal correlation functions of field operators of the various particles. Starting, in § 2, from two condensates in an initial state described by a simple juxtaposition of two Fock states, we will see how the successive particles' position measurements create an increasing entanglement between the two condensates. A more general study of the system's evolution is presented in § 3, showing, in particular, how this growing entanglement leads to a better and better definition of the relative phase between the two condensates. The computations presented in this complement are limited to the case where the number of measured positions remains small compared to the total number of particles in the condensates. Complement D_{XXI}, will go a step further and relax this hypothesis.

1. Probabilities of single, double, etc. position measurements

As we start the successive measurements of the particles' positions, we begin by computing the probability of finding a first particle in an interval[1] of infinitesimal width Δ around position $x = x_1$, then a second particle in the interval of width Δ around position $x = x_2$, etc. The computations we present here are valid for any state $|\Phi_0\rangle$ of the identical particle system. They are, actually, the equivalent of those encountered in the general study of correlation functions in § B-3 of Chapter XVI; nevertheless, we will go through them again in the specific context of the present complement. The results will be applied, in § 2, to the particular case where $|\Phi_0\rangle$ is a double Fock state.

1-a. Single measurement (one particle)

With a measurement of the position yielding a result included in the interval $D_1 = \left[x_1 - \frac{\Delta}{2}, x_1 + \frac{\Delta}{2}\right]$, we can associate the Hermitian operator:

$$P_\Delta(x_1) = \int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x\ \Psi^\dagger(x)\,\Psi(x) \tag{4}$$

where $\Psi(x)$ is the field operator destroying a particle at point x, and $\Psi^\dagger(x)$ its Hermitian conjugate, creating a particle. The average value of $P_\Delta(x_1)$ yields the average particle number in the interval D_1. In what follows, we shall, most of the time, assume that Δ is small enough compared to the other dimensions of the problem to justify the approximation:

$$P_\Delta(x_1) \simeq \Delta \times \Psi^\dagger(x_1)\,\Psi(x_1) \tag{5}$$

[1] To keep the notation simple, we consider a one-dimensional problem and note x_1, x_2,..,x_p, the particles' positions. Generalizing to three dimensions only requires replacing all the x_i by the vectors $\mathbf{r}_i$.

Operator $P_\Delta(x_1)$ is a symmetric one-particle operator of the type described in relation (B-1) in Chapter XV. It can also be written as:

$$P_\Delta(x_1) = \sum_{q=1}^{N} \int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x \ |q:x\rangle\langle q:x| \tag{6}$$

which is the sum over all the $q = 1, 2, .., N$ particles of the projectors into the interval D_1 of the positions of each of them. As all these projectors commute with each other, and since they each have eigenvalues 1 and 0, the eigenvalues of $P_\Delta(x_1)$ are equal to 0, 1, 2, .. N. Now, if Δ is small enough, there can be no more than one particle in the interval D_1; this means that the only accessible eigenvalues are 0 and 1, so that $P_\Delta(x_1)$ becomes the projector associated with the measurement of a particle's presence in the interval D_1:

$$[P_\Delta(x_1)]^2 = P_\Delta(x_1) \qquad \text{if } \Delta \to 0 \tag{7}$$

Suppose now the system is in state $|\Phi_0\rangle$. The probability $\mathcal{P}_{D_1}$ of finding a particle in the infinitesimal interval D_1 of length Δ is:

$$\begin{aligned}\mathcal{P}_{D_1}(x_1) &= \langle\Phi_0| P_\Delta(x_1) |\Phi_0\rangle \\ &= \Delta \ \langle\Phi_0| \Psi^\dagger(x_1)\Psi(x_1)|\Phi_0\rangle\end{aligned} \tag{8}$$

Right after the detection of this first particle, the system is now, according to relation (E-39) of Chapter III (postulate of wave packet reduction), in the normalized state:

$$|\Phi_0'\rangle = \frac{1}{\sqrt{\mathcal{P}_{D_1}(x_1)}} P_\Delta(x_1)|\Phi_0\rangle \tag{9}$$

1-b. Double measurement (two particles)

Let us now focus on the probability $\mathcal{P}_{D_1D_2}(x_2, x_1)$ of detecting a first particle in an interval of width Δ around point x_1, then a second one in an interval Δ around point x_2; we assume the system does not have time to evolve in between the two measurements.

We start by computing the conditional probability[2] $\mathcal{P}_{D_1D_2}(x_2/x_1)$ of detecting a particle in the interval $\left[x_2-\frac{\Delta}{2}, x_2+\frac{\Delta}{2}\right]$ noted D_2, knowing that a particle has been detected in the interval $\left[x_1-\frac{\Delta}{2}, x_1+\frac{\Delta}{2}\right]$ already noted D_1. This probability equals:

$$\begin{aligned}\mathcal{P}_{D_1D_2}(x_2/x_1) &= \langle\Phi_0'| P_\Delta(x_2)|\Phi_0'\rangle \\ &= \frac{1}{\mathcal{P}_{D_1}(x_1)}\langle\Phi_0| P_\Delta(x_1)P_\Delta(x_2)P_\Delta(x_1)|\Phi_0\rangle\end{aligned} \tag{10}$$

where, in the second line, we have used (9); the projector $P_\Delta(x_2)$ may be obtained by replacing x_1 by x_2 in expression (4). We assume that the two detection intervals do

[2]Note the different notation used for the conditional probability $\mathcal{P}_{D_1D_2}(x_2/x_1)$, with a fraction bar between the variables, and the simple probability (a priori probability) $\mathcal{P}_{D_1D_2}(x_2, x_1)$ of obtaining the two results. These two probabilities are related by expression (14).

not overlap in space, so that all the operators appearing in $P_\Delta(x_1)$ commute with those appearing in $P_\Delta(x_2)$. We then have, taking (7) into account:

$$\begin{aligned} P_\Delta(x_1) P_\Delta(x_2) P_\Delta(x_1) &= [P_\Delta(x_1)]^2 P_\Delta(x_2) \\ &= P_\Delta(x_1) P_\Delta(x_2) \end{aligned} \tag{11}$$

If Δ is small enough, we can replace $P_\Delta(x_1)$ and $P_\Delta(x_2)$ by their expressions (5); this leads to:

$$\begin{aligned} P_\Delta(x_1) P_\Delta(x_2) &= \Delta^2 \ \Psi^\dagger(x_1) \Psi(x_1) \Psi^\dagger(x_2) \Psi(x_2) \\ &= \Delta^2 \ \Psi^\dagger(x_1) \Psi^\dagger(x_2) \Psi(x_2) \Psi(x_1) \end{aligned} \tag{12}$$

where, in the second line, we again used the fact that field operators defined in non-overlapping regions of space commute with each other. Inserting this result in (10), we get:

$$\mathcal{P}_{D_1 D_2}(x_2/x_1) = \frac{\Delta^2}{\mathcal{P}_{D_1}(x_1)} \langle\Phi_0| \Psi^\dagger(x_1) \Psi^\dagger(x_2) \Psi(x_2) \Psi(x_1) |\Phi_0\rangle \tag{13}$$

Now, the probability of detecting a particle at point x_1, then a particle at point x_2, is the product of the probability $\mathcal{P}_{D_1}(x_1)$ of detecting a particle at point x_1 and the conditional probability $\mathcal{P}_{D_1 D_2}(x_2/x_1)$ of detecting a particle at point x_2 knowing that a particle has been detected at point x_1:

$$\mathcal{P}_{D_1 D_2}(x_2, x_1) = \mathcal{P}_{D_1}(x_1) \times \mathcal{P}_{D_1 D_2}(x_2/x_1) \tag{14}$$

Taking (13) into account, this leads to:

$$\begin{aligned} \mathcal{P}_{D_1 D_2}(x_2, x_1) &= \Delta^2 \langle\Phi_0| \Psi^\dagger(x_1) \Psi^\dagger(x_2) \Psi(x_2) \Psi(x_1) |\Phi_0\rangle \\ &= \langle\Phi_{21} \ |\Phi_{21}\rangle \end{aligned} \tag{15}$$

where $|\Phi_{21}\rangle$ is the non-normalized state:

$$|\Phi_{21}\rangle = \Delta \times \Psi(x_2) \Psi(x_1) |\Phi_0\rangle \tag{16}$$

The probability we are looking for is simply the squared norm of the ket obtained by destroying in the initial state a particle at point x_1, and a second one at point x_2, multiplied by the width Δ of the infinitesimal measurement interval.

1-c. Generalization: measurement of any number of positions

The previous computations deal with simple and double density measurements; we now generalize them to measurements of higher order densities. From now on and to simplify the notation, we shall omit the D subscript in the probabilities $\mathcal{P}$.

To compute the probability associated with a triple measurement, we start from the expression of the state vector right after the detection of the second particle at x_2. Taking (10) into account, and similarly as for (9), this normalized state is written:

$$|\Phi_0''\rangle = \frac{1}{\sqrt{\langle\Phi_0'| P_\Delta(x_2) |\Phi_0'\rangle}} P_\Delta(x_2) |\Phi_0'\rangle = \frac{1}{\sqrt{\mathcal{P}(x_2/x_1)}} P_\Delta(x_2) |\Phi_0'\rangle \tag{17}$$

or else, if we insert (9) and use (14):

$$|\Phi_0''\rangle = \frac{1}{\sqrt{\mathcal{P}(x_2/x_1)\mathcal{P}(x_1)}}\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle$$
$$= \frac{1}{\sqrt{\mathcal{P}(x_2,x_1)}}\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle \tag{18}$$

The probability of the third measurement at x_3, knowing that the first two measurements gave results at x_1 and x_2, is thus:

$$\mathcal{P}(x_3/x_2,x_1) = \langle\Phi_0''|\, P_\Delta(x_3)\,|\Phi_0''\rangle$$
$$= \frac{1}{\mathcal{P}(x_2,x_1)}\,\langle\Phi_0|\, P_\Delta(x_1)\, P_\Delta(x_2)\, P_\Delta(x_3)\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle \tag{19}$$

As before, we consider that the position measurement zones do not overlap, so that all the projection operators commute with each other:

$$\mathcal{P}(x_3/x_2,x_1) = \frac{1}{\mathcal{P}(x_2,x_1)}\,\langle\Phi_0|\, P_\Delta(x_3)\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle$$
$$= \frac{\Delta^3}{\mathcal{P}(x_2,x_1)}\,\langle\Phi_0|\,\Psi^\dagger(x_1)\,\Psi^\dagger(x_2)\,\Psi^\dagger(x_3)\,\Psi(x_3)\,\Psi(x_2)\,\Psi(x_1)\,|\Phi_0\rangle \tag{20}$$

In the second line, we assumed Δ was small enough to use the approximate relation (5).

As the law of conditional probabilities indicates that the probability of the three measurements at x_1, x_2 and x_3 is given by:

$$\mathcal{P}(x_3,x_2,x_1) = \mathcal{P}(x_2,x_1)\times\mathcal{P}(x_3/x_2,x_1) \tag{21}$$

we simply get:

$$\mathcal{P}(x_3,x_2,x_1) = \Delta^3\,\langle\Phi_0|\,\Psi^\dagger(x_1)\,\Psi^\dagger(x_2)\,\Psi^\dagger(x_3)\,\Psi(x_3)\,\Psi(x_2)\,\Psi(x_1)\,|\Phi_0\rangle \tag{22}$$

which is a direct generalization of (15).

The same line of reasoning allows showing that the probability associated with the measurement of Q positions is proportional to the average value in the system's state of a product of $2Q$ field operators $\Psi^\dagger$ and Ψ arranged in normal order, and evaluated at x_1, x_2,...x_Q. The probabilities are therefore equal to the spatio-temporal correlation functions of the field operators arranged in normal order (and multiplied by Δ^Q).

2. Measurement induced enhancement of entanglement

We have reasoned until now in a general way, without specifying the initial state $|\Phi_0\rangle$ of the system under study. We now assume we are dealing with a double condensate, as in (3), and for simplicity we shall take $N_1 = N_2 = N$ (actually the computation that follows only requires the hypothesis $N_1 \simeq N_2$). We thus have N particles occupying the individual state $|k\rangle$ with a well-defined momentum $\hbar k$, and an equal number of particles occupying the state $|-k\rangle$ with opposite momentum:

$$|\Phi_0\rangle = |+k:N;-k:N\rangle = \frac{1}{N!}\left[a^\dagger_{+k}\right]^N\left[a^\dagger_{-k}\right]^N|0\rangle \tag{23}$$

We propose studying the interference signals that may occur in the single and double counting rates measured on such a state. We shall need the probabilities calculated above as well as expressions (A-3) and (A-6) of Chapter XVI for the field operators:

$$\Psi(x) = \frac{1}{L^{1/2}}\left[e^{ikx}a_k + e^{-ikx}a_{-k} + ..\right]$$
$$\Psi^\dagger(x) = \frac{1}{L^{1/2}}\left[e^{-ikx}a_k^\dagger + e^{ikx}a_{-k}^\dagger + ..\right] \tag{24}$$

where a_k (or a_{-k}) and $a_k^\dagger$ (or $a_{-k}^\dagger$) are the annihilation and creation operators of a particle in mode k (or $-k$), and where L is the edge of the box used to normalize the plane waves. The dots on the right-hand side of these formulas stand for the other terms present in the field operator expansions of these operators. Because of the choice of the initial state (23), these additional terms do not play any role in the following calculations, as will be shown below.

2-a. Measuring the single density $\mathcal{P}(x_1)$

Relation (8) now becomes:

$$\mathcal{P}(x_1) = \Delta\ \langle +k:N;-k:N|\,\Psi^\dagger(x_1)\Psi(x_1)\,|+k:N;-k:N\rangle \tag{25}$$

Using expressions (24) for the field operators, and the fact that the cross terms $a_{-k}^\dagger a_k$ and $a_k^\dagger a_{-k}$ have a zero average value in the double Fock state (23), we get:

$$\mathcal{P}(x_1) = \frac{\Delta}{L}\langle +k:N;-k:N|\,a_k^\dagger a_k + a_{-k}^\dagger a_{-k}\,|+k:N;-k:N\rangle = \frac{2N\Delta}{L} \tag{26}$$

This means that there is no interference in the single density measurement signal. This was to be expected since the initial double Fock state includes no phase that could help determine the eventual position of such fringes.

2-b. Entanglement between the two modes after the first detection

Relation (9) yields the ket $|\Phi_0'\rangle$, right after the first measurement. It can be written as:

$$|\Phi_0'\rangle = \frac{1}{\sqrt{\mathcal{P}(x_1)}}\int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x\ \Psi^\dagger(x)\,\Psi(x)\ \ |\Phi_0\rangle$$
$$= \frac{1}{L\sqrt{\mathcal{P}(x_1)}}\int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x\ \left[e^{-ikx}a_k^\dagger + e^{ikx}a_{-k}^\dagger + ..\right]\left[e^{ikx}a_k + e^{-ikx}a_{-k} + ..\right]|\Phi_0\rangle \tag{27}$$

Taking (23) into account, and for an infinitesimal Δ, we get:

$$|\Phi_0'\rangle \propto \Big\{2N\,|\Phi_0\rangle + \sqrt{N(N+1)}\left[e^{-2ikx_1}\,|k:N+1;-k:N-1\rangle\right.$$
$$\left.+e^{2ikx_1}\,|k:N-1;-k:N+1\rangle\right] + ..\Big\} \tag{28}$$

where +.. stands for the components of $|\Phi_0'\rangle$ where a particle occupies an individual state other than k and $-k$; these components do not play any role in what follows. Relation (28) shows that the entanglement of state $|\Phi_0'\rangle$ has increased as a result of the detection of the first particle. This state now contains a linear superposition of the initial state and two additional states of the global system, $|N-1:+k;N:-k\rangle$ and $|N:+k;N-1:-k\rangle$; the coefficients of this superposition, and in particular their relative phase, depend on the point x_1 where the first particle has been detected.

2-c. Measuring the double density $\mathcal{P}(x_2,x_1)$

We now compute the probability $\mathcal{P}(x_2,x_1)$ associated with a double density measurement. Relations (15) and (16) show that the probability is the squared norm of the ket:

$$|\Phi_{21}\rangle = \Delta \times \Psi(x_2)\,\Psi(x_1)\,|N:+k;N:-k\rangle \tag{29}$$

Inserting in this equality the first relation (24), the terms symbolized by the dots .. disappear (as they involve annihilation operators yielding zero when acting on the $+k$ and $-k$ states, the only initially populated states). We obtain:

$$|\Phi_{21}\rangle = \frac{\Delta}{L} \times \left[e^{ikx_2}a_k + e^{-ikx_2}a_{-k}\right]\left[e^{ikx_1}a_k + e^{-ikx_1}a_{-k}\right]|N:+k;N:-k\rangle \tag{30}$$

or else:

$$\begin{aligned}|\Phi_{21}\rangle = \frac{\Delta}{L} \times \Big\{ \sqrt{N(N-1)}\Big[& e^{ik(x_1+x_2)}\,|N-2:+k;N:-k\rangle \\ & + e^{-ik(x_1+x_2)}\,|N:+k;N-2:-k\rangle\Big] \\ + N\left[e^{ik(x_2-x_1)} + e^{ik(x_1-x_2)}\right]&|N-1:+k;N-1:-k\rangle\Big\}\end{aligned} \tag{31}$$

The squared norm of this state vector yields the probability:

$$\mathcal{P}(x_2,x_1) = \frac{\Delta^2}{L^2}\left\{2N(N-1) + 4N^2\cos^2\left[k(x_2-x_1)\right]\right\} \tag{32}$$

The presence of the cosine of $k(x_2-x_1)$ reveals the existence of a spatial dependence, contrary to what happened for $\mathcal{P}(x_1)$: once a first particle is detected at x_1, the most probable positions x_2 for the second detection are those for which $k(x_2-x_1)$ is a multiple of π. In other words, fringes appear in the double density measurement.

2-d. Discussion

One may wonder which objects interfere in the double counting signal. They are not waves but transition amplitudes associated with two different paths leading the system from the initial state (23) to the same final state $|N-1,+k;N-1,-k\rangle$, where each of the two modes has lost one particle. In the above computation, the first path corresponds to the term $\left(e^{-ikx_2}a_{-k}\right)\left(e^{ikx_1}a_k\right)$, where one particle with momentum $+k$ disappears as it is detected at x_1 and the $-k$ particle disappears as it is detected at x_2; the second path corresponds to the term $\left(e^{ikx_2}a_k\right)\left(e^{-ikx_1}a_{-k}\right)$ where it is now the

particle with momentum $-k$ that disappears as it is detected at x_1 and the $+k$ particle that disappears as it is detected at x_2.

The double counting signal observed on a double condensate is very similar to the double photodetection signal obtained, in § 2-c-γ of Complement E_{XX}, in the study of a product of two one-photon wave packets. In both cases, the signal spatial dependence comes from a quantum interference between the amplitudes of two different paths between the same initial and final states. The difference between the paths comes from a different "switching" between one of the two components of the initial state and one of the two components of the final state.

3. Detection of a large number Q of particles

We now extend the previous reasoning to the case where any number Q of particles' position measurements are performed; we shall limit ourselves to the case where Q remains much smaller than the total particle number N of each condensate.

3-a. Probability of a multiple detection sequence

Generalizing relation (22) allows writing the probability $\mathcal{P}(x_Q, .., x_2, x_1)$ for detecting a particle at x_1, a particle at x_2, .. a particle at x_Q, in the form:

$$\mathcal{P}(x_Q, .., x_2, x_1) = \Delta^Q \ \langle N : +k; N : -k| \, \Psi^\dagger(x_1)\Psi^\dagger(x_2).. \, \Psi^\dagger(x_Q) \\ \Psi(x_Q).. \, \Psi(x_2)\Psi(x_1) \, |N : +k; N : -k\rangle \tag{33}$$

As before, we use relations (24) to replace the field operators and their adjoints by linear combinations of annihilation operators a_k, a_{-k} and creation operators $a_k^\dagger$, $a_{-k}^\dagger$. We then get:

$$\mathcal{P}(x_Q, ..., x_2, x_1) = \frac{\Delta^Q}{L^Q} \ \langle N : +k; N : -k| \, (a_k^\dagger e^{-ikx_1} + a_{-k}^\dagger e^{ikx_1})... \\ ...(a_k^\dagger e^{-ikx_Q} + a_{-k}^\dagger e^{ikx_Q})(a_k e^{ikx_Q} + a_{-k} e^{-ikx_Q})... \\ ... \, (a_k e^{ikx_1} + a_{-k} e^{-ikx_1}) \, |N : +k; N : -k\rangle \tag{34}$$

α. Simplifying hypothesis

When several annihilation operators act successively on the right on the initial ket, each of them introduces a varying factor $\sqrt{N-p}$; this factor depends on the number p of particles already annihilated by the other operators. In the same way, when the creation operators act on the left on the initial bra, they also introduce varying factors. To keep things simple, we shall ignore these variations, assuming that the total detection number Q is always very small compared to the total particle number N in each individual state:

$$Q \ll N \tag{35}$$

One can then replace all the factors $\sqrt{N-p}$ by $\sqrt{N}$:

$$\sqrt{N-p} \simeq \sqrt{N} \tag{36}$$

Apart from multiplication by a fixed factor $\sqrt{N}$, the only effect of each operator is to vary by one unit the occupation number; this result does not depend on the previous actions undergone by the state vector (all the operators commute, once the above approximation has been made). One can then freely move the annihilation and creation operators in the product of operators appearing in (34). Regrouping all the operators associated with the same values of x_i, we get the operator:

$$\begin{aligned} D(x_j) &= (a_k^\dagger e^{-ikx_j} + a_{-k}^\dagger e^{ikx_j})(a_k e^{ikx_j} + a_{-k} e^{-ikx_j}) \\ &= \left[a_k^\dagger a_k + a_{-k}^\dagger a_{-k} + a_k^\dagger a_{-k} e^{-2ikx_j} + a_{-k}^\dagger a_k e^{2ikx_j} \right] \end{aligned} \tag{37}$$

and expression (34) becomes:

$$\mathcal{P}(x_Q, ...x_2, x_1) = \frac{\Delta^Q}{L^Q} \ \langle N:+k; N:-k| \prod_{j=1}^{Q} D(x_j) \, |N:+k; N:-k\rangle \tag{38}$$

We are then left with the computation of the average value in the initial state of a product of operators $D(x_j)$. Expanding each of them according to the second line of (37), we get the sum of 4^Q products, most of them having, nevertheless, a zero average value in the double Fock state $|\Phi_0\rangle$. This is because the only products having a non-zero average value are those for which the repeated effect of the annihilation operators a_k is exactly balanced by the effect of an equal number of operators $a_k^\dagger$ (the particle number in the individual $-k$ state is then also constant, since the total number of particles must be conserved).

Consider then one of those non-zero products. Still using approximation (36), the contribution of each operator $D(x_j)$ will be one of the three factors $NF_{q_j}(x_j)$, with $q_j = 0, \pm 1$ and:

$$\begin{aligned} F_0(x_j) &= 2 \\ F_{\pm 1}(x_j) &= e^{\pm 2ikx_j} \end{aligned} \tag{39}$$

The contribution of F_0 leaves the particle numbers unchanged, the contribution of F_{+1} replaces a $+k$ particle by a $-k$ particle, and finally that of F_{-1} performs the opposite substitution. Relation (38) then becomes:

$$P(x_Q, ..., x_2, x_1) = \frac{\Delta^Q}{L^Q} \ \prod_{j=1}^{Q} \Big[\sum_{q_j = 0, \pm 1} NF_{q_j}(x_j) \Big] \qquad \text{with } S = 0 \tag{40}$$

where, when we expand the product in the right hand side of the equation, we retain only the terms for which the sum of all q_j vanishes:

$$S = \sum_{j=1}^{Q} q_j = 0 \tag{41}$$

This constraint simply expresses the conservation of the particle number in each individual state.

β. Simple expression for the multiple detection probability

An easy way to impose constraint (41) on all the q_j is to introduce a Kronecker delta $\delta_{S,0}$, and write:

$$P(x_Q, ..x_2, x_1) = \frac{\Delta^Q}{L^Q} \sum_{q_1, q_2, .., q_Q} \delta_{S,0} \prod_{j=1}^{Q} N F_{q_j}(x_j) \tag{42}$$

where the summation over the q_j is now free of constraint. We can then use relation:

$$\int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi}\, e^{iS\xi} = \delta_{S,0} \tag{43}$$

which amounts to multiplying each term $F_{q_j}(x_j)$ in (40) by $\exp(iq_j\xi)$ and summing over $\mathrm{d}\xi$. We then get:

$$P(x_Q, ..x_2, x_1) = \frac{\Delta^Q}{L^Q} \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \prod_{j=1}^{Q} \sum_{q_j} N e^{iq_j\xi} F_{q_j}(x_j) \tag{44}$$

Each summation over q_j then yields the quantity:

$$\begin{aligned} &N\left[F_0(x_j) + F_+(x_j)e^{i\xi} + F_-(x_j)e^{-i\xi}\right] \\ &\qquad = N\left[2 + \exp(2ikx_j + i\xi) + \exp(-2ikx_j - i\xi)\right] \end{aligned} \tag{45}$$

which can be written as:

$$2N\left[1 + \cos(2kx_j + \xi)\right] = 4N\cos^2\left(kx_j + \frac{\xi}{2}\right) \tag{46}$$

Finally, we obtain the following simple analytical expression for the multiple detection rate:

$$P(x_Q, ..x_2, x_1) = \left(\frac{4\Delta N}{L}\right)^Q \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \prod_{j=1}^{Q} \cos^2\left(kx_j + \frac{\xi}{2}\right) \tag{47}$$

3-b. Discussion; emergence of a relative phase

Relation (47) allows understanding how the successive measurements enable the progressive emergence of a relative phase.

α. Detecting the first particles

Let us start with the very first detection at $x = x_1$. Equation (47) yields the probability of such an event:

$$P(x_1) \propto \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \cos^2(kx_1 + \frac{\xi}{2}) \tag{48}$$

The term in cosine appearing in the integral yields the fringes that one expects from the interference between two waves, with wave vectors $+k$ and $-k$ along the x axis, and a

phase shift ξ. However, the summation over ξ indicates that the interference pattern must be averaged over all the possible values of ξ, uniformly distributed between 0 and 2π: this means that the fringes are completely blurred out.

The double detection rate at x_1 and x_2 is obtained by keeping the terms $j = 1$ and $j = 2$ in (47). It is equal to:

$$P(x_2, x_1) \propto \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \cos^2(kx_2 + \frac{\xi}{2}) \cos^2(kx_1 + \frac{\xi}{2}) \tag{49}$$

As the first detection has already occurred, x_1 is fixed in this equation, and the product of the two cosines yields the probability of finding the second particle at x_2. But the integral over $\mathrm{d}\xi$, which yields the x_2 dependence of the probability, is no longer over a phase ξ uniformly distributed between 0 and 2π, because of the presence of the $\cos^2(kx_1 + \xi/2)$ associated with the first detection; the blurring of the fringes is not as radical as before. For this second detection, the function $\cos^2(kx_1 + \xi/2)$ actually plays the role of an x_1 dependent phase distribution; the two detections are no longer independent. This confirms the qualitative discussion of § 2.

This mechanism can be generalized to higher order measurements. As an example, the triple detection rate at x_1, x_2 and x_3 is equal to:

$$P(x_3, x_2, x_1) \propto \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \cos^2(kx_3 + \frac{\xi}{2}) \cos^2(kx_2 + \frac{\xi}{2}) \cos^2(kx_1 + \frac{\xi}{2}) \tag{50}$$

Once the first two detections have been made at x_1 and x_2, the relative phase distribution that comes into play for the third detection is the product of two cosine functions $\cos^2(kx_2+\xi/2) \cos^2(kx_1+\xi/2)$ – and no longer a single one as was the case before. As the product of two cosine functions yields a sharper curve than a single cosine function, the relative phase is better defined for the third detection than for the second. The process continues in the same way with the following detections, and the phase is more and more precisely defined. This means that it is the first detections that determine the positions of the fringes appearing in the following detections, each of them contributing to a more and more precise definition of the relative phase distribution.

This argument is of course only valid for a given experiment. If one performs a new experiment with the same experimental conditions, the first detections will not, in general, happen at the same places as in the first experiment. Consequently, after a large number of detections, a fringe pattern will appear, shifted with respect to the pattern observed in the first experiment. Finally, if one adds up the positions measured in a large number of successive experiments, the fringes average to practically zero, and one gets a quasi-uniform position distribution.

β. Emergence of a well-defined relative phase after a large number of detections

After a large number of detections, Q, the relative phase distribution for the $(Q + 1)^{th}$ detection is given by the product of a large number Q of cosine functions, yielding a very narrow phase distribution, centered at a value ξ_M. One can then replace in (47) all the $[1 + \cos(2kx_j + \xi)]$ by $[1 + \cos(2kx_j + \xi_M)]$, so that the probability becomes a product: the detections are now independent, the interference pattern becomes stable with a sharper and sharper contrast. These predictions have been confirmed by numerical simulations based on equation (47).

Narrowing of the relative phase distribution

The narrowing of the relative phase distribution can be explained by an analytical calculation. Let us assume that after Q detections this distribution can be approximated by a Gaussian curve centered at ξ_M, and with a width σ_Q:

$$W_Q(\xi) \simeq e^{-(\xi-\xi_M)^2/\sigma_Q^2} \tag{51}$$

After the $(Q+1)^{th}$ detection, the new distribution will be given by:

$$W_{Q+1}(\xi) \simeq e^{-(\xi-\xi_M)^2/\sigma_Q^2}\,[1+\cos(2kx_{Q+1}+\xi)] \tag{52}$$

As the function $1+\cos(2kx_{Q+1}+\xi)$ is much broader than $W_{Q+1}(\xi)$, it can be expanded in powers of $\xi-\xi_M$ in the vicinity of $\xi=\xi_M$ where the distribution $W_Q(\xi)$ takes on significant values:

$$\begin{aligned}1+\cos(2kx_{Q+1}+\xi) &= 1+\cos(2kx_{Q+1}+\xi_M)\\ &-(\xi-\xi_M)\sin(2kx_{Q+1}+\xi_M)-\frac{1}{2}(\xi-\xi_M)^2\cos(2kx_{Q+1}+\xi_M)\end{aligned} \tag{53}$$

One can also expand $W_Q(\xi)$ in the vicinity of $\xi=\xi_M$:

$$e^{-(\xi-\xi_M)^2/\sigma_Q^2} = 1-\frac{(\xi-\xi_M)^2}{\sigma_Q^2} \tag{54}$$

We then multiply (53) and (54) and obtain an expansion for $W_{Q+1}(\xi)$:

$$\begin{aligned}W_{Q+1}(\xi) &\simeq 1+\cos(2kx_{Q+1}+\xi_M)-(\xi-\xi_M)\sin(2kx_{Q+1}+\xi_M)\\ &-(\xi-\xi_M)^2\left[\frac{1}{2}\cos(2kx_{Q+1}+\xi_M)+\frac{1}{\sigma_Q^2}\Big(1+\cos(2kx_{Q+1}+\xi_M)\Big)\right]\end{aligned} \tag{55}$$

We note that $W_{Q+1}(\xi)$ depends on the position x_{Q+1} of the $(Q+1)^{th}$ detection. We can obtain an average value for $W_{Q+1}(\xi)$ by weighting $W_{Q+1}(\xi)$ by the probability $[1+\cos(2kx_{Q+1}+\xi_M)]$ for the $(Q+1)^{th}$ detection to occur at $x=x_{Q+1}$, and integrate x_{Q+1} over a spatial period of the interference pattern:

$$\overline{W_{Q+1}(\xi)} = \int_0^{2\pi/2k} \mathrm{d}x_{Q+1}\,[1+\cos(2kx_{Q+1}+\xi_M)]\,W_{Q+1}(\xi). \tag{56}$$

Since:

$$\begin{aligned}\overline{\cos(2kx_{Q+1}+\xi_M)} &= \overline{\sin(2kx_{Q+1}+\xi_M)}\\ &= \overline{\cos(2kx_{Q+1}+\xi_M)\sin(2kx_{Q+1}+\xi_M)} = 0\end{aligned} \tag{57}$$

and:

$$\overline{\cos^2(2kx_{Q+1}+\xi_M)} = \frac{1}{2} \tag{58}$$

we finally obtain:

$$\overline{W_{Q+1}(\xi)} \simeq \frac{3}{2}\left[1-(\xi-\xi_M)^2\left(\frac{1}{\sigma_Q^2}+\frac{1}{6}\right)\right] \simeq \frac{3}{2}\,e^{-(\xi-\xi_M)^2)/\sigma_{Q+1}^2} \tag{59}$$

where:

$$\frac{1}{\sigma_{Q+1}^2} = \frac{1}{\sigma_Q^2} + \frac{1}{6}. \tag{60}$$

Equation (60) shows that $\sigma_{Q+1} < \sigma_Q$, meaning that the distribution curve becomes narrower after each detection. One can easily iterate equation (60) to obtain:

$$\frac{1}{\sigma_{Q+n}^2} = \frac{1}{\sigma_Q^2} + \frac{n}{6} \tag{61}$$

where n is a positive integer. This shows that if $n \gg 1/\sigma_Q^2$, the width of the relative phase distribution decreases as $1/\sqrt{n}$.

A similar computation can be made to study the position of $W_{Q+1}(\xi)$'s maximum when Q increases. One finds that the center of the relative phase distribution is shifted by a quantity proportional to $1/n$.

Finally, it is interesting to note the link between the uncertainty on the relative phase (which decreases as the detection number Q increases) and the uncertainty on the difference $N_+ - N_-$ between the numbers $N_\pm$ of particles in the $\pm k$ condensates (which, on the contrary, increases). At the beginning of the experiment (before the first detection), we have $N_+ = N_- = N$. After the first detection, we saw in § 2 that the state of the system contains a linear superposition of states $N_+ = N \pm 1$ and $N_- = N \mp 1$: the difference $N_+ - N_-$ between N_+ and N_- is no longer fixed and equal to zero, but can take on several values 0, ± 2. After the second detection, the state of the system contains a superposition of states having always the same value of $N_+ + N_-$, but values of $N_+ - N_-$ that can be equal to 0, ± 2, .. and so on. After Q detections, the values of $N_+ - N_-$ spread out between $-2Q$ and $+2Q$. This result is an illustration of the fact that the relative phase and the difference between the particle numbers between the two condensates are conjugate quantities.

Conclusion

This complement illustrates how successive measurements on a system having components on two individual states, each with N particles, can build up (from zero) a relative phase between these components; for this to happen, the measurements must depend on the relative phase between those two components. Mathematically, relation (47) shows that the results obtained for an ensemble of Q position measurements (with $Q \ll N$) are exactly the same as if an initial well defined phase ξ had existed from the beginning of the experiment, even though it was totally unknown (and could have taken on any value uniformly distributed between 0 and 2π). The measurements did indeed introduce entanglement and its associated relative phase, but the quantum predictions are equivalent to those obtained by assuming that the measurements only reveal a preexisting phase (as in quantum theories with so-called "additional variables").

The process we have discussed is, however, of an essentially different nature: it is indeed each individual measurement that contributes to a better and better definition of the relative phase for the measurements to come; these will occur at points x whose probability distributions depend on the results of all the previously performed measurements. We shall see in Complement D_{XXI} that if, instead of measuring a fraction of

all the particles, they each undergo a measurement, the phase properties can no longer be understood as that of a classical preexisting (but unknown) phase; these properties clearly become quantum, as shown by the possibility of violations of Bell's inequalities.

Complement D_{XXI}

Emergence of a relative phase with spin condensates; macroscopic non-locality and the EPR argument

1 Two condensates with spins 2254

1-a Spin 1/2: a brief review 2254

1-b Projectors associated with the measurements 2255

2 Probabilities of the different measurement results 2255

2-a A first expression for the probability 2256

2-b Introduction of the phase and the quantum angle 2258

3 Discussion 2259

3-a Measurement number $Q \ll 2N$ 2259

3-b Macroscopic EPR argument 2261

3-c Measuring all the spins, violation of Bell's inequalities 2263

This complement continues the discussion of Complement C_{XXI} on the measurement induced emergence of a relative phase between condensates, but in a more general case. We established in C_{XXI} that as more and more measurement results are obtained, their number still remaining smaller than the total particle number, the relative phase of the two condensates becomes better defined. It soon reaches a classical regime where it is (almost) perfectly determined. This necessarily comes with large fluctuations of the numbers of particles occupying the two individual states (or more precisely of their difference), as required by the uncertainty relation between phase and occupation numbers. In the present complement, a first important difference is that we no longer assume that the number of measurements remains small compared to the total particle number. This will enable us to follow the evolution of the phase properties during the whole series of measurements, including the last moments when the number of particles remaining to be measured is just a few units. For these few remaining particles, the fluctuations in the difference of the occupation numbers is necessarily limited to a few units, meaning that the phase can no longer be precisely determined. The phase then comes back to a quantum regime, where one can no longer interpret the measurement results in a classical context (preexisting but totally unknown phase). Another difference with Complement C_{XXI} is that we now assume the two condensates correspond to different individual spin states. Instead of position measurements yielding continuous results, we can now perform measurements on the spin directions, which yield discrete results. This will make it easier to discuss the quantum effects, which can lead to violations of Bell's inequalities (Chapter XXI, § F-3). Another advantage of dealing with spins is that we can go back to the EPR argument (Chapter XXI, § F-1) in a case where the elements of reality, introduced by EPR, are macroscopic and have, in addition, a simple physical interpretation (spin angular momentum).

1. Two condensates with spins

We now assume that the two individual states populated in the condensates are the two states $|u,\pm\rangle$ corresponding to two different internal states noted $|\pm\rangle$, but to the same orbital state $|u\rangle$:

$$|u,\pm\rangle = |u\rangle \otimes |\pm\rangle \tag{1}$$

If $(a_{u,+})^\dagger$ and $(a_{u,-})^\dagger$ are the creation operators associated with these states, the state $|\Phi_0\rangle$ of the system formed by the juxtaposition of the two condensates can be written as:

$$|\Phi_0\rangle = \frac{1}{N!}\left[(a_{u,+})^\dagger\right]^N \quad \left[(a_{u,-})^\dagger\right]^N \quad |0\rangle \tag{2}$$

which replaces relation (23) of Complement C_{XXI}; the total particle number is $2N$.

By commodity, we will often call "spin states" the two states $|\pm\rangle$, and reason as if they were indeed the two accessible states of a spin-1/2 particle. This is just a manner of speaking: according to the spin-statistic theorem (Chapter XIV, § C-1), bosons cannot be half-integer spin particles. The system we consider is actually an ensemble of bosons that have access to only two internal states; these can be, for example, the two $m = 0$ and $m = 1$ states of a spin equal to 1, or not necessarily spin states.

1-a. Spin 1/2: a brief review

For the reasoning that follows, it may be useful to recall a few relations (Chapter IV, § A-2) concerning a spin 1/2 (with no orbital variables). As pointed out above, we are dealing with a fictitious spin, whose operators act on any two internal states, noted $|\pm\rangle$ by pure commodity. Operator σ_z, associated with the first Pauli matrix (Complement A_{IV}), is the difference between the projector onto the state $|+\rangle$ and the projector onto the state $|-\rangle$:

$$\sigma_z = |+\rangle\langle+| \;-\; |-\rangle\langle-| \tag{3}$$

whereas operators σ_x and σ_y are expressed as a function of the two non-diagonal operators $|+\rangle\langle-|$ and $|-\rangle\langle+|$ as:

$$\begin{aligned}\sigma_x &= \;\; |+\rangle\langle-| \;+\; |-\rangle\langle+| \\ \sigma_y &= -i\,|+\rangle\langle-| \;+\; i\,|-\rangle\langle+|\end{aligned} \tag{4}$$

As for the fictitious spin component along a direction $O\varphi$ in the xOy plane, making an angle φ with the Ox axis, the corresponding operator is written:

$$\sigma_\varphi = \cos\varphi\;\sigma_x + \sin\varphi\;\sigma_y = e^{-i\varphi}\,|+\rangle\langle-| \;+\; e^{i\varphi}\,|-\rangle\langle+| \tag{5}$$

Its eigenvalues are $\eta = \pm 1$, and its eigenvectors can be expressed as:

$$\begin{aligned}|\psi_{\eta=+1}\rangle &= \frac{1}{\sqrt{2}}\left[e^{-i\varphi/2}\,|+\rangle + e^{i\varphi/2}\,|-\rangle\right] \\ |\psi_{\eta=-1}\rangle &= \frac{1}{\sqrt{2}}\left[-e^{-i\varphi/2}\,|+\rangle + e^{i\varphi/2}\,|-\rangle\right]\end{aligned} \tag{6}$$

as can be easily checked by applying operator (5) to these relations. The projector onto the ket with eigenvalue η can thus be written:

$$\begin{aligned} P_\eta(\varphi) &= \frac{1}{2}\left[1+\eta\sigma_\varphi\right] \\ &= \frac{1}{2}\left[1+\eta\left(e^{-i\varphi}\left|+\right\rangle\left\langle-\right| \;+\; e^{i\varphi}\left|-\right\rangle\left\langle+\right|\right)\right] \end{aligned} \tag{7}$$

1-b. Projectors associated with the measurements

For an ensemble of identical bosons having orbital variables, we note $\Psi_\pm(\mathbf{r})$ the two field operators associated with the two internal states $\pm$. The operator associated with the total particle density at point $\mathbf{r}$ is the sum of the local densities, $\Psi_+^\dagger(\mathbf{r})\Psi_+(\mathbf{r})$ corresponding to the + spin state, and $\Psi_-^\dagger(\mathbf{r})\Psi_-(\mathbf{r})$ corresponding to the − state:

$$n(\mathbf{r}) = \Psi_+^\dagger(\mathbf{r})\Psi_+(\mathbf{r}) + \Psi_-^\dagger(\mathbf{r})\Psi_-(\mathbf{r}) \tag{8}$$

As for the operator associated with the σ_z spin component along the Oz quantization axis, relation (3) indicates that it is the difference:

$$\sigma_z(\mathbf{r}) = \;\Psi_+^\dagger(\mathbf{r})\Psi_+(\mathbf{r}) - \Psi_-^\dagger(\mathbf{r})\Psi_-(\mathbf{r}) \tag{9}$$

According to (5), the operator associated with a measurement performed along a direction $O\varphi$ of the xOy plane making an angle φ with the Ox axis is written as:

$$\sigma_\varphi(\mathbf{r}) = e^{-i\varphi}\Psi_+^\dagger(\mathbf{r})\Psi_-(\mathbf{r}) + \;e^{i\varphi}\Psi_-^\dagger(\mathbf{r})\Psi_+(\mathbf{r}) \tag{10}$$

The measurements we are interested in pertain both to the position of the particles and their spin: for each measurement, the position is measured in an infinitesimal volume Δ centered at point $\mathbf{r}_j$ and, when measuring the direction of the spin along the $O\varphi_j$ direction, we obtain the result $\eta_j = \pm 1$. By analogy with (7), the projector associated with such a measurement can be written:

$$\begin{aligned} P_{\eta_j}(\mathbf{r}_j,\varphi_j) &= \frac{1}{2}\int_\Delta \mathrm{d}^3r' \;\left[n(\mathbf{r}') + \eta_j\;\sigma_{\varphi_j}(\mathbf{r}')\right] \\ &\simeq \frac{\Delta}{2}\left[n(\mathbf{r}_j) + \eta_j\;\sigma_{\varphi_j}(\mathbf{r}_j)\right] \end{aligned} \tag{11}$$

where $n(\mathbf{r})$ and $\sigma_\varphi(\mathbf{r})$ are given by (8) and (10). Operator $P_{\eta_j}(\mathbf{r}_j,\varphi_j)$ projects both the orbital variables onto this small domain and the spins onto the eigenstate of the component along the $O\varphi_j$ axis, with eigenvalue $\eta_j = \pm 1$.

2. Probabilities of the different measurement results

Consider now a system of $2N$ bosons in the state (2). Spin measurements are performed in a series of regions of space, which cover the whole extension of the orbital wave function $u(\mathbf{r})$ without overlapping. The measurements are supposed to be ideal, so that every particle is detected. The regions are supposed to be sufficiently small to obtain a negligible probability of double detection in any of them. Those where a particle is actually detected are centered at $\mathbf{r}_j$ (with $j = 1, 2, .., 2N$), as illustrated in Figure 1.

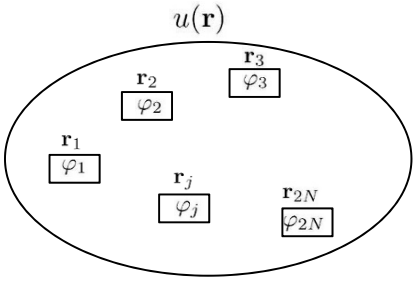


Figure 1: Two condensates, each having N particles, the first one with + spins, and the other with − spins, share the same orbital wave function $u(\mathbf{r})$, represented by the oval in the figure. Measurements of the transverse direction of the spins are performed in $2N$ non overlapping regions of space, centered at points $\mathbf{r}_j$ (with $j = 1, 2, .., 2N$). In each region, the measurement is performed along a transverse direction (perpendicular to the quantization axis), defined by an angle φ_j, which may depend on j; the corresponding result is $\eta_j = \pm 1$.

In each of these regions, one performs a measurement of the spin component along a transverse direction (perpendicular to the quantization axis), defined by the angle φ_j, and the measurement result is $\eta_j = \pm 1$. We now calculate the probability of getting a series of results $\eta_j = \pm 1$ in those $2N$ regions.

2-a. A first expression for the probability

The associated projectors $P_{\eta_j}(\mathbf{r}_j, \varphi_j)$ all commute with each other, since they contain field operators (and their adjoints) at different points in space (we assume that all measurements are done simultanously, or separated by a very short time). The probability $\mathcal{P}_{2N}$ of a result is therefore the average value in the state $|\Phi_0\rangle$ of the product of projectors:

$$\begin{aligned}\mathcal{P}_{2N}\left(\{\eta_i,\varphi_i\}\right) &= \langle\Phi_0| \prod_{j=1}^{2N} P_{\eta_j}(\mathbf{r}_j,\varphi_j)\,|\Phi_0\rangle \\ &= \left(\frac{\Delta}{2}\right)^{2N} \langle\Phi_0| \prod_{j=1}^{2N} \Big\{ \Psi_+^\dagger(\mathbf{r}_j)\Psi_+(\mathbf{r}_j) + \Psi_-^\dagger(\mathbf{r}_j)\Psi_-(\mathbf{r}_j) + \eta_j e^{-i\varphi_j}\Psi_+^\dagger(\mathbf{r}_j)\Psi_-(\mathbf{r}_j) \\ &\qquad + \eta_j e^{i\varphi_j}\Psi_-^\dagger(\mathbf{r}_j)\Psi_+(\mathbf{r}_j)\Big\}\,|\Phi_0\rangle \end{aligned} \tag{12}$$

where $\{\eta_i, \varphi_i\}$ symbolizes the ensemble of the variables $(\eta_1, \varphi_1, .., \eta_N, \varphi_N)$. As the operators commute, we can also move all the field operators $\Psi_\pm(\mathbf{r}_j)$ towards the right, and their adjoints $\Psi_\pm^\dagger(\mathbf{r}_j)$ towards the left. We now introduce a basis $u_i(\mathbf{r})$ for the wave

functions, its first element being the wave function of the populated states (1). This basis allows expanding the field operators according to relation (A-14) of Chapter XVI, and we can write:

$$\Psi_{\pm}(\mathbf{r}_j) = \sum_i u_i(\mathbf{r}_j)\ a_{u_i,\pm} \tag{13}$$

where $a_{u_i,\pm}$ is the annihilation operator of a particle in the individual state $|u_i,\pm\rangle$. Now the only term that actually plays a role in this expansion is the $i=1$ term: all the other $i \neq 1$ terms yield zero when acting on the state $|\Phi_0\rangle$, which only contains particles in the orbital state $u_1(\mathbf{r}) = u(\mathbf{r})$. One can simply replace $\Psi_{\pm}(\mathbf{r}_j)$ by $u(\mathbf{r}_j)\, a_{u,\pm}$. The same is true for the adjoint of the field operators which, acting on the bra placed on their left, can only destroy particles in a state previously populated; consequently, we can also replace $\Psi^{\dagger}_{\pm}(\mathbf{r}_j)$ by $u^*(\mathbf{r}_j)\, a^{\dagger}_{u,\pm}$. Once these replacements have been performed, we get an expression that can be written, in a symbolic way, as:

$$\begin{aligned}\mathcal{P}_{2N}(\{\eta_i,\varphi_i\}) = \\ \left(\frac{\Delta}{2}\right)^{2N} \langle\Phi_0| : \prod_{j=1}^{2N} |u(\mathbf{r}_j)|^2 \Big[a^{\dagger}_{u,+}a_{u,+} + a^{\dagger}_{u,-}a_{u,-} \\ +\eta_j e^{-i\varphi_j} a^{\dagger}_{u,+}a_{u,-} + \eta_j e^{i\varphi_j} a^{\dagger}_{u,-}a_{u,+}\Big] : |\Phi_0\rangle\end{aligned} \tag{14}$$

The two dots surrounding the product over j in this symbolic writing express the following convention, which originates from the rearrangement of the operators $\Psi_{\pm}(\mathbf{r}_j)$ and $\Psi^{\dagger}_{\pm}(\mathbf{r}_j)$ mentioned above: in each of the 4^{2N} terms of the product of sums, all the annihilation operators $a_{u,\pm}$ are regrouped towards the right, and all the creation operators $a^{\dagger}_{u,\pm}$ towards the left. To obtain the probability we are looking for, we have to compute the average values in the state $|\Phi_0\rangle$ of 4^{2N} products, in normal order (Complement $\mathrm{B_{XVI}}$, § 1-a-α), of creation and annihilation operators in the two states $|u_{i=1},\pm\rangle$.

The situation now becomes very similar to that leading to relation (37) in Complement $\mathrm{C_{XXI}}$. The computation that follows is indeed similar, except for the fact that we no longer use the approximation (35) of that complement (number of measurements small compared to the particle number): we now assume that all the particles are measured. Actually, most of the terms of the product over j appearing in (14) have a zero average value in the state $|\Phi_0\rangle$. The only relevant terms are those that contain exactly N annihilation operators $a_{u,+}$ and N other annihilation operators $a_{u,-}$, in which case their action yields the vacuum, hence a normalized ket; if this is not the case, the result is zero. For a similar reason, they must also contain exactly N creation operators $a^{\dagger}_{u,+}$ and N other $a^{\dagger}_{u,-}$, otherwise the result is zero. All these non-zero terms have the same average values, since the product of operators in the normal order introduces each time the same factor $\left(\sqrt{N!}\times\sqrt{N!}\right)^2$, that is $(N!)^2$; we also get the product of $2N$ coefficients F_{q_j,q'_j}, which can take one of these 4 values:

$$F_{+1,+1} = F_{-1,-1} = 1 \tag{15}$$

and:

$$F_{+1,-1} = \eta_j e^{-i\varphi_j} \qquad F_{-1,+1} = \eta_j e^{i\varphi_j} \tag{16}$$

Now $F_{+1,+1}$ corresponds to a term associated with a particle destruction in $|u,+\rangle$, followed by a particle creation in that same state. In the same way, $F_{-1,-1}$ corresponds to an annihilation-creation in the individual state $|u,-\rangle$. Finally, $F_{+1,-1}$ corresponds to an annihilation in state $|u,-\rangle$ followed by a creation in state $|u,+\rangle$, and the opposite for $F_{-1,+1}$. All the non-zero terms therefore correspond to products of $2N$ numbers F_{q_j,q'_j} such that the sum of the q_i and the sum of the q'_j are both zero; this condition automatically ensures the conservation of the particle number in each individual state. The final result is:

$$\mathcal{P}_{2N}\left(\{\eta_i,\varphi_i\}\right) = \left(\frac{\Delta}{2}\right)^{2N} (N!)^2 \prod_{j=1}^{2N} \Big[\, |u(\mathbf{r}_j)|^2 \sum_{q_j,q'_j=\pm 1} F_{q_j,q'_j} \Big] \quad \text{with } S=S'=0 \tag{17}$$

where, in the right hand side, we retain only the terms satisfying the double condition:

$$\begin{aligned} &\sum_j q_j = S = 0 \\ &\sum_j q'_j = S' = 0 \end{aligned} \tag{18}$$

2-b. Introduction of the phase and the quantum angle

Because of the summation constraints, expression (17) is not easy to handle. This is why we introduce two delta functions $\delta_{S,0}$ and $\delta_{S',0}$, which obey:

$$\begin{aligned} \delta_{S,0} &= \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi}{2\pi}\, e^{iS\xi} \\ \delta_{S',0} &= \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi'}{2\pi}\, e^{iS'\xi'} \end{aligned} \tag{19}$$

This amounts to multiplying in (17) each F_{q_j,q'_j} by $e^{i\left(q_j\xi+q'_j\xi'\right)}$ and integrating over the two angles ξ and ξ'. This enables us to write the probability in the form:

$$\begin{aligned} &\mathcal{P}_{2N}\left(\{\eta_i,\varphi_i\}\right) = \\ &\quad \left(\frac{\Delta}{2}\right)^{2N} (N!)^2 \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi}{2\pi} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi'}{2\pi} \prod_{j=1}^{2N} |u(\mathbf{r}_j)|^2 \sum_{q_j,q'_j} e^{i\left(q_j\xi+q'_j\xi'\right)} F_{q_j,q'_j} \end{aligned} \tag{20}$$

where the summations over q_j and q'_j are now independent, thanks to relations (19) that automatically ensure the constraints are obeyed. For each value of j, each sum contributes the factor:

$$\begin{aligned} &F_{+1,+1}\, e^{i(\xi+\xi')} + F_{-1,-1}\, e^{-i(\xi+\xi')} + F_{+1,-1}\, e^{i(\xi-\xi')} + F_{-1,+1}\, e^{i(\xi'-\xi)} \\ &\qquad = 2\cos\left(\xi+\xi'\right) + 2\eta_j \cos\left(\xi-\xi'-\varphi_j\right) \end{aligned} \tag{21}$$

We finally make the change of variables[1]:

$$\begin{aligned}\Lambda &= \xi + \xi' \\ \lambda &= \xi - \xi'\end{aligned} \tag{22}$$

This leads to the simpler expression:

$$\mathcal{P}_{2N}\left(\{\eta_i, \varphi_i\}\right) = \\ (\Delta)^{2N}\,(N!)^2 \int_{-\pi}^{+\pi} \frac{\mathrm{d}\Lambda}{2\pi} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \prod_{j=1}^{2N} \left|u(\mathbf{r}_j)\right|^2 \left[\cos\Lambda + \eta_j \cos\left(\lambda - \varphi_j\right)\right] \tag{23}$$

The following discussion is entirely based on this result. For reasons that will be explained below, λ is called the phase, whereas Λ is called the "quantum angle".

Comment:

It will be useful in what follows to note that the right-hand side of the above equality stays the same if we make the change:

$$\int_{-\pi}^{+\pi} \frac{\mathrm{d}\Lambda}{2\pi} \Rightarrow 2 \int_{-\pi/2}^{+\pi/2} \frac{\mathrm{d}\Lambda}{2\pi} \tag{24}$$

To show this, we can decompose the integral over $\mathrm{d}\Lambda$ in a sum $I = I_1 + I_2$, where I_1 is the integral between $-\pi/2$ and $+\pi/2$, and I_2, the integral between $\pi/2$ and $3\pi/2$ (as the period of the function to be integrated is 2π, any integration domain covering the entire circle is equivalent). Now I_2 is just equal to I_1. This is because the function to be integrated is multiplied by $(-1)^{2N} = 1$ when one changes Λ to $\Lambda' = \Lambda - \pi$ as well as λ to $\lambda' = \lambda - \pi$. Consequently, changing the integration variables Λ, λ to Λ', λ' allows giving to I_2 the same integration domain as I_1.

3. Discussion

Let us examine first the case where the number of measurements is negligible compared to the particle number in each condensate; this will enable us to compare the results obtained with those of Complement $\mathrm{C_{XXI}}$.

3-a. Measurement number $Q \ll 2N$

We first recall a general property of quantum mechanics concerning compatible observables (Chapter III, § C-6-a). When several operators A, B, C, etc. commute with each other, one can build a basis with their common eigenvectors. The scalar product of these eigenvectors and the system state vector yields the probability amplitude for finding the system in each of these eigenvectors. If the eigenvalues are non-degenerate, the squared modulus of this amplitude yields the probability of finding the corresponding eigenvalues upon a series of simultaneous measurements associated with all the operators

[1]The Jacobian of this change of variables is equal to 2, and this factor should be introduced in the denominator. Nevertheless, since the integrated function is periodic, this factor can be taken into account by reducing the integration domain of the two variables λ and Λ to the interval $-\pi$, $+\pi$, which reduces the area of the integration domain by a factor 2.

A, B, C, etc. If they are degenerate, we just have to sum the probabilities over all the orthogonal eigenkets. This is the rule we have followed until now in this complement.

Imagine now that we ignore the measurement results associated with one or several operators of the series, for example B and C; the probabilities of the measurement results we still consider relevant are then simply the sum over the possible results associated with the ignored measurements (sum of the probabilities of exclusive events). One can also imagine another situation where the quantities associated with B and C are actually never measured. The possible series of eigenvalues of the measured operators are then less numerous than in the previous case (since a smaller number of measurements is performed), which increases the degree of degeneracy of these eigenvalues. As for the eigenvalues of B and C, even though they do not correspond to actual measurements, they can still be used as indices to distinguish between the different orthogonal eigenvectors associated with measurement results of operators A, D, etc. Consequently we still have to sum the probabilities on these different eigenvalues, just as we did in the case where these measurements were ignored. This means that quantum mechanics yields the same probabilities whether we assume that the measurement results of B and C are ignored or have never been measured.

Let us now compute the probability $\mathcal{P}_Q$ of obtaining results $\eta_1, .., \eta_Q$ when performing Q measurements on the spins. As we already know the probability (23) corresponding to the case $Q = 2N$, we can consider that all the $2N$ measurements have been performed, but that we ignore the results of $2N-Q$ among them. As we just discussed, this amounts to summing in (23) the probabilities of the two possible results for each of these $2N-Q$ ignored measurements, i.e. the probabilities associated with two opposite values of the η_j. It follows that in the product over j in (23), $\cos(\lambda - \varphi_j)$ will disappear from all the $2N-Q$ terms, leaving only $\cos\Lambda$. We get for $\mathcal{P}_Q$ the following expression (omitting from now on the numerical factors, which are not relevant for our discussion):

$$\mathcal{P}_Q(\{\eta_i, \varphi_i\}) \propto \int_{-\pi/2}^{+\pi/2} \frac{\mathrm{d}\Lambda}{2\pi} [\cos\Lambda]^{2N-Q} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \prod_{j=1}^{Q} |u(\mathbf{r}_j)|^2 [\cos\Lambda + \eta_j \cos(\lambda - \varphi_j)] \tag{25}$$

In this expression, the notation $(\{\eta_i, \varphi_i\})$ now stands for Q pairs of variables (η, φ), instead of $2N$ as before. The integral over $\mathrm{d}\Lambda$ contains the function $\cos\Lambda$ to the power $2N - Q$; if $Q \ll 2N$, this power is very high, and the function becomes a very narrow peak centered at $\Lambda = 0$. This allows us to write:

$$\mathcal{P}_Q(\{\eta_i, \varphi_i\}) \propto \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \prod_{j=1}^{Q} |u(\mathbf{r}_j)|^2 [1 + \eta_j \cos(\lambda - \varphi_j)] \tag{26}$$

This result is very similar to the one obtained in relation (47) of Complement C_{XXI}, namely a product of two positive individual probabilities[2]:

$$p_j(\lambda, \varphi_j) = \frac{1}{2} |u(\mathbf{r}_j)|^2 [1 + \eta_j \cos(\lambda - \varphi_j)] \tag{27}$$

[2]Thanks to the factor 1/2, the sum of the two probabilities ($\eta_j = \pm 1$) is normalized to the probability of presence of a particle in the detection volume.

This product is then averaged over the angle λ, which can take on any value between $-\pi$ and $+\pi$. As we show below, probability (27) is actually the probability of finding the result η_j when measuring the component of a single spin along an axis with direction φ_j, assuming that spin was initially polarized along a direction defined by the angle λ.

Demonstration:

We call Oz the spin quantization axis, and consider a spin polarized in the (transverse) direction in the plane xOy, making an angle λ with the Ox axis. Relation (6) indicates that its state is then:

$$|\psi_\lambda\rangle = \frac{1}{\sqrt{2}}\left[e^{-i\lambda/2}\,|+\rangle + e^{+i\lambda/2}\,|-\rangle\right] \tag{28}$$

When measuring the spin component along an axis defined by the angle φ, the state associated with the $\eta = +1$ measurement result is:

$$|\psi_\varphi\rangle = \frac{1}{\sqrt{2}}\left[e^{-i\varphi/2}\,|+\rangle + e^{+i\varphi/2}\,|-\rangle\right] \tag{29}$$

Consequently, the probability of that result is:

$$\mathcal{P}_{\eta=+1} = |\langle\psi_\varphi|\psi_\lambda\rangle|^2 = \frac{1}{4}\left[e^{i(\lambda-\varphi)/2} + e^{-i(\lambda-\varphi)/2}\right]^2 = \cos^2\frac{\lambda-\varphi}{2} \tag{30}$$

$$= \frac{1}{2}\left[1+\cos(\lambda-\varphi)\right] \tag{31}$$

As for the probability of the $\eta = -1$ result, it is simply the complementary probability, obtained by changing the sign in front of $\cos(\lambda-\varphi)$. In both cases, we get the probability given by (27).

This means that λ can be interpreted as the relative phase between the two condensates. In this case, the predictions of quantum mechanics are identical to the predictions of a theory where the phase would be considered as a classical quantity, perfectly determined but as yet unknown at the beginning of the experiment. From such a point of view, this phase would be revealed more and more precisely by the successive measurements, instead of being created as assumed in the standard quantum mechanics interpretation. Therein lies a link to the heart of the EPR argument.

3-b. Macroscopic EPR argument

The EPR argument was presented in § F-1 of Chapter XXI. It is based on the double hypothesis of reality and locality, as well as on the assumption that all quantum mechanical predictions are correct. The conclusion of the argument is that quantum mechanics is necessarily incomplete; to render it complete, "elements of reality" must be added to it. In an EPRB experiment, involving two spins in a singlet state, these elements of reality can be spin directions, well defined even before any measurements has been performed.

Such an addition necessarily falls outside the framework of standard quantum mechanics. Bohr was opposed to it; he argued that the concept of elements of reality proposed by EPR could not be relevant for microscopic systems, since it was meaningless to try and dissociate them, conceptually, from their experimental surrounding. As we

discussed in Chapter XXI, this position is logically sound; it allows invalidating the conclusions of the EPR argument. However, we are going to show that the double condensates offer another context for applying the EPR reasoning, particularly interesting as it involves macroscopic quantities (as well as the conservation of angular momentum). These physical quantities can, in principle, be on our scale, thereby making it more difficult to deny them an independent physical reality.

We consider a physical system in a quantum state similar to (2), where the two $|\pm\rangle$ internal states of the particles are eigenstates of the spin components along the Oz quantization axis, for example the $m = 0$ and $m = 1$ states of a spin $S = 1$. For the clarity of the discussion, we will assume that the orbital wave functions of each condensate are distinct but overlap in two regions of space, as schematized in Figure 2. In each of these two regions (which may be separated by an arbitrarily large distance), two observers, Alice and Bob, perform measurements of the spin components along transverse directions[3], N_A measurements for Alice, N_B for Bob. For each measurement performed, each of the observers chooses an arbitrary direction defined by an angle φ_i; Alice's choices are completely independent of those of Bob, and vice versa. A first series of measurements ($1 \leq i \leq N_A$) is performed by Alice in a first region of space; right after that, Bob performs another series ($N_A + 1 \leq i \leq N_A + N_B$) in his own laboratory, located very far away.

Now we saw that, as soon as Alice has measured the spins of a few particles, the relative phase of the two condensates in the entire space is fixed with a fairly good precision (the larger the number of measurements, the better the precision). These measurements also fix the transverse direction of the spins. Alice cannot, however, decide what this direction will be, as it is fixed in a totally random way in the measurement process. Standard quantum mechanics then predicts that when Bob will perform his own measurements, it is practically certain (within a negligible error) that he will find the same relative phase. As he can perform a large number of measurements, he can find out, practically instantaneously, the spin direction created and observed by Alice. The EPR argument underlines that, as no interaction had time to propagate from Alice to Bob, it is not possible for this transverse orientation to have been created by Alice's measurements: it necessarily existed prior any measurements.

What is new in our case, compared to the two-spin case, is that Bob's observations may concern an arbitrarily large number of spins; his experiment then amounts to measuring the angular momentum direction of a macroscopic spin system, which has an arbitrarily large angular momentum. As we are now dealing with macroscopic quantities, one can no longer argue, as Bohr did, that the microscopic world is accessible neither to human experimentation nor to human language description. In our present case, it seems more artificial to refuse, as suggested by Bohr, to consider separately the physical properties of systems located in distinct regions of space. The EPR argument becomes harder to refute. Reference [90] contains a discussion of this unexpected situation in terms of conservation of angular momentum.

[3]The longitudinal direction is the direction of the spin polarization in the initial state (2), the transverse directions are all the perpendicular directions.

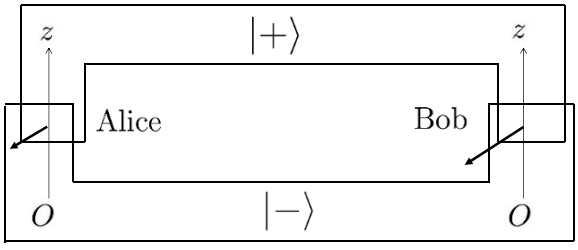


Figure 2: Scheme of an experiment on a double spin condensate, one in a + internal state, the other in a − state. The two condensates have distinct orbital wave functions, overlapping in two regions of space where two observers, Alice and Bob, perform measurements. These two regions can be separated by an arbitrarily large distance. Alice and Bob measure the spins one by one, choosing each time a transverse component (perpendicular to the Oz quantization axis) defined by an angle φ. Whereas, initially, the two condensates have no relative phase, the first measurements performed by Alice create one. The paradox is that this phase propagates instantaneously to the remote region where Bob performs his measurements. This is reminiscent of the EPR argument, but in a more striking case where the EPR "elements of reality" can be macroscopic.

3-c. Measuring all the spins, violation of Bell's inequalities

Imagine now that we measure all the spins; the selective effect around the $\Lambda = 0$ value does not occur any longer. The interpretation in terms of probabilities of individual events is no longer possible: the factor $[\cos\Lambda + \eta_j \cos(\lambda - \varphi_j)]$ in (23) can sometimes take on negative values, which rules out its possibility to be considered as a standard probability. Actually, it is not unusual in quantum mechanics that purely quantum effects arise through "negative probabilities ", as in this case. The angle Λ is called the "quantum angle ", which underlines that its role is to introduce such quantum effects, such as non-locality effects and violations of the Bell's inequalities.

To prove that such violations occur for any value of N requires using relation (23), which involves many parameters (all the measuring angles, which are arbitrary); it is easier to perform a numerical calculation as explained in the second reference of [90]. Our objective here is to simply show that the phase does not always behave in a classical way. This is why, without presenting the numerical calculation, we shall study the behavior of expression (23) for the simple case of two measurements on two spins ($Q = 2$ and $N = 1$). This will enable us to show that this expression does predict the existence of violations of the inequalities, for certain cases (more general cases are treated in the above reference). Clearly, in the case of two spins we could have carried out this computation in a simpler and more direct way.

Using definition (A-7) of Chapter XV for the Fock states, the state (2) can be

written in the form:

$$|\Psi\rangle = \left(a_{u,+}\right)^{\dagger}\left(a_{u,-}\right)^{\dagger}|0\rangle = \frac{1}{\sqrt{2}}\left[|1:u,+;2:u,-\rangle + |1:u,-;2:u,+\rangle\right]$$
$$= |1:u;2:u\rangle \otimes \frac{1}{\sqrt{2}}\left[|1:+;2:-\rangle + |1:-;2:+\rangle\right] \tag{32}$$

This leads to an entangled spin state, very similar to the one considered in § B of Chapter XXI. The only difference is the + sign in the present spin state, instead of the − in the singlet state considered in that chapter, but this difference is of no great consequence (we shall come back to this point more precisely – see note 4). Such a state can be expected to lead to significant quantum effects, as for example to situations violating Bell's inequalities.

We can also go back to the general relation (23) to show that it indeed leads to violations of Bell's inequalities in this simple case. For $N = 1$, this relation becomes:

$$\mathcal{P}_2\left(\eta_1, \varphi_1; \eta_2, \varphi_2\right)$$
$$\propto \int_{-\pi}^{+\pi} \frac{\mathrm{d}\Lambda}{2\pi} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \left[\cos\Lambda + \eta_1 \cos\left(\lambda - \varphi_1\right)\right]\left[\cos\Lambda + \eta_2 \cos\left(\lambda - \varphi_2\right)\right]$$
$$= \frac{1}{2}\left[1 + \eta_1\eta_2 \ \cos\varphi_1 \cos\varphi_2 + \eta_1\eta_2 \ \sin\varphi_1 \sin\varphi_2\right] \tag{33}$$

where we have used the fact that the average value on the circle of cosine squared or sine squared is equal to 1/2, whereas the average value of the product of cosine and sine is zero. We obtain:

$$\mathcal{P}_2\left(\eta_1, \varphi_1; \eta_2, \varphi_2\right) \propto \frac{1}{2}\left[1 + \eta_1\eta_2 \ \cos\left(\varphi_1 - \varphi_2\right)\right] \tag{34}$$

Normalizing to unity the sum of the 4 probabilities obtained for $\eta_1 = \pm 1$ and $\eta_2 = \pm 1$, we finally get:

$$\mathcal{P}_2\left(+1,+1\right) = \mathcal{P}_2\left(-1,-1\right) = \frac{1}{2}\cos^2\left(\frac{\varphi_1 - \varphi_2}{2}\right)$$
$$\mathcal{P}_2\left(+1,-1\right) = \mathcal{P}_2\left(-1,+1\right) = \frac{1}{2}\sin^2\left(\frac{\varphi_1 - \varphi_2}{2}\right) \tag{35}$$

These relations are very similar to equalities (B-10) of Chapter XXI, if we make the change[4]:

$$\varphi_1 \Rightarrow \theta_A + \pi$$
$$\varphi_2 \Rightarrow \theta_B \tag{36}$$

The angles φ_1 and φ_2 now play the role of the analyzer orientation angles in Figure 2 of that chapter. Now we know (Chapter XXI, § F-3) that these equalities lead to significant violations of Bell's inequalities (by a factor $\sqrt{2}$), hence to marked non-local quantum effects; such effects should therefore be expected in our present case.

[4]This change results from the + sign in the spin state (32), instead of the − sign in the singlet state.

In the general case where N can take on any value, measuring the totality of the spins may lead to strong violations of Bell's inequalities, provided the measurement angles are judiciously chosen (for large N, the angular domain leading to such violations decreases as the inverse of the square root of that number [90]). Note however that these violations will disappear as soon as certain spins are no longer measured (or their corresponding results no longer taken into account, which amounts to the same thing).

Conclusion

This complement is an illustration of the limits of the phenomenon studied in the previous complement: the process of successive measurements builds up a phase that has all the properties of a classical phase, but only up to a certain point. If all the particles are measured, and for certain particular choices of the measuring angles, in an ideal experiment the phase should exhibit some distinctly quantum properties.

Furthermore, one could have expected that the extreme quantum properties, as for example their non-local aspects discussed in §§ F-1 and F-3 of Chapter XXI, would only concern systems with a small particle number, or in singlet spin states (these having a special status among all the states accessible to a physical system). The present complement shows that this is not at all the case: in principle, the same properties should exist for systems composed of a very large number of particles, in a fairly simple quantum state (a double condensate).

Appendix IV

Feynman path integral

1 Quantum propagator of a particle 2267
1-a Expressing the propagator as a sum of products 2268
1-b Calculation of the matrix elements 2268
2 Interpretation in terms of classical histories 2272
2-a Expressing the propagator as a function of classical actions . 2272
2-b Generalization: several particles interacting via a potential . 2274
3 Discussion; a new quantization rule 2274
3-a Analogy with classical optics 2274
3-b A new quantization rule 2275
4 Operators 2276
4-a One single operator 2276
4-b Several operators 2279

In Chapter III, we introduced the postulates of quantum mechanics using the Hamiltonian approach, with quantization rules applied to conjugate Hamiltonian variables. It is however possible to introduce quantum mechanics and its quantization rules in an entirely different way, starting from a classical Lagrangian, and using Feynman path integrals. This approach provides an interesting insight to the relationship between classical and quantum physics, reminiscent of the connections between geometric and wave optics. Furthermore, this approach is preferable in a certain number of cases, in particular for situations where we know the classical Lagrangian but not the conjugate variables necessary to define a Hamiltonian[1].

This appendix is an elementary introduction to Feynman path integrals, and some of their properties, without too much concern for mathematical rigor. We first study in § 1 the quantum propagator of a particle, and then show in § 2 how to express it as the sum of contributions coming from different classical histories (possible evolutions) of the physical system. Once these results have been established, we discuss in § 3 how to take the inverse point of view, and start from these classical histories to derive the usual form of quantum mechanics, its quantization rules, its propagators and, in § 4, its operators. For the sake of simplicity, we shall only consider an ensemble of particles interacting via a position dependent potential; for the study of more general cases (vector potential, commutative or non-commutative gauge invariance), the reader may consult references [91], [92], or [93].

1. Quantum propagator of a particle

Consider a spinless particle. The propagator $G(\mathbf{r}', t'; \mathbf{r}, t)$ of this particle is defined as:

$$G(\mathbf{r}', t'; \mathbf{r}, t) = \langle \mathbf{r}' | U(t', t) | \mathbf{r} \rangle \tag{1}$$

[1] This happens, for example, if the Lagrangian does not depend on the time derivative of a coordinate q_i, in which case one cannot define the conjugate momentum p_i.

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

where the kets $|\mathbf{r}\rangle$ are the eigenkets of the position operator, and $U(t',t)$ the evolution operator between the initial time t and the final time t' (Complement $\mathrm{F_{III}}$). This propagator gives the probability amplitude for the particle, starting at time t from a state localized at point $\mathbf{r}$, to be found at a later time t' at point $\mathbf{r}'$.

1-a. Expressing the propagator as a sum of products

We can split the time interval $[t,t']$ into $\mathcal{N}$ equal smaller segments by setting:

$$\delta t = \frac{t'-t}{\mathcal{N}} \tag{2}$$

We therefore introduce $\mathcal{N}-1$ intermediate times $t_1,\ t_2,\ ..,\ t_k,\ ..,t_{\mathcal{N}-1}$ (see Figure 1), which are equal to:

$$t_k = t + k\ \delta t \qquad k = 1,2,..,\mathcal{N}-1 \tag{3}$$

It will also be useful for what follows to set $t_0 = t$ and $t_{\mathcal{N}} = t'$. We can then express $U(t',t)$ as a product of $\mathcal{N}$ terms:

$$U(t',t) = U(t_{\mathcal{N}},t_{\mathcal{N}-1})\ ...\ U(t_k,t_{k-1})\ ...\ U(t_2,t_1)\,U(t_1,t_0) \tag{4}$$

Between each evolution operators, we introduce a closure relation in the $\{|\mathbf{r}\rangle\}$ basis:

$$\begin{aligned} U(t',t) = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 .. \int \mathrm{d}^3 r_k .. \int \mathrm{d}^3 r_{\mathcal{N}-1} \\ U(t_{\mathcal{N}},t_{\mathcal{N}-1})\,|\mathbf{r}_{\mathcal{N}-1}\rangle\,\langle\mathbf{r}_{\mathcal{N}-1}|\,U(t_{\mathcal{N}-1},t_{\mathcal{N}-2})\,|\mathbf{r}_{\mathcal{N}-2}\rangle\ ... \\ \langle\mathbf{r}_k|\,U(t_k,t_{k-1})\,|\mathbf{r}_{k-1}\rangle\ ...\ \langle\mathbf{r}_1|\,U(t_1,t_0) \end{aligned} \tag{5}$$

Inserting this equality in (1), the propagator is expressed as:

$$\begin{aligned} G(\mathbf{r}',t';\mathbf{r},t) = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 .. \int \mathrm{d}^3 r_k .. \int \mathrm{d}^3 r_{\mathcal{N}-1} \\ \langle\mathbf{r}'|\,U(t_{\mathcal{N}},t_{\mathcal{N}-1})\,|\mathbf{r}_{\mathcal{N}-1}\rangle\,\langle\mathbf{r}_{\mathcal{N}-1}|\,U(t_{\mathcal{N}-1},t_{\mathcal{N}-2})\,|\mathbf{r}_{\mathcal{N}-2}\rangle\ ... \\ ...\ \langle\mathbf{r}_k|\,U(t_k,t_{k-1})\,|\mathbf{r}_{k-1}\rangle\ ...\ \langle\mathbf{r}_1|\,U(t_1,t_0)\,|\mathbf{r}\rangle \end{aligned} \tag{6}$$

We now let δt tend towards zero (or, equivalently, $\mathcal{N}$ towards infinity). The number of integrations over the $\mathrm{d}^3 r_k$ will then tend towards infinity but the matrix elements are now those of the evolution operator over an infinitesimal time δt.

1-b. Calculation of the matrix elements

We now compute the matrix elements $\langle\mathbf{r}_k|\,U(t_k,t_{k-1})\,|\mathbf{r}_{k-1}\rangle$ of the evolution operators, with $t_k = t_{k-1} + \delta t$. The particle's Hamiltonian H is the sum of the kinetic energy and the energy associated with an external potential:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R}) \tag{7}$$

where $\mathbf{P}$ and $\mathbf{R}$ are, respectively, the momentum and position operators of the particle; m is its mass.

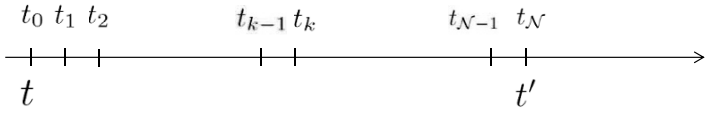


Figure 1: To express the evolution operator as a product of operators, one introduces between the initial time t and the final time t' a whole series of intermediate times t_1, t_2, .., t_k, ..,$t_{\mathcal{N}-1}$; an evolution operator is associated with each of the N time intervals. Introducing at each time t_k a closure relation then leads to a summation over all the possible positions $\mathbf{r}_k$ of the particle at that time.

α. *Free particle*

For a free particle, the Hamiltonian is simply the kinetic energy. Introducing a closure relation on the momentum eigenstates, we can write:

$$\langle \mathbf{r}_k |\, e^{-i\mathbf{P}^2\delta t/2m\hbar}\, |\mathbf{r}_{k-1}\rangle = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k\; e^{i\mathbf{k}\cdot(\mathbf{r}_k-\mathbf{r}_{k-1})}\; e^{-i\hbar k^2\delta t/2m} \tag{8}$$

We show below that this propagator can be computed exactly, whether the time interval δt is infinitesimal or finite; its expression is:

$$\langle \mathbf{r}_k |\, e^{-i\mathbf{P}^2\delta t/2m\hbar}\, |\mathbf{r}_{k-1}\rangle = \left(\frac{m}{2\pi\hbar\;\delta t}\right)^{3/2} e^{-3i\pi/4}\; e^{im(\mathbf{r}_k-\mathbf{r}_{k-1})^2/2\hbar\delta t} \tag{9}$$

Imagine for a moment that in the argument of the last exponential function the factor i is replaced by a simple minus sign (this amounts to switching to an imaginary time). Within a numerical factor, the propagator of the free particle becomes a Gaussian function, decreasing rapidly as soon as the distance $|\mathbf{r}_k - \mathbf{r}_{k-1}|$ becomes large compared to its width $\sqrt{2\hbar\delta t/m}$.

Comment:

This result is not surprising since if one replaces, in Schrödinger's equation, the real time t by an imaginary time it, one gets the diffusion equation whose propagator is indeed a Gaussian. The width of that Gaussian goes to zero as $\delta t \to 0$: as expected, the shorter the time δt, the smaller the distance covered by the particle. It is worth noting, however, that this distance is not proportional to δt, but to the square root of this time; in other words, during very short times, the particle may propagate much further than if it had a constant velocity. This situation is characteristic of a random motion whose correlation time and mean free path tend simultaneously to zero, in a way where the diffusion coefficient remains constant: such a process lead to the classical diffusion equation, and its propagator therefore has the same property.

As for Schrödinger's equation itself (without switching to an imaginary time), when the distance $|\mathbf{r}_k - \mathbf{r}_{k-1}|$ increases, instead of a decrease of the propagator's modulus, we

get an oscillation that gets faster as the distance gets greater; these oscillations are also faster as δt gets shorter. We discuss below how all these phases interfere, as well as the particular role played by the phases associated with classical paths.

Demonstration of relation (9):

We modify the exponent of the function to be integrated in (8) to introduce a perfect square:

$$\frac{\hbar\ \delta t}{2m}\mathbf{k}^2 + \mathbf{k}\cdot(\mathbf{r}_k - \mathbf{r}_{k-1}) = \frac{\hbar\ \delta t}{2m}\left[\mathbf{k} + \frac{m}{\hbar\ \delta t}(\mathbf{r}_k - \mathbf{r}_{k-1})\right]^2 - \frac{m}{2\hbar\ \delta t}(\mathbf{r}_k - \mathbf{r}_{k-1})^2 \tag{10}$$

Making the change of variables:

$$\mathbf{q} = \mathbf{k} + \frac{m}{\hbar\ \delta t}(\mathbf{r}_k - \mathbf{r}_{k-1}) \tag{11}$$

leads us to:

$$\langle \mathbf{r}_k |\, e^{-i\mathbf{P}^2\delta t/2m\hbar}\, |\mathbf{r}_{k-1}\rangle = \frac{1}{(2\pi)^3}\left[\int \mathrm{d}^3q\ e^{-i\hbar q^2\delta t/2m}\right] \times e^{im\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)^2/2\hbar\delta t} \tag{12}$$

The integral that still remains between the brackets on the right-hand side is now a number independent of $\mathbf{r}_k$ and $\mathbf{r}_{k-1}$. As the three components of the vector $\mathbf{q}$ yield the same contribution, this number is the cube of the integral:

$$I = \int_{-\infty}^{+\infty} \mathrm{d}q_x\ e^{-i\hbar q_x^2\ \delta t/2m} \tag{13}$$

Following a classical procedure, we compute the square of that integral I using the polar coordinates where $\rho = \sqrt{q_x^2 + s_x^2}$:

$$\begin{aligned} I^2 &= \int_{-\infty}^{+\infty} \mathrm{d}q_x \int_{-\infty}^{+\infty} \mathrm{d}s_x\ e^{-i\hbar\left(q_x^2+s_x^2\right)\delta t/2m} \\ &= 2\pi\int_0^{\infty} \rho\ d\rho\ e^{-i\hbar\rho^2\delta t/2m} = \pi\frac{2im}{\hbar\ \delta t}\left[e^{-i\infty} - 1\right] \end{aligned} \tag{14}$$

Now, writing the number $e^{-i\infty}$ is mathematically meaningless: as $M \to \infty$, the imaginary exponential e^{iM} oscillates indefinitely between $+1$ and -1, around a zero average value. We shall simply take this number equal to its zero average value[2], so that:

$$I = \sqrt{\frac{2\pi m}{\hbar\ \delta t}}e^{-i\pi/4} \tag{15}$$

Replacing by I^3 the integral over d^3q in (12), we get relation (9).

[2]One can reach this same result by adding a small imaginary part $-i\epsilon$ to the coefficient $\hbar^2\delta t/2m$ of ρ^2 in the exponent. Thanks to this imaginary part, which can be arbitrarily small, the oscillating term does disappear.

β. Effects of the external potential

We must now compute the matrix elements of the exponential operator $e^{-iH\delta t/\hbar}$ when H is the sum of two non-commuting operators. We shall only consider the case where δt is infinitesimal. In that case, the exponential of a sum of non-commuting operators A and B can be written:

$$e^{-i(A+B)\delta t} = 1 - i\left[A+B\right]\delta t - \left[A^2 + AB + BA + B^2\right]\frac{\delta t^2}{2} + 0\left(\delta t^3\right) \tag{16}$$

We can also expand the following product of exponentials as:

$$e^{-iA\delta t/2}\ e^{-iB\delta t}\ e^{-iA\delta t/2} = \left[1 - iA\frac{\delta t}{2} - A^2\frac{\delta t^2}{8} + ..\right]\times \left[1 - iB\delta t - B^2\frac{\delta t^2}{2} + ..\right]\left[1 - iA\frac{\delta t}{2} - A^2\frac{\delta t^2}{8} + ..\right] \tag{17}$$

This leads to:

$$\begin{aligned} &e^{-iA\delta t/2}\ e^{-iB\delta t}\ e^{-iA\delta t/2} \\ &\quad = 1 - i\left[A+B\right]\delta t - \left[\frac{A^2}{2} + \frac{A^2}{2} + \frac{B^2}{2} + \frac{B^2}{2} + AB + BA\right]\frac{\delta t^2}{2} + 0\left(\delta t^3\right) \\ &\quad = e^{-i(A+B)\delta t} + 0\left(\delta t^3\right) \end{aligned} \tag{18}$$

Setting $A = \mathbf{P}^2/2m\hbar$ and $B = V\left(\mathbf{R}\right)/\hbar$ we get, assuming that δt is infinitesimal so that we can neglect the δt^3 terms:

$$\begin{aligned} \left\langle \mathbf{r}_k\right| U\left(t_k, t_{k-1}\right)\left|\mathbf{r}_{k-1}\right\rangle &= \left\langle \mathbf{r}_k\right| e^{-iV(\mathbf{R})\delta t/2\hbar}\ e^{-i\mathbf{P}^2\delta t/2m\hbar}\ e^{-iV(\mathbf{R})\delta t/2\hbar}\left|\mathbf{r}_{k-1}\right\rangle \\ &= e^{-i[V(\mathbf{r}_k)+V(\mathbf{r}_{k-1})]\delta t/2\hbar}\ \left\langle \mathbf{r}_k\right| e^{-i\mathbf{P}^2\delta t/2m\hbar}\left|\mathbf{r}_{k-1}\right\rangle \end{aligned} \tag{19}$$

This result is simply the product of an exponential including the potential and a term that is the propagator of the free particle. Taking (9) into account, we can write:

$$\begin{aligned} &\left\langle \mathbf{r}_k\right| U\left(t_{k-1}, t_k\right)\left|\mathbf{r}_{k-1}\right\rangle \\ &\quad = \left(\frac{m}{2\pi\hbar\ \delta t}\right)^{3/2} e^{-3i\pi/4}\ e^{im(\mathbf{r}_k - \mathbf{r}_{k-1})^2/2\hbar\delta t} \times e^{-i[V(\mathbf{r}_k)+V(\mathbf{r}_{k-1})]\delta t/2\hbar} \end{aligned} \tag{20}$$

The effect of the external potential is straightforward: it adds an exponential including half the sum of the two potential energies associated, one with the position of the bra, the other with the position of the ket.

γ. Final expression

Inserting (20) into (5), we get the following expression for the propagator:

$$G\left(\mathbf{r}', t'; \mathbf{r}, t\right) = \left(\frac{e^{-i\pi/2}\ m}{2\pi\hbar\ \delta t}\right)^{3\mathcal{N}/2} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 .. \int \mathrm{d}^3 r_k .. \int \mathrm{d}^3 r_{\mathcal{N}-1} \exp\left\{ i \sum_{k=1}^{\mathcal{N}-1} \left[\frac{m\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)^2}{2\hbar\delta t} - \frac{V\left(\mathbf{r}_k\right) + V\left(\mathbf{r}_{k-1}\right)}{2\hbar}\delta t\right]\right\} \tag{21}$$

This equality is valid in the limit where $\delta t \to 0$ (or $\mathcal{N} \to \infty$) since, to establish (19), we neglected the δt^3 terms.

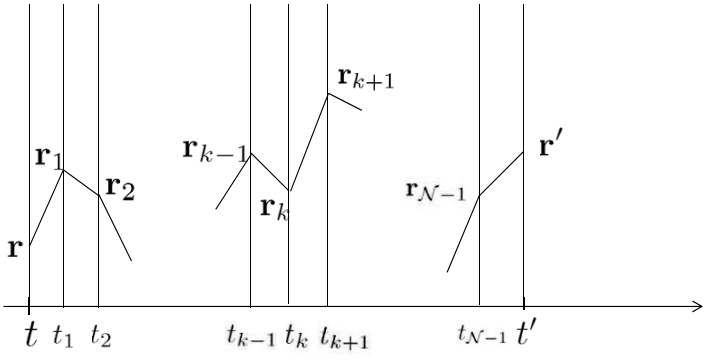


Figure 2: A Feynman path is obtained by associating a position $\mathbf{r}_k$ *with each time* t_k*. The path thus obtained is continuous, but looks in general like a zigzag (the velocity is discontinuous at each time* t_k*). Nevertheless, one can associate a classical action with each of these paths, and get the quantum propagator by summing exponentials of the actions along all these classical paths: while keeping the two end positions* $\mathbf{r}$ *and* $\mathbf{r}'$ *fixed, the summation is performed on all the intermediate positions* $\mathbf{r}_1$, $\mathbf{r}_2$,...,$\mathbf{r}_k$,...$\mathbf{r}_{N-1}$.

2. Interpretation in terms of classical histories

The probability amplitude (6) is obtained by inserting, as many times as necessary, the propagator matrix elements (20). It thus contains as many sums over intermediate positions as there are intermediate times t_k, and this number goes to infinity as $\delta t \to 0$. This expression, somewhat complicated at first sight, actually has a simple interpretation in terms of classical paths the particles can follow between the initial and final times.

2-a. Expressing the propagator as a function of classical actions

Let us go back to classical physics and consider for a moment all the $\mathbf{r}_k$ as being fixed, and a particle that goes through these successive positions at the different times t_k. In between two consecutive times, we assume the particle keeps a constant velocity equal to:

$$\mathbf{v}_k = \frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t} \tag{22}$$

This defines a classical path Γ, with a linear interpolation for times between the discrete instants t_k (*cf.* Figure 2). The particle's Lagrangian $\mathcal{L}$ is written:

$$\mathcal{L} = \frac{1}{2} m \mathbf{v}^2 - V(\mathbf{r}) \tag{23}$$

For any classical path Γ followed by the particle between the initial time t and the final time t', the position and velocity are both functions of time, and so is the Lagrangian $\mathcal{L}(t)$. The corresponding action is written (Appendix III, § 5-b):

$$S_\Gamma = \int_t^{t'} \mathrm{d}t''\ \mathcal{L}(t'') \tag{24}$$

Let us compute this integral using Riemann's method by introducing, between the times t and t', $\mathcal{N}$ time intervals of length δt. We consider that during an infinitesimal time interval δt, the potential energy can be approximated by half the sum of its values at each end of the interval. We then get:

$$S_\Gamma \simeq \sum_{k=1}^{\mathcal{N}} \delta t \left\{ \left[\frac{m\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)^2}{2\ \delta t^2} \right] - \frac{V\left(\mathbf{r}_{k-1}\right) + V\left(\mathbf{r}_k\right)}{2} \right\} \tag{25}$$

This approximate equality becomes exact in the limit $\delta t \to 0$.

On the right-hand side of this equality, we find (to within a factor $i/\hbar$) the argument of the exponential appearing in expression (21) for the quantum propagator. This means that this quantum propagator contains the exponential of an approximate value of the classical action, multiplied by $i/\hbar$. When $\delta t \to 0$ (and hence $\mathcal{N} \to \infty$), the approximate value becomes exact[3], and the sums over d^3r_1, d^3r_2,.., $\mathrm{d}^3r_{\mathcal{N}-1}$ on the right-hand side of (21) introduce a summation over all the paths going from $\mathbf{r}, t$ to $\mathbf{r}', t'$:

$$\sum_{\text{paths } \Gamma} \exp\left[\frac{iS_\Gamma}{\hbar}\right] \tag{26}$$

For this summation over the paths to be meaningful, we must now choose a "path density"; we therefore assume that the number of paths in a "path interval", determined by the set of d^3r_k, is given by the product $C^{3\mathcal{N}/2} \times \mathrm{d}^3r_1 \mathrm{d}^3r_2 .. \mathrm{d}^3r_k .. \mathrm{d}^3r_{\mathcal{N}-1}$, where C is the constant (inverse of a length):

$$C = e^{-i\pi/4} \sqrt{\frac{m}{2\pi\hbar\ \delta t}} \tag{27}$$

This allows us to write:

$$\boxed{G\left(\mathbf{r}', t'; \mathbf{r}, t\right) = \left\langle \mathbf{r}' \right| U\left(t', t\right) \left| \mathbf{r} \right\rangle = \sum_{\text{paths } \Gamma} \exp\left[\frac{iS_\Gamma}{\hbar}\right]} \tag{28}$$

In the limit $\delta t \to 0$, the sum over the paths is no longer discrete. This is why, instead of (28), one often writes:

$$\left\langle \mathbf{r}' \right| U\left(t', t\right) \left| \mathbf{r} \right\rangle = \int \mathcal{D}\left[\mathbf{r}\left(t\right)\right]\ \exp\left[\frac{iS_\Gamma}{\hbar}\right] \tag{29}$$

where the notation $\mathcal{D}\left[\mathbf{r}\left(t\right)\right]$ symbolizes the limit of a sum over the paths:

$$\mathcal{D}\left[\mathbf{r}\left(t\right)\right] = \lim_{\mathcal{N}\to\infty} \left[e^{-i\pi/4} \sqrt{\frac{m}{2\pi\hbar\ \delta t}} \right]^{3\mathcal{N}/2} \int \mathrm{d}^3r_1 \int \mathrm{d}^3r_2 .. \int \mathrm{d}^3r_k .. \int \mathrm{d}^3r_{\mathcal{N}-1} \tag{30}$$

[3]Remember that this complement does not pretend to be mathematically rigorous. This would entail a more careful study of the classical as well as quantum expressions, and of the effects of the simultaneous limits δt goes to zero and $\mathcal{N}$ (number of terms in the products) goes to infinity, keeping in mind that these approximations are done on functions that are arguments of exponentials.

2-b. Generalization: several particles interacting via a potential

The previous considerations can be directly generalized to a system with several particles interacting via a potential depending on their positions. Since the system Hamiltonian is, as above, the sum of two non-commuting terms, we again use the approximate expression (18) for the evolution operator. But rather than inserting the closure relations for a single position, one must now use a basis involving the positions of all the particles: each integral over $\mathrm{d}^3 r$ is then demultiplied into as many integrals as there are particles in the system.

We will not include here the case where the particles are charged and subjected to a magnetic field, which would include terms such as $q\mathbf{v}\cdot\mathbf{A}(\mathbf{r})$, where $\mathbf{A}(\mathbf{r})$ is the vector potential. Even though it is an interesting case, in particular for its relation to gauge invariance, it will not be considered here for the sake of brevity. The interested reader should consult the references given in the introduction.

3. Discussion; a new quantization rule

The path integral approach is particularly fit for developing an analogy with classical optics. In addition, it allows building new quantization rules. These two points will be discussed successively.

3-a. Analogy with classical optics

Relations (28) and (29) allow making a link between two a priori unrelated quantities: the classical mechanics paths with their associated actions, and the quantum propagator. Knowing the wave function $\Psi(\mathbf{r},t)$ at a given time t, it is this quantum propagator that enables computing that wave function at a later time t' as follows:

$$\begin{aligned}\Psi(\mathbf{r}',t') &= \langle \mathbf{r}' \,|\Psi(t')\rangle = \langle \mathbf{r}'|\, U(t',t)\,|\Psi(t)\rangle \\ &= \int \mathrm{d}^3 r \;\langle \mathbf{r}'|\, U(t',t)\,|\mathbf{r}\rangle\,\langle \mathbf{r}\,|\Psi(t)\rangle \end{aligned} \tag{31}$$

Taking into account definition (1) of the propagator, the above expression can be rewritten as:

$$\Psi(\mathbf{r}',t') = \int \mathrm{d}^3 r\; G(\mathbf{r}',t';\mathbf{r},t)\;\Psi(\mathbf{r},t) \tag{32}$$

The propagator is thus the kernel of the integral equation expressing the temporal propagation of the wave function. Now, equality (28) shows that this propagator is equal to a sum of exponentials of the actions corresponding to all the classical paths. Therefore, whereas in classical mechanics a single path (or in certain cases a finite number of paths) is selected by the stationarity condition of the action, in quantum mechanics all the paths come into play to determine the propagation amplitude, each one with its particular phase. In a manner of speaking, one can say that, in quantum mechanics, the particle goes through all the possible intermediate positions $\mathbf{r}_k$ and hence follows all the possible histories between the two end points.

It is worth noting that all these histories (even highly unlikely histories involving totally arbitrary positions) contribute with the same amplitude. On the other hand, the

phases associated with the histories are different from each other and allow understanding how the different histories come into play. This situation can be analyzed in terms of stationary phase conditions. It is easy to understand that in the summation, the histories corresponding to a stationary action will play a particular role since all the neighboring histories will add their contribution in a coherent way. On the other hand, in the vicinity of histories for which the action varies rapidly, the phase oscillates quickly and the corresponding contributions will cancel out through destructive interference. Consequently, classical histories play a privileged role, which becomes more prominent as the phase oscillations become more rapid. In the limit where $\hbar \to 0$, these oscillations become infinitely rapid and only the classical histories prevail.

Finally, this situation reminds us of classical optics and the Huyghens-Fresnel principle, where a light wave is computed as the sum of waves radiated from each point of an intermediate surface, taking into account the phases linked to the propagation along each path. In the geometrical approximation, where the wavelength tends towards zero, the trajectories of the light rays correspond to paths having a stationary phase, i.e. which does not change for infinitely close paths. Geometrical optics is the analog of classical mechanics, whereas Huyghens wave optics is the analog of quantum mechanics. The Feynman integral path is therefore a useful tool to study the link between classical and quantum mechanics, and in particular the semiclassical limit of quantum mechanics (WKB approximation, etc.).

Comment:

The preceding analogy is well founded for a single particle. For a system with N particles, the histories no longer propagate in the ordinary three-dimensional space, but in a $3N$-dimensional configuration space. The analogy with optics described above is no longer as adequate, since in classical optics, electromagnetic waves propagate in ordinary $3D$ space.

3-b. A new quantization rule

At this stage, we can invert the approach. Until now, starting from the rules of Hamiltonian quantum mechanics, we deduced an equivalent expression for the propagator, i.e. another way for finding the solutions of Schrödinger's equation. It is also possible to consider this equivalent expression as the starting point and postulate that the propagator is defined ab initio by a sum over all the classical paths Γ, each contributing an exponential $\exp(iS_\Gamma/\hbar)$. This yields another method for the quantization of a physical system, which offers several advantages. First of all, as we just saw, it highlights the relation of quantum mechanics with classical mechanics, where only a single classical path exists (or sometimes a finite number of paths), as opposed to an infinite number of possible paths in quantum mechanics. Furthermore, it is remarkable that the probability amplitudes thus computed only depend on classical functions (involving only numbers and not operators), the only explicit quantum component being the presence of $\hbar$ in the denominator of the phase. We shall see in § 4 how the concept of operators can be introduced in this approach. In addition, the expressions involving path integrals are symmetric with respect to time and space, since both type of coordinates are integrated

in a similar way[4]. Reasoning directly in space-time makes it easier to include Einstein's relativity, since one can replace the time differential δt by a proper time differential $\delta\tau$. If now the Lagrangian $\mathcal{L}$ is a space-time scalar, so is the action, and the theory acquires relativistic invariance. Finally, Feynman's quantization method only requires the existence of a Lagrangian, with its associated variational principle. Now all the physical systems that have a Lagrangian do not necessarily have the conjugate variables permitting the definition of a Hamiltonian. For such systems, the Feynman path integral method is powerful and this is why it is so important in quantum field theory.

4. Operators

Feynman paths also permit computing the matrix elements of operators in the Heisenberg picture, where they are time-dependent. We shall mostly consider the simplest case where the operators are functions of the position operator $\mathbf{R}$.

4-a. One single operator

Let us insert any operator A "in the middle" of the evolution operator (1) by splitting the time interval $[t, t']$ into two adjacent intervals $[t, t'']$ and $[t'', t']$, with $t < t'' < t'$. We get the expression:

$$\langle \mathbf{r}' | \, U(t', t'') \;\; A \; U(t'', t) \, | \mathbf{r} \rangle = \langle \theta(\mathbf{r}', t'', t') | \, A \, | \theta(\mathbf{r}, t'', t) \rangle \tag{33}$$

with:

$$\begin{aligned} |\theta(\mathbf{r}, t'', t)\rangle &= U(t'', t) \, |\mathbf{r}\rangle \\ |\theta(\mathbf{r}', t'', t')\rangle &= U^{\dagger}(t', t'') \, |\mathbf{r}'\rangle = U(t'', t') \, |\mathbf{r}'\rangle \end{aligned} \tag{34}$$

In this matrix element of A, the ket $|\theta(\mathbf{r}, t'', t)\rangle$ is obtained by the evolution until the time t'' of a state localized at $\mathbf{r}$ at time t; the bra $\langle \theta(\mathbf{r}', t'', t')|$ corresponds to the ket $|\theta(\mathbf{r}', t'', t')\rangle$ which, as it evolves between t'' and t', becomes a ket localized at $\mathbf{r}'$:

$$U(t', t'') \, |\theta(\mathbf{r}', t'', t')\rangle = U(t', t'') \, U(t'', t') \, |\mathbf{r}'\rangle = |\mathbf{r}'\rangle \tag{35}$$

α. Operator function of the position

We now assume operator A is a function $A(\mathbf{R})$ of the particle's position operator. Inserting a closure relation on the positions, we can write the left-hand side of (33) as:

$$\langle \mathbf{r}' | \, U(t', t'') \;\; A(\mathbf{R}) \;\; U(t'', t) \, | \mathbf{r} \rangle = \int \mathrm{d}^3 r'' \, \langle \mathbf{r}' | \, U(t', t'') \, | \mathbf{r}'' \rangle \;\; A(\mathbf{r}'') \;\; \langle \mathbf{r}'' | \, U(t'', t) \, | \mathbf{r} \rangle \tag{36}$$

[4]The sum over all the paths introduces an integral over all the positions $\boldsymbol{r}_k$ in Figure 2, hence differentials of the three space coordinates; in addition, the integral over the times introduces a differential $\mathrm{d}t = t_k - t_{k-1}$. The product of three space differentials by a time differential thus allows one to introduce a differential of space-time volume.

Using the general relation (28), we can then write the two propagators appearing under the integral as the sum over all the paths Γ_1 or Γ_2:

$$\begin{aligned}\langle \mathbf{r}''| U(t'',t) |\mathbf{r}\rangle &= \sum_{\text{paths } \Gamma_1} \exp\left[\frac{iS_{\Gamma_1}}{\hbar}\right] \\ \langle \mathbf{r}'| U(t',t'') |\mathbf{r}''\rangle &= \sum_{\text{paths } \Gamma_2} \exp\left[\frac{iS_{\Gamma_2}}{\hbar}\right]\end{aligned} \tag{37}$$

where Γ_1 is a path linking the initial position $\mathbf{r}$ at time t to the intermediate position $\mathbf{r}''$ at time t'', and Γ_2 the path linking thereafter the intermediate position $\mathbf{r}''$ at time t'' to the final position $\mathbf{r}'$ at time t'. If t'' coincides with the intermediate time t_k, relation (36) becomes:

$$\begin{aligned}&\langle \mathbf{r}'| U(t',t_k)\ A(\mathbf{R})\ U(t_k,t) |\mathbf{r}\rangle \\ &\qquad = \int \mathrm{d}^3 r_k \sum_{\text{paths } \Gamma_1 \text{ and } \Gamma_2} \exp\left[\frac{iS_{\Gamma_2}}{\hbar}\right] A(\mathbf{r}_k) \exp\left[\frac{iS_{\Gamma_1}}{\hbar}\right]\end{aligned} \tag{38}$$

Now the product of the two exponentials yields a single exponential $\exp[iS_\Gamma/\hbar]$ associated with the action S_Γ of a path Γ consisting of the two paths Γ_1 and Γ_2 joined together end to end at $\mathbf{r}_k$. The sum over $\mathrm{d}^3 r_k$ reconstitutes the ensemble of all paths going from the initial position $\mathbf{r}$ at time t to the final position $\mathbf{r}'$ at time t', the only difference being that now each exponential $\exp[iS_\Gamma/\hbar]$ is multiplied by the value $A(\mathbf{r}_k)$ taken by the function A at the intermediate point at time t_k.

We finally obtain:

$$\langle \mathbf{r}'| U(t',t'')\ A(\mathbf{R})\ U(t'',t) |\mathbf{r}\rangle = \sum_{\text{paths } \Gamma} A(\mathbf{r}'')\ \exp\left[\frac{iS_\Gamma}{\hbar}\right] \tag{39}$$

where $A(\mathbf{r}'')$ is the value of $A(\mathbf{r})$ at position $\mathbf{r}''$ which the path Γ traverses at time t''. The matrix elements of the operator in the Heisenberg picture (special case $t' = t$) are thus given by the same summation over the histories as for the propagator, the only difference being that the contribution of each path is now multiplied by the value taken by the operator at the position $\mathbf{r}''$ at the intermediate time t''.

As before, we can now invert the approach and consider relation (39) as the definition of an operator in the framework of the Feynman path quantization method. Here again, it is remarkable that this relation involves only classical functions, without any operator.

β. Velocity operator; canonical commutation relations

In order to define an operator $\mathbf{W}$ associated with the particle's velocity at time t'' (we use the notation $\mathbf{W}$ to avoid any confusion with the potential V), and taking (22) into account, a natural extension of (38) leads to setting:

$$\begin{aligned}&\langle \mathbf{r}'| U(t',t'')\ \mathbf{W}\ U(t'',t) |\mathbf{r}\rangle \\ &\qquad = \int \mathrm{d}^3 r_k \sum_{\text{paths } \Gamma_2} \exp\left[\frac{iS_{\Gamma_2}}{\hbar}\right] \sum_{\text{paths } \Gamma_1} \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t}\right) \exp\left[\frac{iS_{\Gamma_1}}{\hbar}\right]\end{aligned} \tag{40}$$

where the paths Γ_1 are all those going from the initial position $\mathbf{r}$ to the intermediate position $\mathbf{r}_k$ (the preceding intermediate position $\mathbf{r}_{k-1}$ does depend on the path), whereas the paths Γ_2 are all those going from $\mathbf{r}_k$ to the final position $\mathbf{r}'$. Introducing in the middle of the left-hand side a closure relation on the kets $\{|\mathbf{r}_k\rangle\}$, this relation becomes:

$$\int \mathrm{d}^3 r_k \;\; \langle \mathbf{r}' | \, U\left(t', t''\right) | \mathbf{r}_k \rangle \;\; \langle \mathbf{r}_k | \, \mathbf{W} \; U\left(t'', t\right) | \mathbf{r} \rangle$$
$$= \int \mathrm{d}^3 r_k \;\; \langle \mathbf{r}' | \, U\left(t', t''\right) | \mathbf{r}_k \rangle \;\; \Psi_W\left(\mathbf{r}_k, t''\right) \tag{41}$$

with:

$$\Psi_W\left(\mathbf{r}_k, t''\right) = \langle \mathbf{r}_k | \, \mathbf{W} \; U\left(t'', t\right) | \mathbf{r} \rangle = \sum_{\text{paths } \Gamma_1} \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t} \right) \exp\left[\frac{i S_{\Gamma_1}}{\hbar} \right] \tag{42}$$

This wave function is the result of the action of operator $\mathbf{W}$ on the wave function at time t'', equal to:

$$\Psi\left(\mathbf{r}_k, t''\right) = \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle \tag{43}$$

Let us compare the sum over the paths Γ_1 in relation (42) and in the relation (28) used to build the propagator between the times t and t''. In these two equalities, the actions are given by the summations (25) over the intermediate positions. Concerning the contributions of the paths between the initial time and the intermediate time t_{k-1} (the last time over which the summation runs), the two sums in (42) and (28) are identical. Actually, their only difference concerns the very last time interval (between t_{k-1} and $t_k = t''$), which in (42) is multiplied by the factor $\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)/\delta t$. Now this multiplicative factor can also be found by taking the derivative of (28) with respect to $\mathbf{r}_k$ since, using (25), we have:

$$\boldsymbol{\nabla}_{\mathbf{r}_k} \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle = \frac{i \delta t}{\hbar} \sum_{\text{paths } \Gamma_1} \left[m \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t^2} \right) + \frac{1}{2} \boldsymbol{\nabla}_{\mathbf{r}_k} V \right] \exp\left[\frac{i S_{\Gamma_1}}{\hbar} \right] \tag{44}$$

As $\mathbf{r}_k$ is a final fixed point, the term in $\boldsymbol{\nabla}_{\mathbf{r}_k} V$ on the right-hand side can be taken out of the summation over the paths. It yields a contribution in $\delta t \times \boldsymbol{\nabla}_{\mathbf{r}_k} V$ which goes to zero in the limit $\delta t \to 0$. We are left with:

$$\boldsymbol{\nabla}_{\mathbf{r}_k} \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle = \frac{i m}{\hbar} \sum_{\text{paths } \Gamma_1} \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t} \right) \exp\left[\frac{i S_{\Gamma_1}}{\hbar} \right] \tag{45}$$

so that relation (42) becomes:

$$\Psi_W\left(\mathbf{r}_k, t''\right) = \frac{\hbar}{i m} \boldsymbol{\nabla}_{\mathbf{r}_k} \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle = \frac{\hbar}{i m} \boldsymbol{\nabla}_{\mathbf{r}_k} \Psi\left(\mathbf{r}_k, t''\right) \tag{46}$$

This means that the action of the velocity operator $\mathbf{W}$ is simply proportional to a derivative[5] with respect to the position $\mathbf{r}_k$, which is the variable of the wave function at the

[5]The demonstration pertains to a wave function at the instant t'' that is issued from a wave function localized at point $\mathbf{r}$ at time t. By linear superposition, it can be generalized to any wave function at time t, hence confirming the same derivation property.

instant t''. In other words, if $\mathbf{P} = m\mathbf{W}$ is the particle's momentum operator, its action on the wave function is $(\hbar/i)$ times the gradient with respect to the position. We have established a basic result of the usual quantum mechanics, starting from operators introduced in the path integral approach.

The canonical commutation relations between $\mathbf{R}$ and $\mathbf{P}$ are easily derived, since:

$$\frac{\partial}{\partial x}\left[x\Psi\left(x\right)\right] = x\frac{\partial}{\partial x}\left[\Psi\left(x\right)\right] + \Psi\left(x\right) \tag{47}$$

These commutation relations can also be considered as consequences of the path quantization rules.

4-b. Several operators

Feynman postulates also permit introducing products of several operators, acting at the same instant or at different times.

α. Several operators at different times

The previous argument can be generalized to several operators $A\left(\mathbf{R}\right)$, $B\left(\mathbf{R}\right)$, etc. acting at intermediate times t'', t''', etc. As before, we can split the evolution operator into several parts corresponding to the successive time intervals, and insert position closure relations at the intermediate times. Each operator introduces a factor dependent on the corresponding intermediate position, and the time propagation is a sum over histories between successive time intervals. For instance, for two operators, the same reasoning followed above leads to:

$$\begin{aligned}\left\langle \mathbf{r}'\right| U\left(t',t'''\right)\, B\left(\mathbf{R}\right)\, U\left(t''',t''\right)\, A\left(\mathbf{R}\right)\, U\left(t'',t\right)\left|\mathbf{r}\right\rangle \\ = \sum_{\text{paths }\Gamma} B\left(\mathbf{r}'''\right) A\left(\mathbf{r}''\right) \exp\left[\frac{iS_\Gamma}{\hbar}\right]\end{aligned} \tag{48}$$

where $A\left(\mathbf{r}''\right)$ is the value of A for the position $\mathbf{r}''$ crossed by the path Γ at time t'', $B\left(\mathbf{r}'''\right)$ the value of B for the position $\mathbf{r}'''$ crossed by the path Γ at a later time t'''. The result is easily generalized to any number of operators. Note the order in which the operators are arranged in the matrix element on the left-hand side: it corresponds to the order in which the times t'', t''', etc. are arranged in the classical histories used to calculate the actions. The quantum operators are automatically arranged in decreasing times from left to right, even if $A\left(\mathbf{r}''\right)$ and $B\left(\mathbf{r}'''\right)$ are numbers that commute in the right hand side of relation (48).

β. Position and velocity operators, symmetrization

Imagine now we want to introduce, for example, the operator corresponding to the product $\mathbf{R}\cdot\mathbf{P}$ of the position and momentum. We will then proceed as in (41) and (42), however with an added precaution: should we multiply $\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)/\delta t$ by $\mathbf{r}_k$ or by $\mathbf{r}_{k-1}$ (the order does not matter, since we are dealing with numbers). For the sake of symmetry, we multiply by half their sum, so that in (42) $\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)/\delta t$ is replaced by:

$$\frac{1}{\delta t}\frac{\mathbf{r}_k + \mathbf{r}_{k-1}}{2}\cdot\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right) = \frac{1}{2\delta t}\left[\mathbf{r}_k\cdot\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right) + \left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)\cdot\mathbf{r}_{k-1}\right] \tag{49}$$

On the right-hand side, we wrote the two terms so that the times are always decreasing (or stationary) from left to right. Because of the order of the k indices, the first term on the right-hand side introduces the operator $\mathbf{R} \cdot \mathbf{W}$, whereas the second introduces $\mathbf{W} \cdot \mathbf{R}$, i.e. the same operators but in the inverse order. This is an example of how the path integral method leads quite naturally to a symmetrization of the operator order, which automatically ensures the hermiticity of their product.

Conclusion

To be able to use two complementary approaches, the Hamiltonian method and the path integral method, is often quite valuable in the study of numerous physical problems. As an example, path integrals play a fundamental role in field theory. They are for a large part at the base of the implementation of symmetry groups (Abelian or non-commutative) in this theory, which allows building a theory for elementary particles and their interactions. There are, however, other cases where the path integral formalism is very useful, as for example, in the computations of quantum interference with cold atoms. Conceptually, the path integral approach can shed new light on the relations between quantum mechanics and classical mechanics, as well as classical optics as we saw in § 3-a.

In this appendix, we only used path integrals as a method to compute the time propagator of a quantum physical system, which involves imaginary exponentials of the Hamiltonian. Path integrals can also be used in quantum statistical mechanics (Appendix VI), and involve real exponentials of the Hamiltonian (multiplied by the inverse of the temperature). It is the basic tool for many numerical calculations; the interested reader can consult Zinn-Justin's book [92], or reference [94] where, in particular, the PIMC (Path Integral Quantum Monte Carlo) methods are described.

Appendix V

Lagrange multipliers

1 Function of two variables 2281
2 Function of N variables 2283

When a function F depends on non-independent variables (i.e. which are related by constraints), its extrema (maxima or minima) can be found by the Lagrange multiplier method. A brief summary of this method is proposed in this appendix. The first part concerns functions of two variables, and the second part will generalize the concept to any number of variables.

1. Function of two variables

Consider first a real function $F(x_1, x_2)$ of two independent variables x_1 and x_2. We assume the fonction F to be regular, continuous, differentiable with continuous derivatives. The extrema of F correspond to values of the variables for which the two partial derivatives are zero:

$$\frac{\partial F(x_1, x_2)}{\partial x_1} = 0 \qquad ; \qquad \frac{\partial F(x_1, x_2)}{\partial x_2} = 0 \tag{1}$$

These two relations amount to stating that the gradient of F must be zero:

$$\overrightarrow{\nabla F} = 0 \tag{2}$$

Two equations with two unknowns x_1 and x_2 generally admit a finite number of solutions (pairs of values for x_1 and x_2); this number can even be zero if the function F does not present any extrema.

Let us now look for the extrema of F when the variables are no longer independent, but must obey a constraint:

$$E(x_1, x_2) = C \tag{3}$$

where C is a constant and $E(x_1, x_2)$ a regular function (continuous, differentiable, etc.). When this constraint is satisfied, the point M with coordinates x_1 and x_2 is forced to follow a curve in the plane (solid line in Figure 1). Imagine we place the point close to an arbitrary point M of the curve, and move it by varying slightly its coordinates by $\mathrm{d}x_1$ and $\mathrm{d}x_2$. For E to remain constant, $\mathrm{d}x_1$ and $\mathrm{d}x_2$ must necessarily obey:

$$\mathrm{d}E = \frac{\partial E(x_1, x_2)}{\partial x_1}\,\mathrm{d}x_1 + \frac{\partial E(x_1, x_2)}{\partial x_2}\,\mathrm{d}x_2 = \overrightarrow{\nabla E} \cdot \overrightarrow{\mathrm{d}M} = 0 \tag{4}$$

The point M therefore necessarily moves along the tangent to the curve, i.e. perpendicularly to its gradient $\overrightarrow{\nabla E}$, as shown in Figure 1. As for the variation of F, it is given by:

$$\mathrm{d}F = \overrightarrow{\nabla F} \cdot \overrightarrow{\mathrm{d}M} \tag{5}$$

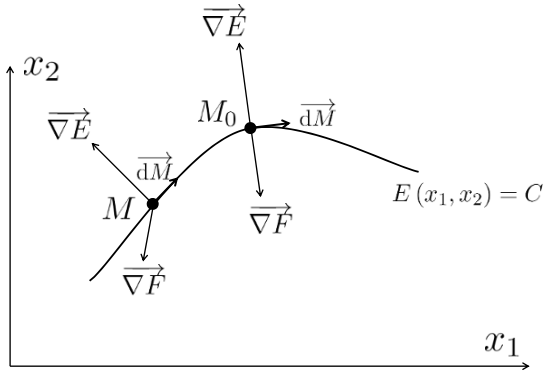


Figure 1: When the constraint $E(x_1, x_2) = C$ is satisfied, the point M, with coordinates x_1 and x_2, is forced to move along a curve in the plane, shown as a solid line. The tangent to this curve is perpendicular to the gradient $\overrightarrow{\nabla E}$ of the function E, meaning that any small displacement of point M along the curve must be perpendicular to this gradient. When the displacement starts from an arbitrary point M, the vectors $\overrightarrow{\nabla E}$ and $\overrightarrow{\nabla F}$ are not parallel, and the function F varies to first order in $dF = \overrightarrow{\nabla F} \cdot \overrightarrow{dM}$. However, if the variation starts from a point, such as M_0, where the two gradients are parallel, the function F is stationary. Geometrically, this parallelism means that the solid line is tangent to a contour line of the surface representing the function $F(x_1, x_2)$.

As in general the vectors $\overrightarrow{\nabla E}$ and $\overrightarrow{\nabla F}$ are not parallel, this scalar product is not zero. The function F thus varies to first order in $\overrightarrow{\mathrm{d}M}$, meaning it is not stationary at that point.

If, however, we start from a point M_0 on the curve where the two gradients are parallel (or antiparallel), condition (4) means that the variation (5) is zero, and stationarity is attained. In such a case, M moves (at constant E) along a curve that is tangent at M_0 to a contour line of the surface representing the function F. Geometrically, it is easy to understand that a displacement along a contour line keeps F constant to first order. Algebraically, imposing the gradients to be parallel amounts to writing that there exists a constant λ, called "Lagrange multiplier", such that:

$$\overrightarrow{\nabla F} = \lambda \, \overrightarrow{\nabla E} \tag{6}$$

which is equivalent to saying that the differential of the function $F - \lambda E$ is zero:

$$\begin{aligned} \mathrm{d}(F - \lambda E) &= \overrightarrow{\nabla}(F - \lambda E) \cdot \overrightarrow{\mathrm{d}M} \\ &= \frac{\partial [F - \lambda E]}{\partial x_1} \mathrm{d}x_1 + \frac{\partial [F - \lambda E]}{\partial x_2} \mathrm{d}x_2 \\ &= 0 \end{aligned} \tag{7}$$

This means that one must simply replace the function F by the fonction $F - \lambda E$ with an arbitrary Lagrange multiplier λ to obtain the stationarity of F when its variables obey the constraint (3).

When λ is fixed, we get as before two equations with two unknowns, so that the variables x_1 and x_2 are determined. Inserting them into (3) yields a value for C, which is thus fixed. If, however, the Lagrange multiplier λ is allowed to vary, the constant C becomes a function of λ, and can be adjusted by changing λ. As an example, when studying the canonical equilibrium (Appendix VI, § 1-b), one maximizes the value of the entropy S (which plays the role of the function F) while keeping the average energy value E constant. A Lagrange multiplier β is then introduced to impose the stationarity of $S - \beta E$; changing β allows controlling the value of E.

2. Function of N variables

We now consider a function $F(x_1, x_2, ..., x_N)$ of N supposedly independent variables x_1, x_2,... ,x_N. The extrema of F are obtained by annulling the N components of the gradient of F (each component being the partial derivative of F with respect to one of the variables):

$$\overrightarrow{\nabla F} = 0 \tag{8}$$

We get N equations to determine N unknown variables, yielding a finite number of extrema.

Imagine now that the N variables are no longer independent, but linked by $P < N$ conditions:

$$E_p(x_1, x_2, ..., x_N) = C_p \qquad \text{with } p = 1, 2, ..., P \tag{9}$$

Consider a point M in an N-dimensional space, with coordinates $x_1, x_2, ..., x_N$. If these N coordinates satisfy the P conditions (9), their infinitesimal variations obey the P relations:

$$\overrightarrow{\nabla E_p} \cdot \overrightarrow{\mathrm{d}M} = 0 \qquad \text{with } p = 1, 2, ..., P \tag{10}$$

For all the functions E_p to remain constant, the displacement $\overrightarrow{\mathrm{d}M}$ of point M in the N-dimensional space must be orthogonal to all the gradients $\overrightarrow{\nabla E_p}$. Two cases are then possible:

(i) either the gradient $\overrightarrow{\nabla F}$ belongs to the sub-space generated by the $\overrightarrow{\nabla E_p}$, in which case the orthogonality condition (10) implies that $\overrightarrow{\mathrm{d}M}$ is also orthogonal to $\overrightarrow{\nabla F}$. Consequently the variation $\mathrm{d}F = \overrightarrow{\nabla E} \cdot \overrightarrow{\mathrm{d}M}$ is zero and the stationarity is ensured.

(ii) or the gradient $\overrightarrow{\nabla F}$ is not contained in that subspace, and it possesses a non-zero component $\overrightarrow{\nabla F}_\perp$ orthogonal to that subspace. One can then choose $\overrightarrow{\mathrm{d}M}$ parallel to $\overrightarrow{\nabla F}_\perp$ and obtain a first order variation of F, while satisfying the P constraints.

In conclusion, the stationarity of F is equivalent to the condition that the gradient $\overrightarrow{\nabla F}$ be contained in the subspace generated by the $\overrightarrow{\nabla E_p}$. This amounts to stating that there exist Lagrange multipliers λ_p (with $p = 1, 2, ..., P$) such that:

$$\overrightarrow{\nabla F} = \lambda_1 \overrightarrow{\nabla E_1} + \lambda_2 \overrightarrow{\nabla E_2} + ... + \lambda_P \overrightarrow{\nabla E_P} \tag{11}$$

In an equivalent way, the stationarity condition can be obtained by annulling the differential:

$$\mathrm{d}\left(F - \lambda_1 E_1 - \lambda_2 E_2 - ... - \lambda_P E_P\right) = 0 \tag{12}$$

and then treating the variables $x_1, x_2, ..., x_N$ as if they were independent.

When the Lagrange multipliers λ_p are fixed, each component of relation (11) yields an equation, so that we have as many equations as variables $x_1, x_2, ..., x_N$. One therefore obtains for the function F a finite number of extrema linked by the constraints, yielding fixed values for the functions E_1, E_2, ..., E_P. A variation in the Lagrange multipliers will change the values of these functions, which can therefore be adjusted to a value that has been chosen in advance.

Appendix VI

Brief review of Quantum Statistical Mechanics

1 Statistical ensembles 2285
1-a Microcanonical ensemble 2285
1-b Canonical ensemble 2289
1-c Grand canonical ensemble 2291
2 Intensive or extensive physical quantities 2292
2-a Microcanonical ensemble 2293
2-b Canonical ensemble 2294
2-c Grand canonical ensemble 2295
2-d Other ensembles 2295

In quantum mechanics, as in classical mechanics, it is not possible to describe a system having a very large number of degrees of freedom (for example a system containing a number of particles that is of the order of the Avogadro number) with highest precision. Such a description would in particular include the value of many quantities, for instance many particle correlations, which fluctuate rapidly and are not necessarily of interest. A less detailed and more probabilistic description must be used, in which the state of the system in known only statistically. The system occupies one on a series of possible states, with a certain probability. One then says that the system is described by a "statistical ensemble". The use of a density operator (Complement E_{III}) to describe the physical system is particularly convenient in this case.

We do not attempt in this appendix to give a general introduction to statistical mechanics and its postulates. We simply summarize a number of quantum statistical mechanics results used in several complements. For example, most of the complements of Chapter XV, as well as B_{XVII} and D_{XVII}, use the concept of chemical potential μ or of "grand potential" Φ; their interpretation in the framework of the different statistical ensembles will be given in this appendix.

1. Statistical ensembles

Several "statistical ensembles" are commonly used to describe physical systems at equilibrium. We shall focus here on the three main ones: the microcanonical, the canonical and the grand canonical ensembles. The first of these ensembles provides the general setting for introducing the two others.

1-a. Microcanonical ensemble

Consider a physical system containing N particles in a box of volume $\mathcal{V}$. The energy E' of the system lies within an interval:

$$E - \Delta E/2 \leq E' \leq E + \Delta E/2 \tag{1}$$

with $\Delta E \ll E$. The system is isolated from its surroundings preventing any exchange of particles or energy. We note $|E', q\rangle$ the eigenstates of its Hamiltonian $\widehat{H}$, where q is an index reflecting the possible degeneracy of each eigenvalue E' of this Hamiltonian.

α. Density operator, entropy

The system is supposed to have the same probability of being in any state whose energy falls within the interval (1); no state is favored over any other. The microcanonical density operator of the system at equilibrium is then:

$$\rho_{\text{eq}} = \frac{1}{Z} P_{\Delta E}(E) \tag{2}$$

where $P_{\Delta E}(E)$ is the projector onto the subspace containing all the accessible states:

$$P_{\Delta E}(E) = \sum_{E'=E-\frac{\Delta E}{2}}^{E+\frac{\Delta E}{2}} \sum_{q} |E', q\rangle \langle E', q| \tag{3}$$

and where Z is the microcanonical partition function defined as:

$$Z = \text{Tr}\{P_{\Delta E}(E)\} \tag{4}$$

What relation (2) means is that the occupation probabilities of the states $|E', q\rangle$ are all equal to $1/Z$. Relation (3) shows that each of the projectors onto a state $|E', q\rangle$ contributes one unit to the trace of relation (4); the partition function Z is simply the number of terms in the summation (3), i.e. the number of levels in the energy interval (1). If $D(E)$ is the density of states, we can write:

$$Z = D(E) \times \Delta E \tag{5}$$

As in § 1 of Complement A_{XXI}, we define the entropy S as:

$$S = -k_B \text{ Tr}\{\rho_{\text{eq}} \ln \rho_{\text{eq}}\} \tag{6}$$

where k_B is the Boltzmann constant. The $|E', q\rangle$ are eigenvectors of ρ_{eq}, with an eigenvalue equal to $1/Z$ if E' belongs to the interval $[E - \frac{\Delta E}{2}, E + \frac{\Delta E}{2}]$, and equal to zero otherwise. If E' belongs to the interval, we have:

$$\rho_{\text{eq}} \ln \rho_{\text{eq}} |E', q\rangle = -\frac{\ln Z}{Z} |E', q\rangle \tag{7}$$

If E' does not belong to the interval, since $\lim_{x \to 0} x \ln x = 0$, we get:

$$\rho_{\text{eq}} \ln \rho_{\text{eq}} |E', q\rangle = 0 \tag{8}$$

We now multiply the two previous relations by the bra $\langle E', q|$ and sum over E' and q to get a trace. Only the bras whose energy falls within the interval will yield a non-zero contribution. As there are Z of them, we obtain:

$$\text{Tr}\{\rho_{\text{eq}} \ln \rho_{\text{eq}}\} = -\ln Z \tag{9}$$

The equilibrium value of the entropy is therefore:

$$S = k_B \ln Z \tag{10}$$

β. Temperature, chemical potential

Suppose we now change the equilibrium energy E by an infinitesimal amount $\mathrm{d}E$, keeping the volume $\mathcal{V}$ and the particle number N constant. Since no work is exchanged with the outside (the external walls are fixed), this amounts solely to a heat change:

$$\mathrm{d}Q = \mathrm{d}E = T\mathrm{d}S \tag{11}$$

where we have used the usual thermodynamic definition of the entropy $\mathrm{d}S = \mathrm{d}Q/T$. In the microcanonical ensemble, the temperature is thus defined as:

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{N,\mathcal{V}} \tag{12}$$

where we have used the partial derivative notation to emphasize that the changes are made keeping both N and $\mathcal{V}$ constant.

Let us now change the particle number N, keeping the volume and the energy E constant. We then define the "chemical potential" μ (which has the dimension of an energy) as:

$$\mu = -T \left.\frac{\partial S}{\partial N}\right|_{E,\mathcal{V}} \tag{13}$$

At fixed temperature, the faster the entropy (which depends on the number of accessible levels in the energy band ΔE) grows with the particle number N, the larger the absolute value of μ. The chemical potential plays an essential role in the grand canonical equilibrium as we shall see (§ 1-c). The third partial derivative of S (with respect to the volume) will be determined in § 2-a.

Comment:

Let us insert (5) in relation (10), but first multiplying $D(E)$ by $k_B T$ and dividing ΔE by this same quantity (this has the advantage of providing dimensionless arguments for the logarithmic functions). This yields:

$$\frac{S}{k_B} = \ln\left[D\left(E\right) k_B T\right] + \ln\left[\frac{\Delta E}{k_B T}\right] \tag{14}$$

In a macroscopic system, the particle number is very large, of the order of the Avogadro number. Let us see then what happens when the particle number N goes to infinity. We assume that the energy E as well as the volume are proportional to N (thermodynamic limit). We then expect the entropy to also be proportional to N. This linear variation of the entropy cannot come from the second term in (14): even if the energy interval ΔE is proportional to N, it will only yield a much slower logarithmic variation. Most of the variation of S actually comes from the first term of (14), and from the fact that the density of states increases with N in an exponential way: as the exponent of $D(E)$ contains N, this variation is phenomenally rapid. In the limit of large systems, the first term in (14) largely dominates the second. This is why it is often said that the entropy characterizes the density of states of a physical system (or more precisely the number of its quantum energy levels in a microscopic energy interval, chosen here to be equal to $k_B T$).

γ. Entropy maximization

We now choose for ρ an arbitrary Hermitian density operator, with positive or zero eigenvalues whose sum is equal to 1. We denote $|\theta_n\rangle$ its eigenvectors and p_n its eigenvalues ($0 \leq p_n \leq 1$) which obey:

$$\sum_n p_n = 1 \tag{15}$$

We assume that ρ is restricted to the energy band (1): all the $|\theta_n\rangle$ for which $p_n \neq 0$ are arbitrary linear combinations of the eigenvectors $|E, q\rangle$ obeying (1).

An entropy can be associated with ρ:

$$S_\rho = -k_B \ \mathrm{Tr}\left\{\rho \ \ln\rho\right\} \tag{16}$$

where k_B is the Boltzmann constant. We are going to show that among all possible ρ operators, the equilibrium one, ρ_{eq}, maximizes this entropy. We can write:

$$\rho = \sum_n p_n \ \ |\theta_n\rangle \langle\theta_n| \qquad\qquad \ln\rho = \sum_n \ln p_n \ \ |\theta_n\rangle \langle\theta_n| \tag{17}$$

and therefore:

$$S_\rho = -k_B \ \mathrm{Tr}\left\{\sum_n p_n \ \ln p_n \ \ |\theta_n\rangle \langle\theta_n|\right\} = -k_B \sum_n p_n \ \ln p_n \tag{18}$$

Any variation of the p_n results in a variation of S written as:

$$\mathrm{d}S = -k_B \sum_n \left[1 + \ln p_n\right] \ \mathrm{d}p_n \tag{19}$$

However relation (15) requires the sum of the variations $\mathrm{d}p_n$ to be zero. To write that relation (19) is zero while taking into account this constraint, we use a Lagrange multiplier λ (Appendix V) and obtain the equation:

$$\sum_n \left[\lambda + 1 + \ln p_n\right] \ \mathrm{d}p_n = 0 \tag{20}$$

which must be satisfied for any $\mathrm{d}p_n$. Canceling the corresponding coefficients leads to:

$$\ln \ p_n = -\lambda - 1 \tag{21}$$

This means that all the non-zero p_n must be equal. Operator ρ is therefore proportional to the projector (3). Once we normalize its trace to 1, we get (2): the microcanonical density operator corresponds to an entropy extremum. As all the p_n are between 0 and 1, relation (18) shows that this extremum is positive. To find out if it is a maximum or a minimum, we consider another ρ operator, whose eigenvalues are all zero except one, equal to 1; its associated entropy S_ρ is zero. Consequently the extremum of S_ρ obtained for the microcanonical equilibrium is an absolute maximum.

This result proves an important theorem: the density operator that maximizes the entropy is the sum of the projectors onto all the accessible states, with equal eigenvalues (the probabilities of finding the physical system in each of these states).

1-b. Canonical ensemble

We now consider a physical system $\mathscr{S}$ no longer isolated but in contact with a reservoir $\mathcal{R}$ with which it exchanges energy; for example $\mathscr{S}$ and $\mathcal{R}$ could be coupled through a wall conducting heat but remaining fixed so that no work can be exchanged between them. Let us call $N_{\mathcal{R}}$, $E_{\mathcal{R}}$ and $\mathcal{V}_{\mathcal{R}}$ the particle number, the energy and the volume of the reservoir $\mathcal{R}$.

α. Density operator

Assuming the reservoir $\mathcal{R}$ to be much larger than the system $\mathscr{S}$, its temperature remains constant as it exchanges energy with $\mathscr{S}$. According to relation (12), this implies that its temperature $T_{\mathcal{R}}$, defined as:

$$\frac{1}{T_{\mathcal{R}}} = \left.\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\right|_{N_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} \tag{22}$$

is a constant. It will be characterized by the constant β:

$$\beta = \frac{1}{k_B T_{\mathcal{R}}} = \frac{1}{k_B}\left.\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\right|_{N_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} \tag{23}$$

As relation (10) showed that $S_{\mathcal{R}} = k_B \ln Z_{\mathcal{R}}$, where $Z_{\mathcal{R}}$ is the number of the reservoir's accessible levels in an energy band ΔE around $E_{\mathcal{R}}$, we deduce:

$$\left.\frac{\partial \ln Z_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\right|_{N_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} = \beta \tag{24}$$

This means that this number of levels varies exponentially as a function of the energy (keeping $N_{\mathcal{R}}$ and $\mathcal{V}_{\mathcal{R}}$ constant):

$$Z_{\mathcal{R}} \propto e^{\beta E_{\mathcal{R}}} \tag{25}$$

The total system $\mathscr{S} + \mathcal{R}$ is described by an equilibrium microcanonical density operator. Its energy eigenvectors are the tensor product[1] of the energy eigenvectors $|E, q\rangle$ of the system $\mathscr{S}$ and the energy eigenvectors $|E_{\mathcal{R}}, q_{\mathcal{R}}\rangle$ of the system $\mathcal{R}$:

$$|E, q\rangle \otimes |E_{\mathcal{R}}, q_{\mathcal{R}}\rangle \tag{26}$$

The microcanonical density operator of $\mathscr{S} + \mathcal{R}$, with a total energy $E_{\text{tot}} = E' + E'_{\mathcal{R}}$ is given by:

$$\rho_{\mathscr{S}+\mathcal{R}} = \frac{1}{Z_{\mathscr{S}+\mathcal{R}}} \sum_{E'+E'_{\mathcal{R}}=E_{\text{tot}}-\frac{\Delta E}{2}}^{E_{\text{tot}}+\frac{\Delta E}{2}} \sum_{q,q_{\mathcal{R}}} \Big(|E', q\rangle \otimes |E'_{\mathcal{R}}, q_{\mathcal{R}}\rangle \Big) \Big(\langle E', q| \otimes \langle E'_{\mathcal{R}}, q_{\mathcal{R}}| \Big) \tag{27}$$

We get the density operator ρ_{eq} of the system $\mathscr{S}$ by taking a trace over the reservoir:

$$\rho_{\text{eq}} = \text{Tr}_{\mathcal{R}} \{\rho_{\mathscr{S}+\mathcal{R}}\} \tag{28}$$

[1] We assume that the coupling between $\mathscr{S}$ and $\mathcal{R}$ is weak, so that its contribution to the total energy is negligible.

In (27), the trace over $\mathcal{R}$ of each projector $|E'_{\mathcal{R}}, q_{\mathcal{R}}\rangle \langle E'_{\mathcal{R}}, q_{\mathcal{R}}|$ is just equal to one. The density operator ρ_{eq} is simply a sum of projectors onto the energy eigenstates $|E', q\rangle \langle E', q|$, multiplied by the number of levels of $\mathcal{R}$ with an energy $E_{\text{tot}} - E'$ within an energy band ΔE. Relation (25) shows that this number of levels varies exponentially as $e^{\beta E_{\mathcal{R}}} = e^{\beta\left(E_{\text{tot}} - E'\right)}$. Omitting the proportionality factors $1/Z_{\mathscr{S}+\mathcal{R}}$ and $e^{\beta E_{\text{tot}}}$, we get:

$$\rho_{\text{eq}} \propto \sum_{E'} \sum_{q} e^{-\beta E'} |E', q\rangle \langle E', q| = e^{-\beta \widehat{H}} \tag{29}$$

where $\widehat{H}$ is the Hamiltonian of the system $\mathscr{S}$. Normalizing the trace of ρ, we obtain:

$$\boxed{\rho_{\text{eq}} = \frac{1}{Z_c}\, e^{-\beta \widehat{H}}} \tag{30}$$

where Z_c is the "canonical partition function" defined as:

$$\boxed{Z_c = \text{Tr}\left\{e^{-\beta \widehat{H}}\right\}} \tag{31}$$

These two relations define the density operator of $\mathscr{S}$ in the canonical thermal equilibrium. Contrary to what happened in the microcanonical equilibrium, the energy of the system $\mathscr{S}$ is no longer restricted to a small interval ΔE, but may spontaneously fluctuate outside this energy band under the effect of the coupling with the reservoir.

The thermodynamic potential of the canonical equilibrium is defined by the function F called the "free energy":

$$F = \left\langle \widehat{H} \right\rangle - TS \tag{32}$$

At equilibrium, when ρ is given by (30), this free energy is equal to:

$$F = \left\langle \widehat{H} \right\rangle + k_B T\ \text{Tr}\left\{\rho_{\text{eq}}\ \left[-\ln Z_c - \beta \widehat{H}\right]\right\} = \left\langle \widehat{H} \right\rangle - \left\langle \widehat{H} \right\rangle - k_B T\ \ln Z_c \tag{33}$$

and we obtain:

$$\boxed{F = -k_B T\ \ln Z_c} \tag{34}$$

β. Minimization of the free energy

Starting from an arbitrary density operator ρ of unit trace, let us show that its associated free energy will be minimal when ρ is equal to its value at the canonical equilibrium (30). We first compute the variation of F:

$$\mathrm{d}F = \text{Tr}\left\{\left[\widehat{H} + k_B T\,(1 + \ln\rho)\right] \mathrm{d}\rho\right\} \tag{35}$$

This variation is zero for any $\mathrm{d}\rho$ only if the operator between the inner brackets is zero, which means:

$$\ln\ \rho = -1 - \beta \widehat{H} \tag{36}$$

This indicates that $\rho \propto e^{-\beta \widehat{H}}$, which is the canonical equilibrium operator. Finally, if we choose for ρ the projector onto a state having a large positive energy, S will be zero, $\left\langle \widehat{H} \right\rangle$ arbitrarily very large, and consequently F will be very large as well. It is thus clear that the extremum of F, which occurs when ρ takes the equilibrium value, is a minimum.

1-c. Grand canonical ensemble

We now assume that the physical system $\mathscr{S}$ can exchange not only energy but also particles with the reservoir $\mathcal{R}$: $\mathscr{S}$ and $\mathcal{R}$ must be coupled through an interface the particles can cross. As above, this reservoir is supposed large enough for its temperature to remain constant when it exchanges energy with $\mathscr{S}$. We also assume it contains a very large number of particles, which barely changes in relative value during the particle exchanges with $\mathscr{S}$. Its chemical potential therefore remains constant:

$$\mu = -T \left.\frac{\partial S_{\mathcal{R}}}{\partial N}\right|_{E_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} = \text{constant} \tag{37}$$

α. Density operator

As the particle numbers N of system $\mathscr{S}$ and N_R of the reservoir are no longer constant, their state spaces are now Fock spaces (Chapter XV). We must add indices N and $N_{\mathcal{R}}$ respectively to the kets $|E, q\rangle$ and $|E_{\mathcal{R}}, q_{\mathcal{R}}\rangle$ and there are now two summations over these indices in the expression for the microcanonical density operator $\rho_{\mathscr{S}+\mathcal{R}}$ written in (27). The microcanonical density operator of the total system is then:

$$\rho_{\mathscr{S}+\mathcal{R}} = \frac{1}{Z_{\mathscr{S}+\mathcal{R}}} \sum_{E'+E'_{\mathcal{R}}=E_{\rm tot}-\frac{\Delta E}{2}}^{E_{\rm tot}+\frac{\Delta E}{2}} \sum_{N'+N'_{\mathcal{R}}=N_{\rm tot}} \sum_{q,q_{\mathcal{R}}} \Big(|E', N', q\rangle \otimes |E'_{\mathcal{R}}, N'_{\mathcal{R}}, q_{\mathcal{R}}\rangle \Big) \Big(\langle E', N', q| \otimes \langle E'_{\mathcal{R}}, N'_{\mathcal{R}}, q_{\mathcal{R}}| \Big) \tag{38}$$

where $E_{\rm tot}$ and $N_{\rm tot}$ are respectively the energy and the particle number of the total system $\mathscr{S} + \mathcal{R}$. The argument then follows the same lines as in § 1-b-α. A partial trace over the reservoir leads to the density operator of $\mathscr{S}$, which is a linear combination of projectors:

$$|E', N', q\rangle \langle E', N', q| \tag{39}$$

with weights corresponding to the number of states of the reservoir in an energy band centered around $E'_{\mathcal{R}} = E_{\rm tot} - E'$, the number of particles in the reservoir being $N'_{\mathcal{R}} = N_{\rm tot} - N'$. Two reservoir variables change simultaneously, instead of one for the canonical equilibrium. As μ and β remain constant in (24) and (37), the entropy $S_{\mathcal{R}}$ varies linearly with respect to these variables:

$$S_{\mathcal{R}} = S^0_{\mathcal{R}} + \frac{1}{T} E'_{\mathcal{R}} - \frac{\mu}{T} N'_{\mathcal{R}} = S^0_{\mathcal{R}} + k_B \beta \left(E'_{\mathcal{R}} - \mu N'_{\mathcal{R}}\right) \tag{40}$$

where $S^0_{\mathcal{R}}$ is a constant that is of no importance in what follows. Using again relation (10) to relate the reservoir entropy to the number of states $Z_{\mathcal{R}}$ accessible to this reservoir, we get:

$$Z_{\mathcal{R}} = e^{S_{\mathcal{R}}/k_B} \propto e^{\beta\left(E'_{\mathcal{R}} - \mu N'_{\mathcal{R}}\right)} = e^{\beta(E_{\rm tot} - \mu N_{\rm tot})} \times e^{-\beta\left(E' - \mu N'\right)} \tag{41}$$

where μ and β remain constant. The same argument as above shows that the trace over the reservoir variables lead to the following density operator ρ for the system $\mathscr{S}$:

$$\boxed{\rho_{\text{eq}} = \frac{1}{Z_{gc}}\ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)}} \tag{42}$$

with:

$$\boxed{Z_{\text{gc}} = \text{Tr}\left\{e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)}\right\}} \tag{43}$$

where $\widehat{N}$ is total particle number operator of $\mathscr{S}$.

β. Grand potential

The thermodynamic potential for the grand canonical ensemble is the "grand potential" Φ defined as:

$$\Phi = \langle\widehat{H}\rangle - TS - \mu\langle\widehat{N}\rangle \tag{44}$$

Following the same demonstration as for F in the canonical ensemble, we can show that the equilibrium value of this potential is:

$$\boxed{\Phi = -k_B T\ \ln Z_{gc}} \tag{45}$$

If we let ρ vary, we can show, as above, that this potential reaches a minimum when ρ is equal to (42); a detailed demonstration is given in § 1 of Complement G_{XV}.

2. Intensive or extensive physical quantities

Take a macroscopic physical system $\mathscr{S}$ at equilibrium, and divide it into two subsystems of equal sizes $\mathscr{S}'$ and $\mathscr{S}''$; one can imagine that a wall separates $\mathscr{S}'$ from $\mathscr{S}''$. Certain physical quantities associated with $\mathscr{S}'$ or $\mathcal{S}''$, taken separately, are half of what they were for $\mathscr{S}$: the volumes, the energies, the particle numbers, the entropies, etc. Such quantities are said to be "extensive". Inversely, other physical quantities do not change upon this division: the particle number per unit volume, the temperature, the chemical potential, etc. Such quantities are said to be "intensive". In a general way, when a macroscopic physical system of volume $\mathcal{V}$ is divided into several macroscopic parts of volumes $\mathcal{V}_1$, $\mathcal{V}_2$, etc., the physical quantities measured in each part and which are proportional to their respective volume are said to be extensive, and those which remain constant are said to be intensive.

As for the ensembles studied above, their description involves a mixture[2] of extensive and intensive variables:

(i) In the microcanonical ensemble, the three independent variables describing the physical system at equilibrium are the three extensive variables $\mathcal{V}$, N and the system's energy E; the other physical quantities (temperature, entropy, chemical potential, etc.) are

[2]including at least one extensive variable, otherwise the system's size would not be determined.

considered functions of these variables. The thermodynamic potential is the entropy S, extensive and directly related to the logarithm of the microcanonical partition function.

(ii) In the canonical ensemble, the three independent variables include two extensive variables $\mathcal{V}$ and N as well as an intensive variable T (or β). The thermodynamic potential is the free energy F, an extensive function directly related to the logarithm of the canonical partition function.

(iii) In the grand canonical ensemble, there is only one independent extensive variable, the volume $\mathcal{V}$, and two intensive variables T and μ. The thermodynamic potential is the function Φ, extensive and directly related to the logarithm of the grand canonical partition function.

For a macroscopic system, the three ensembles are generally considered equivalent. The statistical descriptions are, however, different. In the canonical equilibrium, for example, the energy is not restricted to an interval ΔE but can fluctuate and take on values outside this interval. However, for a macroscopic system, the fluctuations in energy are very small compared to its average value. Assuming that the system's energy is confined within a fixed band ΔE is a valid approximation and allows taking for $\langle \hat{H} \rangle$ the microcanonical energy E:

$$\langle \hat{H} \rangle = E \tag{46}$$

Another example, in the grand canonical ensemble, is the particle number, which fluctuates around its average value $\langle \hat{N} \rangle$. For a macroscopic system, the relative value of these fluctuations is in general[3] very small, and the average value $\langle \hat{N} \rangle$ is practically equal to the particle number N of the microcanonical or canonical ensembles:

$$\langle \hat{N} \rangle = N \tag{47}$$

2-a. Microcanonical ensemble

Relations (12) and (13) give the partial derivatives of the entropy S with respect to the variables E and N; we now compute that derivative with respect to the volume $\mathcal{V}$.

Let us change the physical system volume $\mathcal{V}$ by a small quantity $\mathrm{d}\mathcal{V}$, keeping the particle number N constant, and without any heat exchange (the system is surrounded by isolating walls). The system, having an internal pressure P, is doing the work $P\mathrm{d}\mathcal{V}$, which means that its internal energy varies as:

$$\mathrm{d}E = -P\mathrm{d}\mathcal{V} \tag{48}$$

As there is no heat exchange, $\mathrm{d}Q = 0$, and the thermodynamic relation $\mathrm{d}Q = T\mathrm{d}S$ means that the entropy does not change either:

$$\mathrm{d}S = \left.\frac{\partial S}{\partial \mathcal{V}}\right|_{N,E} \mathrm{d}\mathcal{V} + \left.\frac{\partial S}{\partial E}\right|_{N,\mathcal{V}} \mathrm{d}E = 0 \tag{49}$$

[3]There are exceptions to this rule: for a Bose-condensed ideal gas, the grand canonical fluctuations of the particles' number remain large for a macroscopic system. This is a very special system for which the canonical and grand canonical ensembles are not equivalent for certain physical properties.

Inserting relation (12) in this result and multiplying by T, we obtain:

$$T \left.\frac{\partial S}{\partial \mathcal{V}}\right|_{N,E} \mathrm{d}\mathcal{V} + \mathrm{d}E = 0 \tag{50}$$

As relation (48) shows that the pressure P is given by $-\mathrm{d}E/\mathrm{d}\mathcal{V}$ and taking (10) into account, we finally get:

$$P = T \left.\frac{\partial S}{\partial \mathcal{V}}\right|_{N,E} = k_B T \left.\frac{\partial \ln Z}{\partial \mathcal{V}}\right|_{N,E} \tag{51}$$

which defines the pressure in the microcanonical ensemble.

We already studied, in § 1-a-β, the entropy changes due to variations of either E (keeping $\mathcal{V}$ and N constant) or N (keeping E and $\mathcal{V}$ constant). The present calculation is the last step for obtaining the three partial derivatives of the microcanonical thermodynamic potential, and we can express its total derivative as:

$$\boxed{\mathrm{d}S = \frac{1}{T}\mathrm{d}E + \frac{P}{T}\mathrm{d}\mathcal{V} - \frac{\mu}{T}\mathrm{d}N} \tag{52}$$

2-b. Canonical ensemble

For a macroscopic system, we just saw that $\langle H \rangle$ could be replaced by the microcanonical energy E in the definition (32) of the free energy. Taking the differential of (32) then leads to:

$$\mathrm{d}F = \mathrm{d}E - T\mathrm{d}S - S\mathrm{d}T \tag{53}$$

Using (52) in the $T\mathrm{d}S$ term of this equation, the $\mathrm{d}E$ terms cancel out and we get:

$$\boxed{\mathrm{d}F = -S\mathrm{d}T - P\mathrm{d}\mathcal{V} + \mu \mathrm{d}N} \tag{54}$$

This is the total differential of the thermodynamic potential in the canonical ensemble.

This relation allows a physical interpretation of the chemical potential: it is the gain in free energy when one particle is added to the system[4], keeping constant the temperature and the volume of the system. As for the pressure P, it is given by:

$$P = -\left.\frac{\partial F}{\partial \mathcal{V}}\right|_{N,T} \tag{55}$$

or, using (34):

$$P = k_B T \left.\frac{\partial \ln Z_c}{\partial \mathcal{V}}\right|_{N,T} \tag{56}$$

which is similar to (51). We have obtained the pressure of the physical system as a function of its volume and its temperature, i.e. its "equation of state".

[4]When the temperature is zero, the free energy is just the energy E, and μ is the increase of energy when one particle is added.

To compute the average energy $\left\langle \widehat{H} \right\rangle$, we can use relations (30) and (31), which yields:

$$\left\langle \widehat{H} \right\rangle = \frac{1}{Z_c} \mathrm{Tr} \left\{ \widehat{H}\ e^{-\beta \widehat{H}} \right\} = \frac{-1}{Z_c} \frac{\partial}{\partial \beta} \mathrm{Tr} \left\{ e^{-\beta \widehat{H}} \right\} \tag{57}$$

where the partial derivative is taken keeping N and $\mathcal{V}$ constant. We then have:

$$\left\langle \widehat{H} \right\rangle = -\frac{1}{Z_c} \frac{\partial Z_c}{\partial \beta} = -\frac{\partial \ln Z_c}{\partial \beta} \tag{58}$$

.

2-c. Grand canonical ensemble

In a macroscopic system, as the particle number generally fluctuates very little in relative value, we can replace $\langle N \rangle$ by N in the definition (44) of the thermodynamic grand potential Φ. This leads to:

$$\mathrm{d}\Phi = \mathrm{d}F - \mu \mathrm{d}N - N \mathrm{d}\mu \tag{59}$$

Using (54) in this relation, the $\mu \mathrm{d}N$ terms cancel out and we are left with:

$$\boxed{\mathrm{d}\Phi = -S\mathrm{d}T - P\mathrm{d}\mathcal{V} - N\mathrm{d}\mu} \tag{60}$$

In this ensemble, the volume $\mathcal{V}$ is the only extensive variable. For a fixed temperature and chemical potential, and for a large volume, we get a macroscopic system whose energy, entropy and particle number are proportional to $\mathcal{V}$. We simply get:

$$\Phi = -P\mathcal{V} \tag{61}$$

The grand potential divided by $\mathcal{V}$ yields the pressure directly, without any partial derivative.

Taking (45) into account, the average particle number and the pressure obey:

$$\begin{aligned} N &= -\left. \frac{\partial \Phi}{\partial \mu} \right|_{\mathcal{V},T} = k_B T \left. \frac{\partial \ln Z_{gc}}{\partial \mu} \right|_{\mathcal{V},T} \\ P &= -\frac{\Phi}{\mathcal{V}} = \frac{k_B T}{\mathcal{V}} \ln Z_{gc} \end{aligned} \tag{62}$$

Using these two equalities to eliminate the chemical potential μ, we get the particle number in a given volume as a function of the pressure P and the temperature T (equation of state for the physical system).

2-d. Other ensembles

We have studied the three most commonly used statistical ensembles, but there are others, as for example the isothermal-isobaric ensemble. In this ensemble, the system $\mathscr{S}$ is coupled with a reservoir allowing exchanges of energy and volume, but not particles; the number N remains fixed. The only extensive variable is precisely this variable N,

the other two, the temperature T and the pressure P, being intensive. The thermodynamic potential associated with this isothermal-isobaric ensemble is the Gibbs function G defined as:

$$G = E + P\mathcal{V} - TS \tag{63}$$

As before, we take the differential of this function and note, using (52), that the terms $\mathrm{d}E$ and $P\mathrm{d}\mathcal{V}$ cancel out. We then get:

$$\mathrm{d}G = \mathcal{V}\mathrm{d}P - S\mathrm{d}T + \mu\mathrm{d}N \tag{64}$$

The function G is extensive. It increases as the particle number N gets larger (for fixed pressure and temperature), and for a macroscopic system it is proportional to the system size:

$$G = \mu N \tag{65}$$

Varying both the temperature and the pressure of an ensemble of N particles, we can get the resulting variation of the chemical potential μ by dividing (64) by N and then setting $\mathrm{d}N = 0$. This yields the Gibbs-Duhem relation:

$$\boxed{\mathrm{d}\mu = \frac{\mathcal{V}}{N}\,\mathrm{d}P - \frac{S}{N}\,\mathrm{d}T} \tag{66}$$

This ensemble is particularly useful in the study of a two-phase equilibrium such as a liquid and its vapor, both at the same pressure and temperature.

We have presented a brief review of the general principles of statistical mechanics. For more details, the reader may consult, for example, the following references [95, 96, 97, 3].

Appendix VII

Wigner transform

1 Delta function of an operator 2299

2 Wigner distribution of the density operator (spinless particle) 2299

2-a Definition of the distribution, Weyl operators 2300

2-b Expressions for the Wigner transform 2301

2-c Reality, normalization, operator form 2304

2-d Gaussian wave packet 2305

2-e Semiclassical situations 2306

2-f Quantum situations where the Wigner distribution is not a probability distribution 2309

3 Wigner transform of an operator 2310

3-a Average value of a Hermitian operator (observable) 2311

3-b Special cases 2312

3-c Wigner transform of an operator product 2313

3-d Evolution of the density operator 2316

4 Generalizations 2318

4-a Particle with spin 2318

4-b Several particles 2319

5 Discussion: Wigner distribution and quantum effects 2319

5-a An interference experiment 2319

5-b General discussion; "ghost" component 2322

Introduction

In classical mechanics, it is possible to specify with an arbitrary precision both the position $\mathbf{r}$ and the momentum $\mathbf{p}$ (and hence the velocity) of a particle. If the state of the particle is defined in a statistical way, one describes the classical particle by a distribution $\rho_{\text{cl}}(\mathbf{r}, \mathbf{p})$ in the phase space (Appendix III, § 3-a) , which can be any positive function, normalized to unity. This distribution can, for example, include correlations between the particle's position and velocity. In quantum mechanics, the situation is different. It is true that one often uses two representations, one in position space, the other in momentum space, and that we go from one to the other via a Fourier transformation. But in quantum mechanics these two representations are exclusive: in the position representation, one loses all information on the particle's momentum, and conversely, in the momentum representation one loses all information on the particle's position; consequently, no information on an eventual correlation between position and momentum can be obtained.

It is interesting to introduce a quantum point of view intermediate between these two extremes, to keep at the same time information about position and momentum,

while obeying the general rules of quantum mechanics; these rules impose a limitation on the information precision. The Wigner transformation offers this intermediate point of view as it introduces a quantum mechanical function $\rho_W(\mathbf{r},\mathbf{p})$ that allows computing average values in the same way as with a classical distribution $\rho_{\rm cl}(\mathbf{r},\mathbf{p})$. Historically, this transformation was introduced[1] in 1932 by Wigner [98, 99] as he was working on the quantum corrections to thermal equilibrium, but it turned out to be a much more general tool. It yields very naturally semiclassical expansions in powers of $\hbar$ while studying, for instance, the temporal evolution of a quantum system. We shall also show that it provides a quantization method, leading in particular to correctly symmetrized expressions for quantum operators starting from classical functions of the positions and momenta. There are numerous domains in physics where the Wigner transform has proven useful, and sometimes indispensable.

This appendix will show how to associate with any quantum density operator a *Wigner distribution* $\rho_W(\mathbf{r},\mathbf{p})$, sometimes called a "semiclassical distribution", and we will discuss a certain number of its properties. In a similar way, to any operator (observable) acting in the state space, we can associate a Wigner transform $A_W(\mathbf{r},\mathbf{p})$ that is a simple function of $\mathbf{r}$ and $\mathbf{p}$.

In classical or semiclassical situations (i.e. when the spatial variations of physical quantities occur over large enough distances), the function $\rho_W(\mathbf{r},\mathbf{p})$ possesses all the properties of a classical distribution: it is a positive function that, multiplied by $A_W(\mathbf{r},\mathbf{p})$ and then integrated over the variables $\mathbf{r}$ and $\mathbf{p}$, does yield the average value of the operator A. In this case, we will show that the Wigner distribution simply describes the flow of the probability fluid (Chapter III, § D-1-c-β). As $\rho_W(\mathbf{r},\mathbf{p})$ allows keeping track of the correlations between position and momentum, this function is particularly useful in a number of cases, such as in the theory of quantum transport, with its numerous applications.

In the general case where the quantum effects are important (rapid spatial variations), the classical relations between the distribution and the average values are still valid, meaning the Wigner distribution continues to be quite useful. This distribution shows, however, a significant difference with a probability distribution: the function $\rho_W(\mathbf{r},\mathbf{p})$ can sometimes take on negative values. Furthermore, as we shall point out later, it can sometimes present "ghost" components where it is different from zero at points where the probability of finding the particle is zero. It is thus not possible to interpret the product $\rho_W(\mathbf{r},\mathbf{p})\mathrm{d}^3r\mathrm{d}^3p$ as the probability for the particle to occupy an infinitesimal cell $\mathrm{d}^3r\mathrm{d}^3p$ of the phase space centered at $(\mathbf{r},\mathbf{p})$; moreover, when dealing with infinitesimal volumes such an interpretation would be in clear contradiction with Heisenberg's uncertainty relations. Therefore, one should consider $\rho_W(\mathbf{r},\mathbf{p})$ to be a "quasi-distribution", a tool for computing all the average values without being a real probability distribution, even though we shall use the word distribution, following common usage.

This appendix introduces the tools necessary for the study of these different situations. We shall obtain, in particular, gradient expansions directly yielding expansions in powers of $\hbar$. We first introduce in § 1 a convenient form for the delta function of an operator. This form will be useful, in § 2, for defining the Wigner distribution of a spinless particle density operator. Several of its characteristics will be studied, in particular

[1] Wigner does mention in his article that the same transformation had already been used by L. Szilard, but in another context.

when dealing with a Gaussian wave packet. We then focus in § 3 on the *Wigner transform* of operators and show how it can be associated with the Wigner distribution of the density operator, leading to computations similar to those corresponding to a classical distribution. An important step will be the computation of the Wigner transform associated with a product of operators. Generalization of these concepts to take into account the spin, as well as the possible presence of several interacting particles, is discussed in § 4. Finally, we focus in § 5 on the physical properties of the Wigner transformation, using it to analyze a quantum interference experiment. We will show, in particular, how "ghost components" of the Wigner transform can appear, rapidly change sign, and are the signature of quantum effects.

1. Delta function of an operator

Consider a Hermitian operator A having a continuous spectrum, and whose eigenvalues are noted α. We define the operator $D_A(a)$, depending on a real continuous parameter a, as:

$$D_A(a) = \frac{1}{2\pi} \int \mathrm{d}x \; e^{ix(A-a)} \tag{1}$$

It is, in a way, a "delta function of an operator" associated with the difference between operator A and the constant a. We note $|\alpha, s\rangle$ the eigenvectors of A; the index s (assumed to be discrete) accounts for the possible degeneracy of the eigenvalue α. In a quantum state defined by the density operator ρ, the average value of this operator is:

$$\begin{aligned}
\langle D_A(a) \rangle &= \frac{1}{2\pi} \int \mathrm{d}x \; \mathrm{Tr}\left\{ \rho \; e^{ix(A-a)} \right\} \\
&= \frac{1}{2\pi} \int \mathrm{d}x \int \mathrm{d}\alpha \sum_s \langle \alpha, s | \, e^{ix(A-a)} \, | \alpha, s \rangle \, \langle \alpha, s | \, \rho \, | \alpha, s \rangle \\
&= \frac{1}{2\pi} \int \mathrm{d}x \int \mathrm{d}\alpha \sum_s e^{ix(\alpha-a)} \;\; \langle \alpha, s | \, \rho \, | \alpha, s \rangle
\end{aligned} \tag{2}$$

The integral over $\mathrm{d}x$ yields $2\pi\delta(\alpha - a)$, and we obtain:

$$\langle D_A(a) \rangle = \sum_s \langle a, s | \, \rho \, | a, s \rangle = \mathcal{P}(a) \tag{3}$$

where $\mathcal{P}(a)\,\mathrm{d}a$ is the probability, in a measurement associated with operator A, of finding a result in the interval $[a, a + \mathrm{d}a]$.

2. Wigner distribution of the density operator (spinless particle)

Imagine now that the physical system under study is a spinless particle, described by a density operator ρ. Our purpose is to introduce a function $\rho_W(\mathbf{r}, \mathbf{p})$ yielding simultaneously information on the probability of measurement results on either the position operator $\mathbf{R}$, or the momentum operator $\mathbf{P}$. As these are incompatible observables, we should not expect to make highly accurate predictions concerning both types of measurements: the precision is limited by Heisenberg's uncertainty principle. This function

will have to make a compromise between the two types of information, resulting from an unavoidable quantum uncertainty. As already mentioned in the introduction, this function is actually a "quasi-distribution", even though it is commonly called the "Wigner distribution".

2-a. Definition of the distribution, Weyl operators

By analogy with (1), we can introduce the "Weyl operator", which will be noted $W(\mathbf{r},\mathbf{p})$. It is, in a way, a "delta function" of an operator, associated at the same time with the operators $\mathbf{R}-\mathbf{r}$ and $\mathbf{P}-\mathbf{p}$. We set:

$$W(\mathbf{r},\mathbf{p}) = \frac{1}{\hbar^3}\int \frac{\mathrm{d}^3\kappa}{(2\pi)^3}\int \frac{\mathrm{d}^3x}{(2\pi)^3}\ \mathrm{e}^{i[\boldsymbol{\kappa}\cdot(\mathbf{R}-\mathbf{r})+\mathbf{x}\cdot(\mathbf{P}-\mathbf{p})/\hbar]} \tag{4}$$

where the integration variable $\boldsymbol{\kappa}$ has the dimension of a wave vector (the inverse of a length), and $\mathbf{x}$ the dimension of a length. This operator is Hermitian, as can be shown by changing the sign of the two integration variables. For each quantum state, the Wigner distribution $\rho_W(\mathbf{r},\mathbf{p};t)$ is defined as the average value in that state of the Weyl operator:

$$\boxed{\rho_W(\mathbf{r},\mathbf{p};t) = \langle W(\mathbf{r},\mathbf{p})\rangle} \tag{5}$$

If the system under study is characterized by a density operator $\rho(t)$ whose trace equals 1, this definition amounts to:

$$\rho_W(\mathbf{r},\mathbf{p};t) = \mathrm{Tr}\left\{\rho(t)\,W(\mathbf{r},\mathbf{p})\right\} \tag{6}$$

On the other hand, if the system is described by a normalized state vector $|\Psi\rangle$, this definition becomes:

$$\rho_W(\mathbf{r},\mathbf{p};t) = \langle\Psi(t)|\,W(\mathbf{r},\mathbf{p})\,|\Psi(t)\rangle \tag{7}$$

Note that the density operator as well as the state vector are dimensionless. Taking into account the factors $\hbar$ introduced in (4), $W(\mathbf{r},\mathbf{p})$ and $\rho_W(\mathbf{r},\mathbf{p})$ have the dimensions of $\hbar^{-3}$; the product $\rho_W(\mathbf{r},\mathbf{p})\mathrm{d}^3r\mathrm{d}^3p$ is thus dimensionless, as a probability should be. We are now going to show that this distribution has numerous useful properties.

Later, we shall need the matrix elements $\langle\mathbf{r}_1|\,W(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle$ of the operator $W(\mathbf{r},\mathbf{p})$ in the position representation. As demonstrated below, they are written:

$$\langle\mathbf{r}_1|\,W(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle = \frac{1}{(2\pi\hbar)^3}\ e^{-i\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)/\hbar}\ \delta\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2}-\mathbf{r}\right) \tag{8}$$

Demonstration:

The demonstration uses relation (63) of Complement $\mathrm{B_{II}}$, which expresses the exponential of the sum of two operators A and B as a product of exponentials, provided both operators commute with their commutator $[A,B]$:

$$e^{A+B} = e^{A}\ e^{B}\ e^{-\frac{1}{2}[A,B]} \tag{9}$$

We choose:

$$A = i\,(\mathbf{R}-\mathbf{r})\cdot\boldsymbol{\kappa} \qquad \text{and} \qquad B = i\,(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar \tag{10}$$

The commutator of these two operators:

$$[A,B] = -i\;\boldsymbol{\kappa}\cdot\mathbf{x} \tag{11}$$

is a number, so that both A and B commute with their commutator. Inserting relation (9) into (4), we get:

$$\langle\mathbf{r}_1|\,W\,(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle = \frac{1}{\hbar^3}\int\frac{\mathrm{d}^3\kappa}{(2\pi)^3}\int\frac{\mathrm{d}^3x}{(2\pi)^3}\;e^{i\boldsymbol{\kappa}\cdot\mathbf{x}/2}\,\langle\mathbf{r}_1|\,e^{i(\mathbf{R}-\mathbf{r})\cdot\boldsymbol{\kappa}}e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2\rangle \tag{12}$$

Now relation (13) of Complement $\mathrm{E_{II}}$ tells us that the action of operator $e^{i\mathbf{P}\cdot\mathbf{x}/\hbar}$ is a simple translation of the position eigenvalue by $-\mathbf{x}$, which means that:

$$e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2\rangle = e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2-\mathbf{x}\rangle \tag{13}$$

The function to be integrated in relation (12) then becomes:

$$e^{i\boldsymbol{\kappa}\cdot\mathbf{x}/2}\;\langle\mathbf{r}_1|\,e^{i(\mathbf{R}-\mathbf{r})\cdot\boldsymbol{\kappa}}e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2\rangle = e^{i\boldsymbol{\kappa}\cdot\mathbf{x}/2}\;e^{i(\mathbf{r}_2-\mathbf{x}-\mathbf{r})\cdot\boldsymbol{\kappa}}\;e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}\;\delta\,(\mathbf{r}_1-\mathbf{r}_2+\mathbf{x}) \tag{14}$$

The delta function, integrated over d^3x in (12), leads to replacing in the exponent $\mathbf{x}$ by $\mathbf{r}_2-\mathbf{r}_1$, and $\mathbf{r}_2-\mathbf{x}$ by $\mathbf{r}_1$. We then get:

$$\langle\mathbf{r}_1|\,W\,(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle = \frac{1}{(2\pi\hbar)^3}\int\frac{\mathrm{d}^3\kappa}{(2\pi)^3}\;e^{i\left[\boldsymbol{\kappa}\cdot\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2}-\mathbf{r}\right)-\mathbf{p}\cdot\left(\frac{\mathbf{r}_2-\mathbf{r}_1}{\hbar}\right)\right]} \tag{15}$$

The integration over $\mathrm{d}^3\kappa$ then yields a delta function, and we obtain relation (8).

2-b. Expressions for the Wigner transform

Definition (5) of the Wigner transformation may lead to various expressions for the Wigner transform, depending on the representation used in the state space.

α. Position representation

Using the position representation to calculate the trace appearing in (6), we get:

$$\begin{aligned}\rho_W\,(\mathbf{r},\mathbf{p};t) &= \int\mathrm{d}^3r_2\;\int\mathrm{d}^3r_1\;\;\langle\mathbf{r}_2|\,\rho\,(t)\,|\mathbf{r}_1\rangle\,\langle\mathbf{r}_1|\,W\,(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle \\ &= \int\mathrm{d}^3R\;\int\mathrm{d}^3y\;\left\langle\mathbf{R}+\frac{\mathbf{y}}{2}\right|\rho\,(t)\left|\mathbf{R}-\frac{\mathbf{y}}{2}\right\rangle\left\langle\mathbf{R}-\frac{\mathbf{y}}{2}\right|W\,(\mathbf{r},\mathbf{p})\left|\mathbf{R}+\frac{\mathbf{y}}{2}\right\rangle\end{aligned} \tag{16}$$

where, on the second line, we used the integration variables $\mathbf{R} = (\mathbf{r}_1+\mathbf{r}_2)/2$ and $\mathbf{y} = \mathbf{r}_2-\mathbf{r}_1$. Using relation (8) in this expression, we get the function $\delta\,(\mathbf{R}-\mathbf{r})$, to be integrated over d^3R, which leads to:

$$\boxed{\rho_W\,(\mathbf{r},\mathbf{p};t) = \frac{1}{(2\pi\hbar)^3}\int\mathrm{d}^3y\;e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\;\left\langle\mathbf{r}+\frac{\mathbf{y}}{2}\right|\rho\,(t)\left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle} \tag{17}$$

This relation is often used as a definition of the Wigner distribution. Its integration over $\mathrm{d}^3p/\left(2\pi\hbar\right)^3$ yields a function $\delta\left(\mathbf{y}\right)$, which leads to:

$$\int \mathrm{d}^3p\ \rho_W\left(\mathbf{r},\mathbf{p};t\right) = \langle\mathbf{r}|\,\rho\left(t\right)|\mathbf{r}\rangle = n\left(\mathbf{r},t\right) \tag{18}$$

We confirm a property of the classical distributions in phase space: the integral of the distribution over the momenta yields the probability density $n\left(\mathbf{r},t\right)$ of finding the particle at point $\mathbf{r}$.

In the particular case where the particle is described by a pure state $|\Psi\rangle$ – see relation (7) – the definition of the Wigner distribution becomes:

$$\boxed{\rho_W\left(\mathbf{r},\mathbf{p};t\right) = \frac{1}{\left(2\pi\hbar\right)^3}\int \mathrm{d}^3y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\ \Psi\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)\Psi^*\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right)} \tag{19}$$

where $\Psi\left(\mathbf{r};t\right)$ is the wave function in the position representation:

$$\Psi\left(\mathbf{r};t\right) = \langle\mathbf{r}\,|\Psi\left(t\right)\rangle \tag{20}$$

β. Momentum representation

As position and momentum play symmetrical roles in the argument, we expect to find a similar relation for the Wigner distribution, involving now the matrix elements of the density operator ρ in the momentum representation. This is indeed the case, and we are going to show that:

$$\boxed{\rho_W\left(\mathbf{r},\mathbf{p};t\right) = \frac{1}{\left(2\pi\hbar\right)^3}\int \mathrm{d}^3q\ e^{i\mathbf{q}\cdot\mathbf{r}/\hbar}\left\langle\mathbf{p}+\frac{\mathbf{q}}{2}\right|\rho\left(t\right)\left|\mathbf{p}-\frac{\mathbf{q}}{2}\right\rangle} \tag{21}$$

This expression is the exact analog of (17); it can be considered as an alternative definition of the Wigner distribution. As for the analogy of the property expressed by (18), it is easy to show that:

$$\int \mathrm{d}^3r\ \rho_W\left(\mathbf{r},\mathbf{p};t\right) = \langle\mathbf{p}|\,\rho\left(t\right)|\mathbf{p}\rangle \tag{22}$$

Just as for a classical distribution, the integral over the positions of the Wigner distribution yields the probability density of finding a given momentum $\mathbf{p}$.

Demonstration:

Inserting in the matrix element of (17) two closure relations on normalized momentum plane waves $\mathbf{q}'$ and $\mathbf{q}''$, yields the two scalar products:

$$\left\langle\mathbf{r}+\frac{\mathbf{y}}{2}\,\right|\left.\mathbf{q}'\right\rangle = \frac{1}{\left(2\pi\hbar\right)^{3/2}}e^{i\mathbf{q}'\cdot\left(\mathbf{r}+\frac{\mathbf{y}}{2}\right)/\hbar} \qquad \left\langle\mathbf{q}''\,\right|\left.\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle = \frac{1}{\left(2\pi\hbar\right)^{3/2}}e^{-i\mathbf{q}''\cdot\left(\mathbf{r}-\frac{\mathbf{y}}{2}\right)/\hbar} \tag{23}$$

and we can write:

$$\rho_W\left(\mathbf{r},\mathbf{p};t\right) = \frac{1}{\left(2\pi\hbar\right)^6}\int \mathrm{d}^3y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3q'\int \mathrm{d}^3q''\ e^{i\mathbf{q}'\cdot\left(\mathbf{r}+\frac{\mathbf{y}}{2}\right)/\hbar}e^{-i\mathbf{q}''\cdot\left(\mathbf{r}-\frac{\mathbf{y}}{2}\right)/\hbar}\left\langle\mathbf{q}'\right|\rho\left(t\right)\left|\mathbf{q}''\right\rangle \tag{24}$$

The summation over d^3y yields a delta function:

$$\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{i\left(\frac{\mathbf{q}'+\mathbf{q}''}{2}-\mathbf{p}\right)\cdot\mathbf{y}/\hbar}=\delta\left(\frac{\mathbf{q}'+\mathbf{q}''}{2}-\mathbf{p}\right) \tag{25}$$

so that if we take $\mathbf{Q}=\left(\mathbf{q}'+\mathbf{q}''\right)/2$ and $\mathbf{q}=\mathbf{q}'-\mathbf{q}''$ as integration variables, we obtain (21).

In the particular case where the particle is described by a pure state $|\Psi\rangle$, relation (21) becomes:

$$\boxed{\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3q\; e^{i\mathbf{q}\cdot\mathbf{r}}\;\overline{\Psi}\left(\mathbf{p}+\frac{\mathbf{q}}{2};t\right)\overline{\Psi}^*\left(\mathbf{p}-\frac{\mathbf{q}}{2};t\right)} \tag{26}$$

where $\overline{\Psi}\left(\mathbf{p};t\right)$ is the wave function in the momentum representation:

$$\overline{\Psi}\left(\mathbf{p};t\right)=\langle\mathbf{p}\,|\Psi\left(t\right)\rangle \tag{27}$$

If the wave function is factored[2]:

$$\overline{\Psi}\left(\mathbf{p};t\right)=\overline{\psi}_x\left(p_x;t\right)\times\overline{\psi}_y\left(p_y;t\right)\times\overline{\psi}_z\left(p_z;t\right) \tag{28}$$

relation (26) shows that the Wigner transform is also factored:

$$\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\rho_W^x\left(x,p_x;t\right)\times\rho_W^y\left(y,p_y;t\right)\times\rho_W^z\left(z,p_z;t\right) \tag{29}$$

with:

$$\rho_W^x\left(x,p_x;t\right)=\frac{1}{2\pi\hbar}\int \mathrm{d}q\; e^{iqx/\hbar}\;\overline{\psi}_x\left(p_x+\frac{q}{2};t\right)\overline{\psi}_x^*\left(p_x-\frac{q}{2};t\right) \tag{30}$$

The other two components $\rho_W^y\left(y,p_y;t\right)$ and $\rho_W^z\left(z,p_z;t\right)$ are defined in a similar way. In this particular case, one can reason independently for the three dimensions.

γ. Inverting the relations

We just saw that to each density operator corresponds a unique and well defined Wigner distribution. Inversely, starting from this distribution, one can reconstruct the corresponding density operator via its matrix elements. To take the inverse Fourier transform of (17), we multiply this relation by $e^{i\mathbf{p}\cdot\mathbf{z}/\hbar}$ and integrate over d^3p:

$$\int \mathrm{d}^3p\; e^{i\mathbf{p}\cdot\mathbf{z}/\hbar}\;\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3p\int \mathrm{d}^3y\; e^{i\mathbf{p}\cdot(\mathbf{z}-\mathbf{y})/\hbar}\left\langle\mathbf{r}+\frac{\mathbf{y}}{2}\right|\rho\left(t\right)\left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle \tag{31}$$

On the right-hand side, the integral over d^3p yields a function $(2\pi\hbar)^3\,\delta\left(\mathbf{z}-\mathbf{y}\right)$, so that this equality becomes:

$$\int \mathrm{d}^3p\; e^{i\mathbf{p}\cdot\mathbf{z}/\hbar}\;\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\left\langle\mathbf{r}+\frac{\mathbf{z}}{2}\right|\rho\left(t\right)\left|\mathbf{r}-\frac{\mathbf{z}}{2}\right\rangle \tag{32}$$

[2]Whether the wave function is factored in the momentum or position representation is equivalent.

Setting $\mathbf{r}_1 = \mathbf{r} + \mathbf{z}/2$ and $\mathbf{r}_2 = \mathbf{r} - \mathbf{z}/2$, we obtain the matrix elements of ρ in the position representation:

$$\boxed{\langle \mathbf{r}_1 | \, \rho \, (t) \, | \mathbf{r}_2 \rangle = \int \mathrm{d}^3 p \; e^{i \mathbf{p} \cdot (\mathbf{r}_1 - \mathbf{r}_2)/\hbar} \; \rho_W \left(\frac{\mathbf{r}_1 + \mathbf{r}_2}{2}, \mathbf{p}; t \right)} \tag{33}$$

Knowing the Wigner distribution $\rho_W (\mathbf{r}, \mathbf{p})$ thus defines the operator ρ in a unique way.

In a similar way, multiplying (21) by $e^{-i \mathbf{p}' \cdot \mathbf{r}/\hbar}$, integrating over $\mathrm{d}^3 r$, then setting $\mathbf{p}' = \mathbf{p}_1 - \mathbf{p}_2$ and $\mathbf{p} = (\mathbf{p}_1 + \mathbf{p}_2)/2$, yields the inversion relation in the momentum representation:

$$\boxed{\langle \mathbf{p}_1 | \, \rho \, (t) \, | \mathbf{p}_2 \rangle = \int \mathrm{d}^3 r \; e^{i (\mathbf{p}_2 - \mathbf{p}_1) \cdot \mathbf{r}/\hbar} \; \rho_W \left(\mathbf{r}, \frac{\mathbf{p}_1 + \mathbf{p}_2}{2}; t \right)} \tag{34}$$

2-c. Reality, normalization, operator form

Let us take the Hermitian conjugate of relation (17). As the density operator is Hermitian, the matrix element on the right-hand side becomes:

$$\left\langle \mathbf{r} + \frac{\mathbf{y}}{2} \right| \rho \, (t) \left| \mathbf{r} - \frac{\mathbf{y}}{2} \right\rangle^* = \left\langle \mathbf{r} - \frac{\mathbf{y}}{2} \right| \rho \, (t) \left| \mathbf{r} + \frac{\mathbf{y}}{2} \right\rangle \tag{35}$$

Changing the sign of the integration variable $\mathbf{y}$ then yields again relation (17); the distribution $\rho_W (\mathbf{r}, \mathbf{p})$ is therefore equal to its complex conjugate, meaning it is real.

We now compute the integral of $\rho_W (\mathbf{r}, \mathbf{p})$ over the entire phase space. Summing (18) over $\mathrm{d}^3 r$, we get:

$$\int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W (\mathbf{r}, \mathbf{p}; t) = \int \mathrm{d}^3 r \;\; \langle \mathbf{r} | \, \rho \, (t) \, | \mathbf{r} \rangle = \mathrm{Tr} \, \{ \rho \, (t) \} = 1 \tag{36}$$

where the last equality comes from the fact that the density operator has a trace equal to one. The Wigner distribution of the density operator is thus a real function, normalized to one in phase space, as is the case for a classical distribution.

Comment:

The density operator must obey a stronger constraint than having its trace equal to unity. It is defined as positive definite, meaning that for any ket $|\varphi\rangle$, we must have:

$$\langle \varphi | \, \rho \, (t) \, | \varphi \rangle \geq 0 \tag{37}$$

Now this condition is not merely equivalent to a positivity condition for the Wigner transform. In fact, we shall see below that the Wigner distribution of a density operator can become negative at certain points of phase space. However, the only Wigner distributions $\rho_W (\mathbf{r}, \mathbf{p}; t)$ acceptable for describing quantum systems are those that lead to a density operator obeying constraint (37). To know if a distribution in phase space is acceptable or not, is thus more difficult to decide in quantum mechanics than in classical mechanics.

We now show that relations (33) and (34) can be written in a simple operator form:

$$\rho(t) = (2\pi\hbar)^3 \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \; W(\mathbf{r},\mathbf{p}) \tag{38}$$

In this relation, $\rho_W(\mathbf{r},\mathbf{p};t)$ is the Wigner distribution, hence a function of position and momentum, but $W(\mathbf{r},\mathbf{p})$ is the Weyl operator defined in (4). To prove relation (38), we calculate the matrix element of this equality between the bra $\langle \mathbf{r}_1|$ and the ket $|\mathbf{r}_2\rangle$ to verify that we indeed obtain relation (33). Taking (8) into account, the matrix elements of the right-hand side are:

$$\begin{aligned}
(2\pi\hbar)^3 \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \;\; & \langle \mathbf{r}_1 | \, W(\mathbf{r},\mathbf{p}) \, |\mathbf{r}_2\rangle \\
&= \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \;\; e^{-i\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)/\hbar} \; \delta\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2} - \mathbf{r}\right) \\
&= \int \mathrm{d}^3 p \; \rho_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p};t\right) \;\; e^{-i\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)/\hbar}
\end{aligned} \tag{39}$$

which is equivalent to the right-hand side of (33).

2-d. Gaussian wave packet

A particular case where the computation can be completed is the one-dimensional Gaussian wave packet, studied in Complement $\mathrm{G_I}$. Relation (1) of this complement yields the normalized wave function $\psi_x(x)$ of such a wave packet in the position representation. We slightly modify it to center the wave packet at an arbitrary non-zero position x_0, and replace the variable k by $p = \hbar k$ (setting in particular $p_0 = \hbar k_0$). We obtain:

$$\begin{aligned}
\psi_x(x) &= \frac{\sqrt{a}}{(2\pi)^{3/4}} \int \frac{\mathrm{d}p}{\hbar} \; e^{-a^2(p-p_0)^2/4\hbar^2} \;\; e^{ip(x-x_0)/\hbar} \\
&= \left(\frac{2}{\pi a^2}\right)^{1/4} e^{ip_0(x-x_0)/\hbar} \; e^{-(x-x_0)^2/a^2}
\end{aligned} \tag{40}$$

where the second equality corresponds to relation (9) of Complement $\mathrm{G_I}$ (within an x_0 translation along Ox). The wave functions $(2\pi\hbar)^{-1/2} e^{ipx/\hbar}$ correspond to plane waves with momentum p (normalized with respect to that momentum); looking at the first equality in (40) we recognize the wave function in the momentum representation:

$$\overline{\psi}_x(p) = \frac{1}{(2\pi)^{1/4}} \sqrt{\frac{a}{\hbar}} \; e^{-a^2(p-p_0)^2/4\hbar^2} \;\; e^{-ipx_0/\hbar} \tag{41}$$

The Wigner distribution (30) can then be written (to simplify, we temporarily ignore the time dependence):

$$\begin{aligned}
\rho_W^x(x,p_x) &= \frac{a}{(2\pi)^{3/2}\hbar^2} \int \mathrm{d}q \; e^{-\frac{a^2}{4\hbar^2}\left[\left(p_x-p_0+\frac{q}{2}\right)^2+\left(p_x-p_0-\frac{q}{2}\right)^2\right]} \;\; e^{i\frac{q(x-x_0)}{\hbar}} \\
&= \frac{a}{(2\pi)^{3/2}\hbar^2} e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2} \int \mathrm{d}q \; e^{-\frac{a^2q^2}{8\hbar^2}} \;\; e^{i\frac{q(x-x_0)}{\hbar}}
\end{aligned} \tag{42}$$

The integral over dq is a Fourier transform whose value can be obtained by replacing a by $a/\sqrt{2}$ in relation (50) de l'Appendice I. We get:

$$\rho_W^x(x,p_x) = \frac{a}{(2\pi)^{3/2}\hbar^2} e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2} \times \sqrt{2\pi\hbar} \times \frac{\sqrt{2\hbar}}{\left(a/\sqrt{2}\right)} e^{-2(x-x_0)^2/a^2}$$
$$= \frac{1}{\pi\hbar} e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2} e^{-2\frac{(x-x_0)^2}{a^2}} \tag{43}$$

Taking (40) and (41) into account, we see that the Wigner distribution is simply the product of the probability densities in the position and momentum spaces:

$$\rho_W^x(x,p_x) = |\psi_x(x)|^2 \times |\overline{\psi}_x(p)|^2 \tag{44}$$

This result is particularly simple and shows that the Wigner transform of a Gaussian wave packet (40) contains no correlations between the variables x and p. It can be factored into two Gaussian functions, one concerning the momentum, the other the position. The first one is centered on the average momentum p_0, and has a width of the order of $\hbar/a$; the second, on the average position x_0, with a width of the order of a. These two widths are within the boundaries imposed by Heisenberg relations. Note that, in this case, the Wigner transform remains positive for all the values of its variables, as will also be the case for the semiclassical situations we consider in § 2-e.

Comment:

In the preceding paragraph, we ignored the time dependence of the wave packet. To take it into account and assuming we are dealing with a free particle, one can multiply, in (41), $\overline{\psi}_x(p)$ by $e^{-i\omega_p t}$, with:

$$\omega_p = \frac{p^2}{2m\hbar} \tag{45}$$

where m is the mass of the particle. This introduces, in the integral on the second line of (42), an additional exponential $e^{-ip_x qt/(m\hbar)}$ whose effect on the Fourier transform is to make the following substitution:

$$x \Rightarrow x - \frac{p_x}{m}\, t \tag{46}$$

This simply corresponds to the motion of the particle with a velocity p/m. Making this substitution in (43), we find that the Wigner distribution is still a product of two Gaussian functions, but no longer the product of a function of momentum by a function of position: correlations have appeared between the momentum and position variables.

2-e. Semiclassical situations

To what extent is it possible to consider the Wigner transform to be a true probability distribution? Relations (18) and (22) seem to be in favor of it, as they show that integrating that distribution over the momenta (or over the positions) actually yields a probability distribution of finding the particle at a given point (or with a given momentum). These two "marginal distributions" thus obtained by integration are both probability distributions. But this is not sufficient to ensure that the function $\rho_W(\mathbf{r},\mathbf{p})$

itself (before integration) has the same property. Actually, we already mentioned in the introduction that it is not possible, in general, to interpret the product $\rho_W(\mathbf{r},\mathbf{p})\,\mathrm{d}^3r\mathrm{d}^3p$ as yielding directly the probability for a particle to occupy an infinitesimal cell $\mathrm{d}^3r\mathrm{d}^3p$ of phase space, centered at $(\mathbf{r},\mathbf{p})$. Such a probability distribution is meaningless in quantum mechanics, as Heisenberg's relations forbid the existence of a quantum state defined with an arbitrary precision both in position and momentum spaces.

There are, however, some simple cases, that we shall call "semiclassical", where the Wigner transform is very similar to a classical probability distribution. They correspond to situations we will now study, where the physical quantities vary sufficiently slowly in space compared to a scale we shall define explicitly. In the following section, we will consider more general situations, where the properties of the Wigner transform are radically different. In particular, the Wigner transform can become negative, which immediately excludes any interpretation in terms of probability density.

α. Wave packet with slow spatial variations

Consider the wave function:

$$\psi(\mathbf{r}) = C(\mathbf{r})\, e^{iS(\mathbf{r})} \tag{47}$$

where $C(\mathbf{r})$ is the modulus of the wave function and $S(\mathbf{r})$ its phase. The probability density of presence is then $[C(\mathbf{r})]^2$, while relation (D-17) of Chapter III yields the probability current $\mathbf{J}(\mathbf{r})$:

$$\mathbf{J}(\mathbf{r}) = \frac{\hbar}{m}\,[C(\mathbf{r})]^2\ \boldsymbol{\nabla} S(\mathbf{r}) \tag{48}$$

The matrix elements of the corresponding density operator are written:

$$\langle \mathbf{r}'|\,\rho(t)\,|\mathbf{r}''\rangle = C(\mathbf{r}')\,C(\mathbf{r}'')\ e^{i\left[S(\mathbf{r}')-S(\mathbf{r}'')\right]} \tag{49}$$

We assume that, in the vicinity of each point $\mathbf{r}$, the wave function behaves locally as a plane wave:

$$\psi(\mathbf{r}) \simeq C(\mathbf{r})\, e^{i[\mathbf{K}(\mathbf{r})\cdot\mathbf{r}+\alpha(\mathbf{r})]} \qquad \text{in the vicinity of each point } \mathbf{r} \tag{50}$$

and that the two functions $C(\mathbf{r})$ and $\mathbf{K}(\mathbf{r})$, as well as the phase $\alpha(\mathbf{r})$, vary slowly in space: their variations are negligible over distances of the order of the de Broglie wavelength $2\pi/k(\mathbf{r})$. When $\mathbf{r}'$ and $\mathbf{r}''$ are close enough, one can expand the argument of the exponential in (49); the matrix elements $\langle \mathbf{r}'|\,\rho(t)\,|\mathbf{r}''\rangle$ of the density operator are then written:

$$\langle \mathbf{r}'|\,\rho(t)\,|\mathbf{r}''\rangle \simeq C(\mathbf{r}')\,C(\mathbf{r}'')\ e^{i\mathbf{K}\cdot\left(\mathbf{r}'-\mathbf{r}''\right)} \tag{51a}$$

where $\mathbf{K}$ is defined as:

$$\mathbf{K} = \boldsymbol{\nabla} S(\mathbf{r} = \frac{\mathbf{r}'+\mathbf{r}''}{2}) \tag{51b}$$

β. Density operator; link with the probability fluid

To characterize a semiclassical situation in a more general way, we will now reason in terms of a density operator, without restricting our study to a pure state as we did earlier. To start, we assume there is no long-range non-diagonal order[3]:

$$|\langle \mathbf{r}'| \rho(t) |\mathbf{r}''\rangle| \to 0 \qquad \text{if} \qquad |\mathbf{r}' - \mathbf{r}''| \gg l \tag{52}$$

where l is a macroscopic coherence length. For a pure state, l would be determined by the size of the domain where the wave function has a non-zero modulus $C(r)$. For a statistical mixture of states, we have a different situation: the phases of the various wave functions may interfere destructively at shorter distances, so that l can be much smaller. Nevertheless, we shall assume that l remains larger than a few de Broglie wavelengths $2\pi/k(\mathbf{r}')$ and that when $|\mathbf{r}' - \mathbf{r}''| \lesssim l$, the non-diagonal matrix elements of the density operator vary in a similar way as (51a):

$$\langle \mathbf{r}'| \rho(t) |\mathbf{r}''\rangle \simeq F(\mathbf{r}', \mathbf{r}'')\ e^{i\mathbf{K}\cdot(\mathbf{r}'-\mathbf{r}'')} \qquad \text{if } |\mathbf{r}' - \mathbf{r}''| \lesssim l \tag{53}$$

This expression is simply the generalization of (51a), only valid for a pure state; the real function $F(\mathbf{r}', \mathbf{r}'')$ replaces the modulus product $C(\mathbf{r}')\,C(\mathbf{r}'')$. Both functions $F(\mathbf{r}', \mathbf{r}'')$ and $\mathbf{K}$ are supposed to remain practically constant as the variables $\mathbf{r}'$ and $\mathbf{r}''$ vary by a quantity of the order of l.

With these assumptions, the values of the integration variable giving a significant contribution to the integral in relation (17) correspond to $|\mathbf{y}| \lesssim l$, so that we can write:

$$\rho_W(\mathbf{r}, \mathbf{p}) \simeq \frac{1}{(2\pi\hbar)^3} \int_D \mathrm{d}^3 y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\ F\left(\mathbf{r} + \frac{\mathbf{y}}{2}, \mathbf{r} - \frac{\mathbf{y}}{2}\right)\ e^{i\mathbf{K}(\mathbf{r})\cdot\mathbf{y}} \tag{54}$$

where the integration domain D is centered at $\mathbf{y} = 0$ and extends over a few coherence lengths l. As the function F is practically constant in this domain, and since $F(\mathbf{r}, \mathbf{r}) = \langle \mathbf{r}| \rho(t) |\mathbf{r}\rangle$, we get:

$$\rho_W(\mathbf{r}, \mathbf{p}) \simeq \frac{1}{(2\pi\hbar)^3} \langle \mathbf{r}| \rho(t) |\mathbf{r}\rangle \int_D \mathrm{d}^3 y\ e^{i[\mathbf{K}(\mathbf{r})\cdot\mathbf{y} - \mathbf{p}\cdot\mathbf{y}/\hbar]} \tag{55}$$

or else:

$$\rho_W(\mathbf{r}, \mathbf{p}) \simeq \langle \mathbf{r}| \rho(t) |\mathbf{r}\rangle \times \overline{g}[\mathbf{p} - \mathbf{p}_0(\mathbf{r})] \tag{56}$$

To write this expression, we have used the following definitions:

$$\overline{g}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^3} \int_D \mathrm{d}^3 y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \tag{57}$$

and:

$$\mathbf{p}_0(\mathbf{r}) = \hbar\, \mathbf{K}(\mathbf{r}) \tag{58}$$

[3]The concept of long-range non-diagonal order is introduced in Complement $\mathrm{A_{XVI}}$, §§ 2-a and 3-c, where, in particular, its relation with Bose-Einstein condensation is established. The present hypothesis concerning the absence of long-range order prevents $\rho(t)$ from being the one-body density operator of a system of condensed bosons.

The function $\overline{g}(\mathbf{p})$ is a momentum distribution centered at $\mathbf{p} = 0$, with a width $\Delta p \sim \hbar/l$. It is normalized to unity as the integral of $\overline{g}(\mathbf{p})$ over the momenta yields a function $\delta(\mathbf{y})$, which integrated over $\mathrm{d}^3 y$ is equal to one. Note that in (56) this function takes on its value for a momentum equal to $\mathbf{p} - \mathbf{p}_0(\mathbf{r})$, which means the momentum distribution is centered at the value $\mathbf{p}_0(\mathbf{r})$. As this momentum value depends on $\mathbf{r}$, correlations between position and momenta are now introduced in $\rho_W(\mathbf{r}, \mathbf{p})$.

Expression (56) for the distribution $\rho_W(\mathbf{r}, \mathbf{p})$ can be interpreted as a classical distribution in the probability fluid phase space: it is the product of the local probability density $\langle \mathbf{r} | \rho(t) | \mathbf{r} \rangle$ by a function of momentum $\overline{g}[\mathbf{p} - \mathbf{p}_0(\mathbf{r})]$ centered around the value $\mathbf{p}_0(\mathbf{r})$ defined in (58). Now this $\mathbf{p}_0(\mathbf{r})$ value is precisely the momentum value that, divided by m (to go from momentum to velocity) and multiplied by the probability density, yields the fluid probability current $\mathbf{J}(\mathbf{r})$. Note that the distribution keeps a certain width around $\mathbf{p}_0(\mathbf{r})$, of the order of $\hbar/l$, as required by Heisenberg's uncertainty relation. To sum up, in such semiclassical situations, the Wigner distribution directly reflects the spatial variation of the probability, and of its associated local current. It simply describes the flow of a "probability fluid" (III, § D-1-c-β), as does the distribution in phase space of an ensemble of classical particles forming a moving fluid.

2-f. Quantum situations where the Wigner distribution is not a probability distribution

In the previous examples, the properties of the Wigner distribution are very similar to those of a classical distribution. This is, however, not always the case: as surprising as it may seem, the Wigner transform can, in general, become negative.

α. *Odd wave function*

A very simple case offers such an example. In a one-dimension problem, imagine that the system has an odd wave function, as is the case for example for the first excited state of the harmonic oscillator. We then have, according to relation (19):

$$\begin{aligned} \rho_W^x(x = 0, p_x = 0) &= \frac{1}{2\pi\hbar} \int \mathrm{d}y \ \psi_x\left(\frac{y}{2}\right) \psi_x^*\left(-\frac{y}{2}\right) \\ &= -\frac{1}{2\pi\hbar} \int \mathrm{d}y \ \left|\psi_x\left(\frac{y}{2}\right)\right|^2 \end{aligned} \tag{59}$$

which is obviously negative. As odd wave functions often occur in quantum mechanics, we see that there exist numerous situations where the Wigner distribution has some properties unexpected for a distribution. Strictly speaking, the term "quasi-distribution" should always be used.

β. *Two-peak wave function*

Imagine now the particle wave function is the sum of two wave packets, one localized around $x = +b$, the other around $x = -b$:

$$\psi_x(x) = \frac{1}{\sqrt{2}} \left[\varphi(x - b) + e^{i\chi} \varphi(x + b)\right] \tag{60}$$

where the wave function $\varphi(x)$ is normalized; the relative phase factor χ is arbitrary. For the sake of simplicity, we assume that $\varphi(x)$ is zero when $|x| > a$ and that it is even. We

also suppose that in our case $b \gg a$, meaning that the two wave packets forming the total wave function are well separated.

Let us compute the Wigner distribution at point $x = 0$, therefore at a point where the wave function $\psi_x(x)$ is zero. In one dimension, relation (19) is written as:

$$\rho_W^x(x=0,p_x) = \frac{1}{4\pi\hbar}\int \mathrm{d}y\; e^{-ip_x y/\hbar}\left[\varphi\left(\frac{y}{2}-b\right)+e^{i\chi}\varphi\left(\frac{y}{2}+b\right)\right] \times\left[\varphi^*\left(-\frac{y}{2}-b\right)+e^{-i\chi}\varphi^*\left(-\frac{y}{2}+b\right)\right] \tag{61}$$

In this expression, the functions φ are zero if their argument's modulus is larger than a. As an example, $\varphi\left(\frac{y}{2}-b\right)$ is different from zero only if $y \simeq 2b$, whereas $\varphi^*\left(-\frac{y}{2}-b\right)$ is different from zero only if $y \simeq -2b$; consequently their product is always zero. Actually, in the product of the two brackets, only the "crossed" terms are non-zero, and we obtain (with our assumption that the function φ is even):

$$\rho_W^x(x=0,p_x) = \frac{1}{4\pi\hbar}\int \mathrm{d}y\; e^{-ip_x y/\hbar}\left[e^{-i\chi}\left|\varphi\left(b-\frac{y}{2}\right)\right|^2+e^{i\chi}\left|\varphi\left(b+\frac{y}{2}\right)\right|^2\right] \tag{62}$$

Changing the sign of the integration variable y for the second term in the bracket, we can write:

$$\rho_W^x(x=0,p_x) = \frac{1}{2\pi\hbar}\int \mathrm{d}y\;\cos\left(\frac{p_x y}{\hbar}+\chi\right)\left|\varphi\left(b-\frac{y}{2}\right)\right|^2 \tag{63}$$

In the limit where the width a becomes very narrow, the squared modulus of the wave function φ in the integral behaves as a delta function $\delta(b-y/2)$, and we get:

$$\rho_W^x(x=0,p_x) \simeq \frac{1}{\pi\hbar}\cos\left(\frac{2bp_x}{\hbar}+\chi\right) \tag{64}$$

This result illustrates two properties of the Wigner distributions that both seem quite surprising. The first is that the distribution is non-zero at point $x = 0$, whereas the probability of finding the particle at this position is strictly zero. The second is that the distribution is an oscillating function of momentum, taking successively positive and negative values, whereas a classical distribution always remain positive or zero. These two properties are actually related: integrating the distribution over all possible momenta p_x yields zero, which is in agreement with relation (18) stating that the integral of the Wigner distribution over the momenta yields the probability of the particle's presence at each point. More details on the properties of a two-peak wave function will be given in § 5-a.

3. Wigner transform of an operator

Consider now any operator A acting in the particle state space. We define its Wigner transform $A_W(\mathbf{r},\mathbf{p})$ in the same way as for a density operator, but without the prefactor $1/(2\pi\hbar)^3$ that appears in front of the integrals in (17) and (21):

$$\begin{aligned} A_W(\mathbf{r},\mathbf{p}) &= \int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\left\langle \mathbf{r}+\frac{\mathbf{y}}{2}\right| A \left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle \\ &= \int \mathrm{d}^3q\; e^{i\mathbf{q}\cdot\mathbf{r}/\hbar}\left\langle \mathbf{p}+\frac{\mathbf{q}}{2}\right| A \left|\mathbf{p}-\frac{\mathbf{q}}{2}\right\rangle \end{aligned} \tag{65}$$

To simplify the notation, this definition does not include a time dependence; one can, however, directly replace A by $A(t)$ and $A_W(\mathbf{r},\mathbf{p})$ by $A_W(\mathbf{r},\mathbf{p};t)$, without any other change. The inversion relations (33) and (34) now become:

$$\begin{aligned}\langle \mathbf{r}_1 | A | \mathbf{r}_2 \rangle &= \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 p \; e^{i\mathbf{p}\cdot(\mathbf{r}_1-\mathbf{r}_2)/\hbar} \; A_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p}\right) \\ \langle \mathbf{p}_1 | A | \mathbf{p}_2 \rangle &= \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 r \; e^{i(\mathbf{p}_2-\mathbf{p}_1)\cdot\mathbf{r}/\hbar} \; A_W\left(\mathbf{r},\frac{\mathbf{p}_1+\mathbf{p}_2}{2}\right)\end{aligned} \tag{66}$$

Taking the complex conjugate of relation (65) shows that the Wigner transform of a Hermitian operator A is necessarily a real function. Similarly, the fact that the complex conjugate of (66) is real indicates that it is a sufficient condition for hermiticity.

As the prefactor $1/(2\pi\hbar)^3$ is no longer included in the definition (65), the equivalent of relation (38) is now:

$$A = \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; A_W(\mathbf{r},\mathbf{p}) \; W(\mathbf{r},\mathbf{p}) \tag{67}$$

We saw previously that the operator $W(\mathbf{r},\mathbf{p})$ is Hermitian. The above relation then allows building a Hermitian operator A from any real function $A_W(\mathbf{r},\mathbf{p})$ of position and momentum. In other words, we found a quantization procedure for any classical function, often called "Weyl quantization" or "phase space quantization" [100, 101, 102]. Starting from two functions $A_W(\mathbf{r},\mathbf{p})$ and $B_W(\mathbf{r},\mathbf{p})$, whose product obviously commutes, this procedure yields two operators A and B that, in general, do not commute. Such an operation, which introduces in phase space a non-commutative structure, is sometimes referred to as "geometric quantization".

3-a. Average value of a Hermitian operator (observable)

We can now compute the average value of operator A in the quantum state defined by the density operator $\rho(t)$:

$$\langle A \rangle = \mathrm{Tr}\{\rho(t)\,A\} = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \; \langle \mathbf{r}_1 | \rho(t) | \mathbf{r}_2 \rangle \langle \mathbf{r}_2 | A | \mathbf{r}_1 \rangle \tag{68}$$

We are going to show that:

$$\boxed{\langle A \rangle = \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \; A_W(\mathbf{r},\mathbf{p})} \tag{69}$$

This relation is the exact analog of the relation one would obtain with a classical distribution. It is the reason the Wigner transform of the density operator is referred to as a "quasi-classical distribution", or more simply as a "distribution".

Demonstration:

Inserting in (68) the equalities (33) and (66) leads to:

$$\begin{aligned}\langle A \rangle = \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \int \mathrm{d}^3 p \; e^{i\mathbf{p}\cdot(\mathbf{r}_1-\mathbf{r}_2)/\hbar} \int \mathrm{d}^3 p' \; e^{-i\mathbf{p}'\cdot(\mathbf{r}_1-\mathbf{r}_2)/\hbar} \\ \times \; \rho_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p};t\right) A_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p}'\right)\end{aligned} \tag{70}$$

We now replace the integration variables $\mathbf{r}_1$ and $\mathbf{r}_2$ by the following variables:

$$\mathbf{r} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \qquad ; \qquad \mathbf{r}' = \mathbf{r}_1 - \mathbf{r}_2 \tag{71}$$

The summation over $\mathrm{d}^3 r'$ introduces a delta function:

$$(2\pi)^3 \ \delta\left(\frac{\mathbf{p} - \mathbf{p}'}{\hbar}\right) = (2\pi\hbar)^3 \ \delta\left(\mathbf{p} - \mathbf{p}'\right) \tag{72}$$

which takes care of the integration over $\mathrm{d}^3 p'$. We then finally obtain (69).

3-b. Special cases

In the special case in which the operator A depends only on the position operator:

$$A = F(\mathbf{R}) \qquad \text{and hence:} \qquad \langle \mathbf{r}_1 | A | \mathbf{r}_2 \rangle = F(\mathbf{r}_1) \ \delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{73}$$

the first line of (65) leads to:

$$A_W(\mathbf{r}, \mathbf{p}) = F(\mathbf{r}) \tag{74}$$

The Wigner transform of the operator is then simply the function $F(\mathbf{r})$, which does not depend on the momentum $\mathbf{p}$.

In a similar way, if A depends only on the momentum operator:

$$A = G(\mathbf{P}) \qquad \text{and hence:} \qquad \langle \mathbf{p}_1 | A | \mathbf{p}_2 \rangle = G(\mathbf{p}_1) \ \delta(\mathbf{p}_1 - \mathbf{p}_2) \tag{75}$$

the second line of (65) leads to:

$$A_W(\mathbf{r}, \mathbf{p}) = G(\mathbf{p}) \tag{76}$$

As a further illustration, let us find an operator A whose Wigner transform involves both position and momentum, for example:

$$A_W(\mathbf{r}, \mathbf{p}) = \mathbf{r} \cdot \mathbf{p} \tag{77}$$

Relation (66) yields its matrix elements:

$$\begin{aligned} \langle \mathbf{r}_1 | A | \mathbf{r}_2 \rangle &= \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 p \ e^{i\mathbf{p}\cdot(\mathbf{r}_1 - \mathbf{r}_2)/\hbar} \ \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \cdot \mathbf{p} \\ &= \frac{\hbar}{i} \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \cdot \boldsymbol{\nabla}_{\mathbf{r}_1} \delta(\mathbf{r}_1 - \mathbf{r}_2) \end{aligned} \tag{78}$$

We recognize in this expression the matrix elements of the operator $\mathbf{P}$, equal to the gradient of a delta function of the positions, multiplied by $\hbar/i$. Note, in addition, that $\mathbf{r}_1$ is the result of the action of the position operator on the bra, whereas $\mathbf{r}_2$ is the result of the action of the position operator on the ket. This means that:

$$A = \frac{1}{2} [\mathbf{R} \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{R}] \tag{79}$$

We thus get a Hermitian operator, as expected since its Wigner transform is real. It is however remarkable that building a quantum operator via the Wigner transforms spontaneously introduces an arrangement of the operators' order leading to the necessary symmetry. This property is quite general: starting from real classical functions, the Wigner transform allows building operators symmetrized with respect to position and momentum. This method is a real quantization procedure.

3-c. Wigner transform of an operator product

We are going to show that, in general, the Wigner transform associated with the product of two operators A and B is not simply the product of the Wigner transforms of each operator.

α. General expression

Let us apply relation (65) to obtain the Wigner transform of a product of two operators A and B. Inserting a closure relation on the kets $|\mathbf{z}\rangle$ leads to:

$$[AB]_W(\mathbf{r},\mathbf{p}) = \int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \int \mathrm{d}^3z\; \left\langle \mathbf{r}+\frac{\mathbf{y}}{2}\right| A \left|\mathbf{z}\right\rangle \left\langle \mathbf{z}\right| B \left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle \tag{80}$$

We can then replace the matrix elements of A and B by their expressions (66), which leads to:

$$\begin{aligned}&[AB]_W(\mathbf{r},\mathbf{p})\\ &= \frac{1}{(2\pi\hbar)^6}\int \mathrm{d}^3y \int \mathrm{d}^3z\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \int \mathrm{d}^3p_1 \int \mathrm{d}^3p_2\; e^{i\mathbf{p}_1\cdot\left(\mathbf{r}+\frac{\mathbf{y}}{2}-\mathbf{z}\right)/\hbar} e^{i\mathbf{p}_2\cdot\left(-\mathbf{r}+\frac{\mathbf{y}}{2}+\mathbf{z}\right)/\hbar}\\ &\qquad\qquad \times A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}+\frac{\mathbf{y}}{4},\mathbf{p}_1\right) B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}-\frac{\mathbf{y}}{4},\mathbf{p}_2\right)\end{aligned} \tag{81}$$

Instead of using the position representation, one can use the momentum representation; we then must use the relations on the second lines of (65) and (66). A reasoning similar to that used before leads to:

$$\begin{aligned}&[AB]_W(\mathbf{r},\mathbf{p})\\ &= \frac{1}{(2\pi\hbar)^6}\int \mathrm{d}^3q\; e^{i\mathbf{q}\cdot\mathbf{r}/\hbar} \int \mathrm{d}^3q' \int \mathrm{d}^3x \int \mathrm{d}^3y\; e^{i\left(\mathbf{q}'-\mathbf{p}-\frac{\mathbf{q}}{2}\right)\cdot\mathbf{x}/\hbar} e^{i\left(\mathbf{p}-\frac{\mathbf{q}}{2}-\mathbf{q}'\right)\cdot\mathbf{y}/\hbar}\\ &\qquad\qquad \times A_W\left(\mathbf{x},\frac{\mathbf{q}'+\mathbf{p}}{2}+\frac{\mathbf{q}}{4}\right) B_W\left(\mathbf{y},\frac{\mathbf{q}'+\mathbf{p}}{2}-\frac{\mathbf{q}}{4}\right)\end{aligned} \tag{82}$$

Depending on the case, it will be easier to use either (81) or (82). These two expressions are exact, but fairly complicated. They can be simplified, however, in a certain number of cases.

β. A few simple cases

As a first example, imagine that operator A is simply the position operator $\mathbf{R}$ while B can be any operator. As A_W is no longer dependent on $\mathbf{p}_1$, the integration over $\mathrm{d}^3p_1/(2\pi\hbar)^3$ in (81) yields a delta function $\delta\left(\mathbf{r}+\frac{\mathbf{y}}{2}-\mathbf{z}\right)$; this allows integrating over d^3z to obtain:

$$[\mathbf{R}B]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \int \mathrm{d}^3p_2\; e^{i\mathbf{p}_2\cdot\mathbf{y}/\hbar} \times \left[\mathbf{r}+\frac{\mathbf{y}}{2}\right] B_W(\mathbf{r},\mathbf{p}_2) \tag{83}$$

For the term in $\mathbf{r}$, the integral over d^3y of exponential $e^{i(\mathbf{p}_2-\mathbf{p})\cdot\mathbf{y}/\hbar}$ introduces a function $\mathbf{r}\delta(\mathbf{p}_2-\mathbf{p})$ with the coefficient $(2\pi\hbar)^3$. As for the term in $\mathbf{y}/2$, it yields $\boldsymbol{\nabla}_{\mathbf{p}_2}\delta(\mathbf{p}_2-\mathbf{p})$,

with the coefficient $(\hbar/2i)\,(2\pi\hbar)^3$. After integrating over d^3p_2, we get:

$$[\mathbf{R}B]_W(\mathbf{r},\mathbf{p}) = \mathbf{r}\ B_W(\mathbf{r},\mathbf{p}) - \frac{\hbar}{2i}\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p}) \tag{84}$$

If we now reverse the order of the operators $\mathbf{R}$ and B, the roles of $\mathbf{p}_1$ and $\mathbf{p}_2$ are interchanged in (81); the integration over $\mathrm{d}^3p_2/\,(2\pi\hbar)^3$ yields a function $\delta\left(-\mathbf{r}+\frac{\mathbf{y}}{2}+\mathbf{z}\right)$ and the integration over d^3z leads to:

$$[B\mathbf{R}]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3p_1\ e^{i\mathbf{p}_1\cdot\mathbf{y}/\hbar}\times B_W(\mathbf{r},\mathbf{p}_1)\left[\mathbf{r}-\frac{\mathbf{y}}{2}\right] \tag{85}$$

Compared to (83), the only change is the sign of $\mathbf{y}$ in the final bracket, so that we simply obtain the final result by changing the sign of the gradient on the right-hand side of (84). This means that the Wigner transform of the commutator is:

$$[\mathbf{R},B]_W(\mathbf{r},\mathbf{p}) = -\frac{\hbar}{i}\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p}) \tag{86}$$

Starting from (82), the same reasoning leads to:

$$[\mathbf{P}B]_W(\mathbf{r},\mathbf{p}) = \mathbf{p}\ B_W(\mathbf{r},\mathbf{p}) + \frac{\hbar}{i}\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) \tag{87}$$

This relation can now be iterated to obtain:

$$\left[\mathbf{P}^2B\right]_W(\mathbf{r},\mathbf{p}) = \mathbf{p}^2\ B_W(\mathbf{r},\mathbf{p}) + 2\frac{\hbar}{i}\mathbf{p}\cdot\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) - \hbar^2\Delta_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) \tag{88}$$

We then get the expression for the Wigner transform of the commutator of the momentum squared and any operator B:

$$\left[\mathbf{P}^2,B\right]_W(\mathbf{r},\mathbf{p}) = \frac{2\hbar}{i}\ \mathbf{p}\cdot\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) \tag{89}$$

This relation will be useful for what follows.

γ. *Gradient expansions*

We now show that relation (81) can be expressed as a series expansion of higher order derivatives of the two functions A_W and B_W, of the form:

$$[AB]_W(\mathbf{r},\mathbf{p}) = A_W(\mathbf{r},\mathbf{p})\times B_W(\mathbf{r},\mathbf{p}) + \frac{\hbar}{2i}\left\{B_W(\mathbf{r},\mathbf{p}),A_W(\mathbf{r},\mathbf{p})\right\} + ... \tag{90}$$

where we have used the classical definition of the "*Poisson bracket*" [103, 104] of classical Lagrangian mechanics:

$$\begin{aligned}&\left\{B_W(\mathbf{r},\mathbf{p}),A_W(\mathbf{r},\mathbf{p})\right\}\\ &\quad= \boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}A_W(\mathbf{r},\mathbf{p}) - \boldsymbol{\nabla}_{\mathbf{r}}A_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p})\end{aligned} \tag{91}$$

This shows that, to lowest order in $\hbar$, the Wigner function of an operator product is simply the product of the Wigner transforms of these operators. To first order, a correction must be added, which contains the Poisson bracket of the two Wigner transforms. It

is remarkable that purely quantum considerations bring in this classical Poisson bracket definition; this explains why these results are well suited for the study of the classical limit of quantum mechanics.

In (90), the expansion is limited to the contribution of the first order derivatives of the two functions. The following terms involve higher order derivatives and, consequently, higher powers of $\hbar$ (the corresponding result is called the "Groenewold's formula"; see for example [99]).

Demonstration:

Let us make in (81) the following change of momentum integration variables:

$$\begin{aligned} \mathbf{P} &= \frac{\mathbf{p}_1 + \mathbf{p}_2}{2} \\ \mathbf{q} &= \mathbf{p}_1 - \mathbf{p}_2 \end{aligned} \tag{92}$$

(despite the notation with a capital letter, $\mathbf{P}$ is a classical variable, not an operator). This leads to the new expression:

$$\begin{aligned} [AB]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^6} \int \mathrm{d}^3y \int \mathrm{d}^3z \int \mathrm{d}^3P \int \mathrm{d}^3q \; e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{y}/\hbar} e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{z})/\hbar} \\ \times A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}+\frac{\mathbf{y}}{4},\mathbf{P}+\frac{\mathbf{q}}{2}\right) B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}-\frac{\mathbf{y}}{4},\mathbf{P}-\frac{\mathbf{q}}{2}\right) \end{aligned} \tag{93}$$

If the two Wigner transforms A_W and B_W vary slowly with position and momentum, we can use the expansions:

$$\begin{aligned} A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}+\frac{\mathbf{y}}{4},\mathbf{P}+\frac{\mathbf{q}}{2}\right) &= A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2},\mathbf{P}\right) + \frac{\mathbf{y}}{4}\cdot\boldsymbol{\nabla}_{\mathbf{r}} A_W + \frac{\mathbf{q}}{2}\cdot\boldsymbol{\nabla}_{\mathbf{p}} A_W + ... \\ B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}-\frac{\mathbf{y}}{4},\mathbf{P}-\frac{\mathbf{q}}{2}\right) &= B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2},\mathbf{P}\right) - \frac{\mathbf{y}}{4}\cdot\boldsymbol{\nabla}_{\mathbf{r}} B_W - \frac{\mathbf{q}}{2}\cdot\boldsymbol{\nabla}_{\mathbf{p}} B_W + ... \end{aligned} \tag{94}$$

Keeping only the first term in each of these two expansions (zero-order term in the gradient expansion), the integrals over d^3y and d^3q introduce the delta functions $\delta(\mathbf{P}-\mathbf{p})$ and $\delta(\mathbf{r}-\mathbf{z})$ respectively, each with a coefficient $(2\pi\hbar)^3$. We then get:

$$[AB]_W(\mathbf{r},\mathbf{p}) = A_W(\mathbf{r},\mathbf{p}) \times B_W(\mathbf{r},\mathbf{p}) + ... \tag{95}$$

In this approximation, the Wigner transform of the product of two operators is thus simply the product of the Wigner transforms.

We now take into account the first order terms in the gradient expansion (94). The $\boldsymbol{\nabla}_{\mathbf{r}} A_W$ term on the first line contains a summation over d^3y modified by the presence of $\mathbf{y}$ in the integral:

$$\frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3y \; e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{y}/\hbar} \; \mathbf{y} = \frac{\hbar}{i}\boldsymbol{\nabla}_{\mathbf{P}}\delta(\mathbf{P}-\mathbf{p}) \tag{96}$$

The integral over d^3P in (93) is now modified and leads to a derivation with respect to $\mathbf{P}$ of the function to be integrated, a multiplication by the coefficient $-\hbar/i$, and finally the replacement of $\mathbf{P}$ by $\mathbf{p}$. On the other hand, the integral over $\mathrm{d}^3q/(2\pi\hbar)^3$ is unchanged and leads to the replacement of $\mathbf{z}$ by $\mathbf{r}$. The corresponding term is therefore written:

$$-\frac{\hbar}{4i}\boldsymbol{\nabla}_{\mathbf{p}} \cdot [B_W(\mathbf{r},\mathbf{p})\,\boldsymbol{\nabla}_{\mathbf{r}} A_W(\mathbf{r},\mathbf{p})] \tag{97}$$

As for the $\boldsymbol{\nabla}_{\mathbf{p}} A_W$ term on the first line of (94), it can be handled in the same way. The presence of the variable $\mathbf{q}$ transforms $\delta(\mathbf{r}-\mathbf{z})$ into $\boldsymbol{\nabla}_{\mathbf{z}}\delta(\mathbf{r}-\mathbf{z})$, with a coefficient $-\hbar/i$, where the sign change of this coefficient comes from the $-\mathbf{z}$ in the exponent $\mathbf{q}\cdot(\mathbf{r}-\mathbf{z})$; the integral over d^3P is unchanged. This yields the term:

$$\frac{\hbar}{4i}\boldsymbol{\nabla}_{\mathbf{r}}\cdot\left[B_W(\mathbf{r},\mathbf{p})\,\boldsymbol{\nabla}_{\mathbf{p}}A_W(\mathbf{r},\mathbf{p})\right] \tag{98}$$

which, added to (97), leads to the contribution (the terms involving a double derivation of A_W cancel each other):

$$\frac{\hbar}{4i}\left[\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}A_W(\mathbf{r},\mathbf{p})-\boldsymbol{\nabla}_{\mathbf{r}}A_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p})\right] \tag{99}$$

Finally, the terms coming from the second line of (94) are obtained by exchanging the roles of A_W and B_W, and changing the signs because of the opposite values of $\mathbf{y}$ and $\mathbf{q}$ in the of relation (94). We thus double the result (99), and finally obtain expression (90) to first order in the gradients.

3-d. Evolution of the density operator

The Schrödinger evolution of the density operator obeys the von Neumann equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\rho(t)=[H(t),\ \rho(t)] \tag{100}$$

Taking its Wigner transformation, this equation becomes:

$$i\hbar\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t)=\frac{1}{(2\pi\hbar)^3}\quad[H,\rho]_W(\mathbf{r},\mathbf{p};t) \tag{101}$$

where, on the right-hand side, is written the Wigner transform associated with the commutator of $H(t)$ and $\rho(t)$; the factor $1/(2\pi\hbar)^3$ comes from the definition of the Wigner distribution of the density operator, remembering that no such coefficient appears in the transform of an arbitrary operator. We already saw that the general expression of the Wigner transform of an operator product is somewhat complex, and the same is of course true for their commutator.

α. Classical limit

If we only keep, as in (90), the first order terms in the gradients, we see that the zero-order terms disappear, and that the terms in $H(t)\,\rho(t)$ and $\rho(t)\,H(t)$ double up; in addition, factors $\hbar$ on each side of the equations cancel out. Using this approximation, we get:

$$\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t)=\{H_W(\mathbf{r},\mathbf{p};t),\rho_W(\mathbf{r},\mathbf{p};t)\}+\hbar... \tag{102}$$

where the Poisson bracket of $H_W(\mathbf{r},\mathbf{p}_1;t)$ and $\rho_W(\mathbf{r},\mathbf{p}_1;t)$ is defined in (91). As noticed earlier in § 3-c-γ, the neglected terms are proportional to $\hbar$, and vanish in the classical limit $\hbar\to 0$. We find in this limit, where the gradients of the Wigner transforms with respect to position and momentum are small, the usual equations of classical dynamics.

β. Particle in an external potential

An exact calculation can be made if the particle's Hamiltonian is simply the sum of a kinetic energy and an external potential energy:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R};t) \tag{103}$$

where m is the mass of the particle. The contribution of the kinetic energy to the right-hand side of (101) comes directly from relation (89):

$$\left.\frac{\partial}{\partial t}\right|_{\text{kinetic}} \rho_W(\mathbf{r},\mathbf{p};t) = -\frac{\mathbf{p}}{m}\cdot\boldsymbol{\nabla}_{\mathbf{r}}\rho_W(\mathbf{r},\mathbf{p};t) \tag{104}$$

The evolution of the Wigner distribution induced by the kinetic energy operator is thus given by a "drift term" just as in classical physics.

As for the contribution of the potential energy, the computation is very similar to the one conducted at the beginning of § 3-c-β, except that instead of dealing with the operator $\mathbf{R}$ itself, we are now dealing with a function $V(\mathbf{R})$ of that operator. Taking $B = \rho$ in relations (83) and (85), they become:

$$[V(\mathbf{R})\,\rho]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3p_2\; e^{i\mathbf{p}_2\cdot\mathbf{y}}\times V\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)\rho_W(\mathbf{r},\mathbf{p}_2;t) \tag{105}$$

and:

$$[\rho V(\mathbf{R})]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3p_1\; e^{i\mathbf{p}_1\cdot\mathbf{y}}\times \rho_W(\mathbf{r},\mathbf{p}_1;t)\,V\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right) \tag{106}$$

Finally, the evolution of the Wigner distribution $\rho_W(\mathbf{r},\mathbf{p};t)$ obeys the following equation:

$$\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t) + \frac{\mathbf{p}}{m}\cdot\boldsymbol{\nabla}_{\mathbf{r}}\rho_W(\mathbf{r},\mathbf{p};t) = \frac{1}{i\hbar}\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\int \mathrm{d}^3p'\; e^{i(\mathbf{p}'-\mathbf{p})\cdot\mathbf{y}/\hbar} \times\left[V\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right) - V\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right)\right]\rho_W(\mathbf{r},\mathbf{p}';t) \tag{107}$$

This is an exact equation. It contains all the quantum effects that play a role in the particle's evolution. It obeys a local conservation law for the probability:

$$\frac{\partial}{\partial t}n(\mathbf{r},t) + \boldsymbol{\nabla}_{\mathbf{r}}\cdot\mathbf{J}(\mathbf{r},t) = 0 \tag{108}$$

where the local probability density $n(\mathbf{r},t)$ is defined in (18), and its associated current $\mathbf{J}(\mathbf{r},t)$ is defined as:

$$\mathbf{J}(\mathbf{r},t) = \int \mathrm{d}^3p\;\frac{\mathbf{p}}{m}\;\rho_W(\mathbf{r},\mathbf{p};t) \tag{109}$$

This can be shown by integrating (107) over d^3p, as the left-hand side then becomes identical to the left-hand side of (108), just as in classical mechanics; as for the right-hand side, the integration over d^3p introduces a function $\delta(\mathbf{y})$ that cancels the bracket in the remaining integral.

When the external potential varies slowly enough, one can use in (107) the following approximation:

$$V\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)-V\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right)=\mathbf{y}\cdot\boldsymbol{\nabla}_{\mathbf{r}}V(\mathbf{r};t)+... \tag{110}$$

The integration over $\mathrm{d}^3y/(2\pi\hbar)^3$ then leads to a function $(\hbar/i)\nabla_{\boldsymbol{p}'}\delta(\mathbf{p}'-\mathbf{p})$ and we get:

$$\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t)+\frac{\mathbf{p}}{m}\cdot\boldsymbol{\nabla}_{\mathbf{r}}\rho_W(\mathbf{r},\mathbf{p};t)=\boldsymbol{\nabla}_{\mathbf{r}}V(\mathbf{r};t)\cdot\boldsymbol{\nabla}_{\mathbf{p}}\rho_W(\mathbf{r},\mathbf{p};t)+... \tag{111}$$

One recognizes here the Liouville equation of classical mechanics. The dots at the end of the equation symbolize the possible contributions of terms containing higher order spatial derivatives of the potential $V(\mathbf{r};t)$. They come with a power of $\hbar$ increasing with the order of the derivative. This means that they correspond to quantum corrections: the faster the potential varies in space, the more terms need to be taken into account. On the other hand, when the potential varies slowly, only keeping the classical evolution term is a good approximation.

4. Generalizations

The above considerations can be directly generalized to particles with spin, or to an N-particle system.

4-a. Particle with spin

For a particle with spin, a basis in state space is formed by the kets $|\mathbf{r},\nu\rangle$, where $\mathbf{r}$ is the eigenvalue of the position operator, and ν the eigenvalue of the spin component on the quantization axis. The matrix elements of the density operator are then written:

$$\langle\mathbf{r},\nu|\,\rho(t)\,|\mathbf{r}',\nu'\rangle \tag{112}$$

For each value of ν and ν' we can perform a Wigner transformation and define, as in (17), the functions:

$$\rho_W^{\nu\nu'}(\mathbf{r},\mathbf{p};t)=\frac{1}{(2\pi\hbar)^3}\int\mathrm{d}^3y\;e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\;\left\langle\mathbf{r}+\frac{\mathbf{y}}{2},\nu\right|\rho(t)\left|\mathbf{r}-\frac{\mathbf{y}}{2},\nu'\right\rangle \tag{113}$$

As an example, for a spin 1/2 the two indices ν and ν' can take on two different values, noted $\pm$. We thus define four Wigner functions, which can be arranged in a 2×2 spin matrix:

$$\begin{pmatrix}\rho_W^{++}(\mathbf{r},\mathbf{p};t) & \rho_W^{+-}(\mathbf{r},\mathbf{p};t)\\ \rho_W^{-+}(\mathbf{r},\mathbf{p};t) & \rho_W^{--}(\mathbf{r},\mathbf{p};t)\end{pmatrix} \tag{114}$$

It is easy to show that this matrix is Hermitian:

$$\left[\rho_W^{+-}(\mathbf{r},\mathbf{p};t)\right]^*=\rho_W^{-+}(\mathbf{r},\mathbf{p};t) \tag{115}$$

Such a matrix is frequently used when studying the quantum properties of spin polarization transport in fluids (spin waves for example).

4-b. Several particles

For two spinless particles, relation (17) is easily generalized to:

$$\rho_W(\mathbf{r}_1,\mathbf{p}_1;\mathbf{r}_2,\mathbf{p}_2;t) = \frac{1}{(2\pi\hbar)^6}\int \mathrm{d}^3y_1 \int \mathrm{d}^3y_2\; e^{-i\mathbf{p}_1\cdot\mathbf{y}_1/\hbar}e^{-i\mathbf{p}_2\cdot\mathbf{y}_2/\hbar} \times \left\langle \mathbf{r}_1+\frac{\mathbf{y}_1}{2},\mathbf{r}_2+\frac{\mathbf{y}_2}{2}\middle|\,\rho(t)\,\middle|\mathbf{r}_1-\frac{\mathbf{y}_1}{2},\mathbf{r}_2-\frac{\mathbf{y}_2}{2}\right\rangle \tag{116}$$

Actually, any number of particles can be treated this way. Including the spin can be done as in the previous section, but it rapidly leads to a great number of Wigner functions (4^N for N particles each having a spin 1/2).

The Wigner distribution for a system including a large particle number N therefore depends on $6N$ variables when the particles have no spin; when the particles have a spin 1/2, it is no longer a single distribution that must be studied, but rather 4^N distributions which are the matrix elements of a spin operator. In practice, one usually uses the Wigner distribution of the one-particle density operator, resulting from the partial trace over the $N-1$ other particles, or sometimes the Wigner distribution of the two-particle density operator.

5. Discussion: Wigner distribution and quantum effects

Knowledge of the Wigner distribution allows computing the average values of observables, as seen from relation (69). It can be used to obtain the probability of any measurement result, since this probability is simply the average value of the projector onto the eigensubspace associated with this result. We simply have to compute the Wigner transform of this projector, multiply it by $\rho_W(\mathbf{r},\mathbf{p};t)$, and integrate the result over the two variables. From a practical point of view, all the information is contained in $\rho_W(\mathbf{r},\mathbf{p};t)$. However, and as already underlined with the examples given in § 2-f, that does not mean we should attribute too much physical content to the Wigner distribution itself. Strictly speaking, the Wigner distribution is rather a useful and powerful computation tool than a direct representation of the physical properties of the system.

To highlight the behavior of the Wigner transform in a situations where quantum effects are predominant, we now study an interference experiment.

5-a. An interference experiment

When the wave function of a particle goes through a screen pierced with two holes, it is split into two coherent wave packets propagating in space, and interfering when they overlap. Figure 1 represents these two wave packets after the screen, as they both propagate towards the region R where they will interfere. As they propagate in free space, the Wigner distribution associated with the particle simply obeys relation (104), which is just a classical equation of motion. What causes the interference effects in region R? To answer this question, we shall use relation (19), or its equivalent (26), which allow computing the Wigner transform associated with the particle's wave function.

This wave function is now the sum of two components, $\Psi_1(\mathbf{r},t)$ for the wave packet emerging from the first hole in the screen, and $\Psi_2(\mathbf{r},t)$ for the wave packet emerging from the second hole:

$$\Psi(\mathbf{r},t) = \Psi_1(\mathbf{r},t) + \Psi_2(\mathbf{r},t) \tag{117}$$

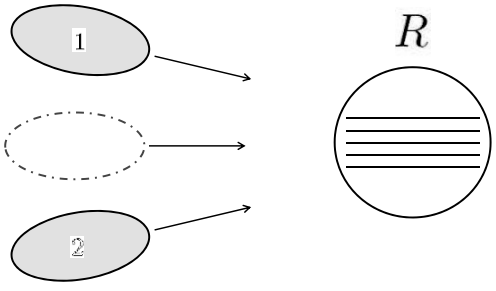


Figure 1: The wave function of a quantum particle can be split into two coherent components 1 *and* 2*, after passing, for example, through a screen pieced with two holes, or through an interferometer. As long as the two wave packets do not overlap, the Wigner distribution is the sum of three components, schematically drawn in ordinary space in the figure: a first one localized with wave packet* 1*, a second with wave packet* 2*, and finally a third one (circled with dashed lines) remaining at mid-distance from the two wave packets. This third component is called the "ghost component": when measuring its position, the particle can never be found in this component. The value of the ghost Wigner distribution oscillates rapidly as a function of the momentum* $\boldsymbol{p}$*.*
Later on, as the two wave packets 1 *and* 2 *overlap, the three components are different from zero in the same region of space; in addition, the momentum oscillations of the ghost component slow down and even vanish. This component now plays an essential role: as it is added to the terms* 1 *and* 2*, it is responsible for introducing the density oscillations producing the fringe pattern (schematized as horizontal lines in region R). It plays a virtual role as long as the wave packets are well separated, but an essential one when they overlap, as it leads to quantum interference effects.*

A similar situation has already been studied in § 2-f. Inserting (117) into relation (19), which is quadratic in Ψ, four contributions will come into play:

$$\rho_W(\mathbf{r},\mathbf{p};t) = \rho_W^1(\mathbf{r},\mathbf{p};t) + \rho_W^2(\mathbf{r},\mathbf{p};t) + \rho_W^{1,2}(\mathbf{r},\mathbf{p};t) + \rho_W^{2,1}(\mathbf{r},\mathbf{p};t) \tag{118}$$

In this equality, $\rho_W^1(\mathbf{r},\mathbf{p};t)$ is obtained when we replace in (19) the functions $\Psi(\mathbf{r},t)$ and $\Psi^*(\mathbf{r},t)$ by $\Psi_1(\mathbf{r},t)$ and $\Psi_1^*(\mathbf{r},t)$ respectively. The contribution $\rho_W^2(\mathbf{r},\mathbf{p};t)$ is obtained by replacing them by $\Psi_2(\mathbf{r},t)$ and $\Psi_2^*(\mathbf{r},t)$ respectively. Finally, the "crossed" contributions $\rho_W^{1,2}(\mathbf{r},\mathbf{p};t)$ and $\rho_W^{2,1}(\mathbf{r},\mathbf{p};t)$ come from replacing $\Psi(\mathbf{r},t)$ by $\Psi_1(\mathbf{r},t)$ and $\Psi^*(\mathbf{r},t)$ by $\Psi_2^*(\mathbf{r},t)$, and conversely. For example, relation (19) leads to:

$$\rho_W^{1,2}(\mathbf{r},\mathbf{p};t) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\;\Psi_1\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)\Psi_2^*\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right) \tag{119}$$

whereas the equivalent relation (26) yields another expression as a function of the Fourier transforms $\overline{\Psi}_1$ and $\overline{\Psi}_2^*$. It can easily be shown that the two distributions $\rho_W^{1,2}(\mathbf{r},\mathbf{p};t)$ and $\rho_W^{2,1}(\mathbf{r},\mathbf{p};t)$ are complex conjugates of each other. Their sum is real, as is, consequently, $\rho_W(\mathbf{r},\mathbf{p};t)$.

As an example, imagine that the two wave packets are Gaussian, as were the wave packets studied in § 2-d. We saw in Complement G_I shows that a Gaussian wave packet, as it propagates in free space, remains Gaussian at all times; its momentum dispersion remains constant, while its spatial width changes with time. For the sake of simplicity, we shall consider a one-dimensional problem and will not explicitly write the time dependence. We assume one of the wave packet to be centered at $+x_0$, and the other at $-x_0$. Relation (41) then leads to:

$$\begin{aligned}\overline{\psi}_1(p_x) &= \frac{1}{(2\pi)^{1/4}}\sqrt{\frac{a}{2\hbar}}\; e^{-a^2(p_x-p_0)^2/4\hbar^2}\; e^{-ip_x x_0/\hbar} \\ \overline{\psi}_2(p_x) &= \frac{1}{(2\pi)^{1/4}}\sqrt{\frac{a}{2\hbar}}\; e^{-a^2(p_x-p_0)^2/4\hbar^2}\; e^{ip_x x_0/\hbar}\end{aligned} \tag{120}$$

(a factor $1/\sqrt{2}$ has been added to ensure the normalization of the total wave function; we assume $a \ll x_0$, so that the spatial overlap of the two wave packets is negligible, and the squared norm of the sum is the sum of the squared norms). The same computation as in § 2-d then yields:

$$\begin{aligned}\rho_W^1(x,p_x) &= \frac{1}{2\pi\hbar}e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2}e^{-2\frac{(x-x_0)^2}{a^2}} \\ \rho_W^2(x,p_x) &= \frac{1}{2\pi\hbar}e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2}e^{-2\frac{(x+x_0)^2}{a^2}}\end{aligned} \tag{121}$$

As for the crossed contributions, the computation is slightly different. Since the two lines of relation (120) have different signs in front of x_0, the product $\overline{\psi}_1(p_x+q/2)\,\overline{\psi}_2^*(p_x-q/2)$ contains the exponential $\mathrm{e}^{-2ipx_0/\hbar}$, whereas the product $\overline{\psi}_1^*(p_x+q/2)\,\overline{\psi}_2(p_x-q/2)$ contains $\mathrm{e}^{2ipx_0/\hbar}$. The computation of (42) then becomes:

$$\begin{aligned}&\rho_W^{1,2}(x,p_x)+\rho_W^{2,1}(x,p_x) \\ &= \frac{1}{2}\frac{a}{(2\pi)^{3/2}\hbar^2}\int \mathrm{d}q\; e^{-\frac{a^2}{4\hbar^2}\left[\left(p-p_0+\frac{q}{2}\right)^2+\left(p-p_0-\frac{q}{2}\right)^2\right]}\; e^{iqx/\hbar}\left[e^{-2ipx_0/\hbar}+e^{2ipx_0/\hbar}\right] \\ &= \frac{a}{(2\pi)^{3/2}\hbar^2}\cos\frac{2px_0}{\hbar}e^{-\frac{a^2}{2\hbar^2}(p-p_0)^2}\int \mathrm{d}q\; e^{-\frac{a^2q^2}{8\hbar^2}}\; e^{i\frac{qx}{\hbar}}\end{aligned} \tag{122}$$

or else:

$$\rho_W^{1,2}(x,p_x)+\rho_W^{2,1}(x,p_x) = \cos\frac{2px_0}{\hbar}e^{-\frac{a^2}{2\hbar^2}(p-p_0)^2}\; e^{-\frac{2x^2}{a^2}} \tag{123}$$

Finally, the total Wigner transform is:

$$\rho_W(x,p_x) = \frac{1}{2\pi\hbar}e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2}\left[e^{-2\frac{(x-x_0)^2}{a^2}}+e^{-2\frac{(x+x_0)^2}{a^2}}+2\cos\frac{2p_xx_0}{\hbar}\; e^{-\frac{2x^2}{a^2}}\right] \tag{124}$$

The first two terms in the bracket are easy to understand: they are simply half the sum of the Wigner transforms associated with each of the wave packet. Each of these two terms is centered on the wave packet it corresponds to, that is at $x = \pm x_0$. The third term is the crossed term, which corresponds to an interference between the two wave packets, and is centered at $x = 0$, half way between them. In addition, this term oscillates as a function of p_x with a frequency proportional to the distance between the two wave packets.

5-b. General discussion; "ghost" component

The distribution $\rho_W^1(\mathbf{r},\mathbf{p};t)$ propagates as if it were the distribution of a free particle described by the single wave packet $\Psi_1(\mathbf{r},t)$; the distribution $\rho_W^2(\mathbf{r},\mathbf{p};t)$ corresponds to the second wave packet, here again as if it were isolated. If these were the only contributions, when the two wave packets overlap these two Wigner distributions would simply add to each other, since they follow a classical evolution; no quantum interference effects would result from this addition.

However, we saw that in (118) we must also include crossed terms (interference terms) whose properties are radically different from the first two terms. A first significant difference comes from their oscillations as a function of momentum, which necessarily involves positive and negative values of the distribution. This is definitely a quantum effect since a classical distribution must always be positive or zero. Another difference is that this crossed term in the Wigner transform propagates in a region of space where the wave function is zero, and consequently cannot correspond to any probability of the particle's presence; the integral over momentum of the last term on the right-hand side of (124) is indeed zero (in the limit $x_0 \gg a$ of well separated wave packets corresponding to the assumption made for our computation). The sum $\rho_W^{1,2}(x,p_x)+\rho_W^{2,1}(x,p_x)$ is sometimes called the "ghost component" of the Wigner distribution (or sometimes, in quantum optics, the "tamasic component"); when measuring the particle's position, it can never be found in this component[4]. Its value is always real, but not necessarily positive, because of its oscillations.

This means that, as long as the two wave packets 1 and 2 are well separated, the Wigner transform associated with the particle is the sum of three independent components: two components separately associated with each wave packet and propagating with them; one "ghost component", also propagating but remaining at mid-distance from the two wave packets. However, when the wave packets meet in region R, the three components of the Wigner transform overlap in space. The ghost component, which has a changing sign, combines with the other two components to modulate the particle's probability of presence, hence producing the interference pattern predicted by quantum mechanics. In a certain sense, one can say that the ghost component carries the quantum effects associated with the particle.

Conclusion

Quantum mechanics and classical mechanics are two very different theories. It was not obvious that, using the Wigner transforms, one could write the quantum equations in a form so akin to the classical equations of a distribution in phase space. Furthermore, we showed that any real classical function of position and momentum could be used in this formalism to generate a Hermitian operator acting in state space. In the limit where $\hbar \to 0$, the quantum equations of motion lead to the same Poisson brackets as the classical equations; quantum and classical theories then show strong similarities. Quantum effects, however, can manifest themselves in several ways:

- the evolution of the Wigner distribution can be significantly different from the classical evolution when the potentials vary rapidly on a scale on the order of $\hbar/p$ (de

[4]It is also known as the "empty component" stressing the fact that this component contains no particle.

Broglie wavelength), as higher order terms in the gradient expansion become essential.

- the Wigner transform is not always positive. We saw an example of this with the ghost component in an interference experiment, which, in a manner of speaking, carries the quantum effects to the usual components.

- whereas in classical physics any distribution in phase space, as long as it is positive and normalized, can be accepted, this is no longer the case in quantum mechanics. The only acceptable Wigner distributions are those which correspond to a density operator that is positive definite, a condition that is not expressed simply in terms of the distribution.

The Wigner transformation is frequently used in quantum physics. We already mentioned that it was introduced in 1932, while studying quantum corrections to thermal equilibrium [98]. It probably plays an even more important role in the study of transport properties where Boltzmann type equations contain simultaneous information on particles' positions and momenta. Furthermore, the Wigner transform is also useful for understanding and characterizing quantum effects, as its negativity in certain regions of phase space is a sensitive indicator of the existence of such effects. One can even use the Wigner transforms to introduce a "phase space formulation of quantum mechanics" [100, 101], totally equivalent to the usual formalism in terms of state space and operators, and which is a real quantization procedure. In a general way, the Wigner transformation belongs to the class of the so-called Liouville formulations of quantum mechanics [105], which have many uses.

Finally, there are many domains of physics (such as signal processing, in particular) in which the Wigner transformation is part of a larger class of mixed time-frequency transformations. Numerous types of such transformations exist (such as sliding window or envelope transforms, wavelets, etc.) chosen to best fit the problem at hand. Even in quantum mechanics there are other quasi-classical transforms, beside the Wigner transform, as for example the Husimi or the Kirkwood transforms, or the Glauber transform expressed in terms of creation and annihilation operators of the electromagnetic field; a review on that subject can be found in [99]. The Wigner transform still remains one of the most useful transforms, allowing, in particular, analytical calculations for many interesting cases.

Bibliography of volume III

[1] C. Cohen-Tannoudji, B. Diu and F. Laloë, *Quantum mechanics, Volume I*, Wiley (1977).

[2] C. Cohen-Tannoudji, B. Diu and F. Laloë, *Quantum mechanics, Volume II*, Wiley (1977). 1591

[3] R.K. Pathria, *Statistical mechanics*, Pergamon press (1972). 2296

[4] E.J. Mueller, Tin-Lun Ho, M. Ueda and G. Baym, "Fragmentation of Bose-Einstein condensates", *Phys. Rev.* **A 74**, 033612 (2006). 1656

[5] J.P. Blaizot et G. Ripka, *Quantum theory of finite systems*, the MIT Press (1986). 1678, 1701, 1809

[6] Wikipedia, "Density functional theory", https://en.wikipedia.org/wiki/Density_functional_theory 1699

[7] L.P. Kadanoff et G. Baym, *Quantum statistical mechanics*, Benjamin (1976). 1798

[8] A.J. Leggett, *Quantum liquids*, Oxford University Press, 2006. 1820, 1890, 1926

[9] J. Bardeen, L.N. Cooper and J.R. Schrieffer, "Theory of superconductivity", *Phys. Rev.* **108**, 1175-1204 (1957). 1889

[10] W. Ketterle and M. Zwierlein, "Making, probing and understanding ultracold Fermi gases" in *Proceedings of the international school of physics Enrico Fermi, Course CLXIV, Varenna*, edited by M. Inguscio, W. Ketterle and C. Salomon, IOS Press (Amsterdam), 2008; arXiv:0801.2500v1. 1926

[11] W. Zwerger, "The BCS-BEC crossover and the unitary Fermi gas", Springer, 2012. 1926

[12] M. Tinkham, "Introduction to superconductivity", Dover books on physics, 2004. 1926

[13] R.D. Parks, "Superconductivity", Volume 1 and 2, Dekker, 1969. 1926

[14] M. Combescot and S-Y Shiau, "Excitons and Cooper pairs", Oxford Univedrsity Press, 2016. 1926

[15] L. Pitaevskii and S. Stringari, "Bose-Einstein condensation and superfluidity", Oxford University Press, 2016. 1944

© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

[16] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Photons and atoms, introduction to quantum electrodynamics*, Wiley (1997). 1960, 1963, 1968, 1980, 1990, 2007, 2008, 2014, 2053, 2063

[17] T.A. Welton, "Some observable effects of the quantum-mechanical fluctuations of the electromagnetic field", *Phys. Rev.* **74**, 1157-1167 (1948). 2008

[18] J.V. Prodan, W. D. Phillips and H. Metcalf, "Laser production of a very slow monoenergetic atomic beam", *Phys. Rev. Lett.* **49**, 1149-1153 (1982). 2025

[19] T.W. Hänsch and A. Schawlow, "Cooling of gases by laser radiation" *Opt. Comm.* **13**, 68-69 (1975). 2026

[20] D.J. Wineland and H. Dehmelt, "Proposed $10^{14}\delta\nu < \nu$ laser fluorescence spectroscopy on Tl^+ mono-ion oscillator III (sideband cooling)", *Bull. Am. Phys. Soc.* **20**, 637 (1975). 2026

[21] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Atom-Photon Interactions. Basic Processes and Applications*, Wiley-Interscience (1992). 2028, 2107, 2117, 2130, 2136, 2145

[22] Special issue on laser cooling and trapping, *JOSA B, Optical Physics*, **6**, Number 11 (1989). 2034

[23] W. Ketterle and N.L Van Druten, "Evaporative cooling of trapped atoms", *Advances in atomic, molecular, and optical physics*, **37**, 181-236 (1996). 2034

[24] C. Cohen-Tannoudji and D. Guéry-Odelin, *Advances in atomic physics. An Overview*, World Scientific, Singapour (2011). 1779, 2034, 2036, 2039, 2065, 2127, 2152, 2153

[25] W.E. Lamb, "Capture of neutrons by atoms in a crystal", *Phys. Rev.* **55**, 190-197 (1939). 2040

[26] R.V. Pound and G. A. Rebka Jr., "Apparent weight of photons", *Phys. Rev. Lett.* **4**, 337-341 (1960). 2040

[27] L.S. Vasilenko,V. P. Chebotayev and A. V. Shishaev, "Line shape of two-photon absorption in a standing-wave field in a gas", *JETP Lett.* **12**, 113-116 (1970). 2041

[28] B. Cagnac, G. Grynberg and F. Biraben, "Spectroscopie d'absorption multiphotonique sans effet Doppler", *J. Phys. (Paris)* **34**, 845-858 (1973). 2041

[29] T.W. Hänsch, "Passion for precision", *Rev. Mod. Phys.* **78**, 1297-1309 (2006). 2041

[30] A. Kastler, "Projet d'expérience sur le moment cinétique de la lumière", Société des Sciences physiques et naturelles de Bordeaux, Jan. 28 (1932). 2052

[31] R.A. Beth, "Mechanical Detection and Measurement of the Angular Momentum of Light", *Phys. Rev.* **50**, 115-125 (1936). 2052

[32] J.W. Simmons and M.J. Guttmann, *States, waves and photons: a modern introduction to light*, Addison-Wesley (1970), Chap. 9. 2052

[33] A.M. Yao and M. J. Padgett, "Orbital angular momentum: origins, behavior and applications", *Advances in Optics and Photonics*, IOP Publishing **3**, 161-204 (2011). 2052

[34] J.D. Jackson, *Classical electrodynamics, 3rd ed.*, Wiley (1999). 2053, 2063

[35] J. Brossel et A. Kastler, "La détection de la résonance magnétique des niveaux excités ", *C. R. Acad. Sci.* **229**, 1213 (1949). 2059

[36] J. Brossel, and F. Bitter, "A new 'Double Resonance' method for investigating atomic energy levels. Application to Hg $^3P_1$", *Phys. Rev.* **86**, 308 (1952). 2059, 2061

[37] J.N. Dodd, W. N. Fox, G. W. Series and M. J. Taylor, "Light beats as indicators of structure of atomic energy levels", *Proc. Phys. Soc.* **74**, 789 (1959). 2061

[38] E. Majorana, "Atomi orientati in campo magnetico variabile", *Nuovo Cimento* **9**, 43-50 (1932). 2061

[39] A. Kastler, "Quelques suggestions concernant la production optique et la détection optique d'une inégalité de population des niveaux de quantifigation spatiale des atomes. Application à l'expérience de Stern et Gerlach et à la résonance magnétique", *J. Phys. Radium* **11**, 255-265 (1950). 2062

[40] N. F. Ramsey, *Molecular beams*, Oxford University Press (1956). 2064

[41] L. Allen, S.M. Barnett and M.J. Padgett, *Optical angular momentum*, IOP Publishing (2003). 2065

[42] M.F. Andersen, C. Ryu, P. Cladé, V. Natarajan, A. Vaziri, K. Helmerson, and W.D. Phillips, "Quantized rotation of atoms from photons with orbital angular momentum" *Phys. Rev. Lett.* **97**, 170406 (2006). 2066

[43] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt", *Annalen der Physik* **17**, 132-149 (1905). 2110

[44] R.A. Millikan, "On the elementary electric charge and the Avogadro constant", *Physical Review* **2**, 109-143 (1913). 2111

[45] W.E. Lamb and M.O. Scully, "The photoelectric effect without photons", in *Polarisation, Matière et Rayonnement*, (Presses Universitaires de France), Jubilee volume in honour of Alfred Kastler, p. 363 (1969). 2110

[46] E. Hanbury Brown and R.Q. Twiss, "A test of a new type of stellar interferometer on Sirius", *Nature*, **178**, 1046-1048 (1956). 2120

[47] P. Grangier, G. Roger and A. Aspect, "Experimental evidence for a photon anticorrelation effect on a beam splitter: a new light on single-photon interferences", *Europhys. Lett.* **1**, 173-179 (1986). 2121

[48] J.T. Höffges, H.W. Baldauf, W. Lange and H. Walther "Heterodyne measurements of the resonance fluorescence of a single ion", *J. of Mod. Optics* **44**, 1999-2010 (1997). 2121, 2122

[49] H.J. Kimble, M. Dagenais and L. Mandel, "Photon antibunching in resonance fluorescence", *Phys. Rev. Lett.* **39**, 691 (1977). 2121

[50] S. Pancharatnam, "Light shifts in semiclassical dispersion theory", *J. Opt. Soc. Am.* **56**, 1636 (1966). 2139

[51] C. Cohen-Tannoudji, "Théorie quantique du cycle de pompage optique. Vérification expérimentale des nouveaux effets prévus", *Ann. Phys.* **13**, 423-461 and 469-504(1962). 2140

[52] C. Cohen-Tannoudji and J. Dupont-Roc, "Experimental study of light shifts in weak magnetic fields", *Phys. Rev. A* **5**, 968-984 (1972). 2140

[53] C. Cohen-Tannoudji, "Observation d'un déplacement de raie de résonance magnétique causé par l'excitation optique", *C. R. Acad. Sci.* **252**, 394-396 (1961). 2140, 2141

[54] B.R. Mollow, "Power spectrum of light scattered by two-level systems", *Phys. Rev.* **188**, 1969-1975 (1969). 2144

[55] S.H. Autler and C. H. Townes, "Stark effect in rapidly varying fields", *Phys. Rev.* **100**, 703-722 (1955). 2144

[56] S. Chu, J. Bjorkholm, A. Ashkin, and A. Cable, "Experimental observation of optically trapped atoms", *Phys. Rev. Lett.* **57**, 314-317 (1986). 2152

[57] R.J. Cook and R.K. Hill, "An electromagnetic mirror for neutral atoms", *Opt. Commun.* **43**, 258-260 (1982). 2153

[58] M.Greiner, O. Mandel, T. Esslinger, T.W. Hänsch and I. Bloch, I. *Nature* **415**, "Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms", 39-44 (2002). 2154

[59] M. Ben Dahan, E. Peik, J. Reichel, Y. Castin and C. Salomon, "Bloch oscillations of atoms in an optical potential", *Phys. Rev. Lett.* **76**, 4508-4511 (1996). 2155

[60] P.D. Lett, R.N. Watts, C.I. Westbrook, W.D. Phillips, P.L. Gould and H.J. Metcalf, "Observation of atoms laser coooled below the Doppler limit", *Phys. Rev. Lett.* **61**, 169-172 (1988). 2155

[61] J. Dalibard and C. Cohen-Tannoudji, "Laser cooling below the Doppler limit by polarization gradients: simple theoretical models", *J. Opt. Soc. Am. B* **6**, 2023-2045 (1989). 2159

[62] P.J. Ungar, D.S.Weiss, E. Riis and S. Chu, "Optical molasses and multilevel atoms: theory", *J. Opt. Soc. Am. B* **6**, 2058-2071 (1989). 2159

[63] C. Salomon, J. Dalibard, W.D. Phillips, A. Clairon and S. Guellati, *Europhys. Lett.*, "Laser cooling of Cesium atoms below 3 μK", **12**, 683-688 (1990). 2159

[64] S. Gleyzes, S. Kuhr, C. Guerlin, J. Bernu, S. Deléglise, U.B. Hoff, M. Brune, J-M. Raimond and S. Haroche, "Quantum jumps of light recording the birth and death of a photon in a cavity", *Nature* **446**, 297-300 (2007). 2160

[65] G. Grynberg, A. Aspect et C. Fabre, *Introduction to quantum optics*, with contributions from F. Bretenaker and A. Browaeys, Cambridge University Press (2010). 2186

[66] D.F. Walls and G.J. Milburn, *Quantum optics*, Springer (1994). 2186

[67] E. Schrödinger, "Discussion of probability relations between separated systems", *Proc. Cambridge Phil. Soc.* **31**, 555 (1935); "Probability relations between separated systems", *Proc. Cambridge Phil. Soc.* **32**, 446 (1936). 2190

[68] F. Laloë, *Do we really understand quantum mechanics?* , Cambridge University Press (2012); second expanded edition (2019). 2202, 2213, 2330

[69] A. Einstein, B. Podolsky and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?", *Phys. Rev.* **47**, 777–780 (1935); *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek eds., Princeton University Press (1983), pp. 138–141. 2204, 2207

[70] A. Einstein, "Quantenmechanik und Wirklichkeit", *Dialectica* **2**, 320–324 (1948). 2207

[71] N. Bohr, "Can quantum-mechanical description of physical reality be considered complete?", *Phys. Rev.* **48**, 696–702 (1935). 2207

[72] N. Bohr, "On the notions of causality and complementarity", *Dialectica* **2**, 312–319 (1948); Science, New Series **111**, 51-54 (Jan20, 1950). 2207

[73] J.S. Bell, "On the problem of hidden variables in quantum mechanics", *Rev. Mod. Phys.* **38**, 447–452 (1966); reprinted in *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), 396–402 and in chapter 1 of [74]. 2208

[74] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press (1987); second augmented edition (2004), which contains the complete set of J. Bell's articles on quantum mechanics. 2329

[75] A. Peres, "Unperformed experiments have no results", *Am. J. Phys.* **46**, 745–747 (1978). 2212

[76] J.A. Wheeler, "Niels Bohr in today's words" in *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 182–213. 2212

[77] A list of references describing Bell's experiments can be found in: A. Aspect, "Closing the door on Einstein and Bohr's quantum debate", *Physics* **8**, 123 (2015). 2212

[78] D. Mermin, *Quantum computer science*, Cambridge University Press (2007). 2212

[79] N. Gisin, G. Ribordy, W. Tittel and H. Zbinden, "Quantum cryptography", *Rev. Mod. Phys.* **74**, 145-195 (2002). 2213

[80] V. Coffman, J. Kundu and W.K. Wootters, "Distributed entanglement", *Phys. Rev.* **A 61**, 052306 (2000). 2223

[81] R.F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden variable model", *Phys. Rev.* **A 40**, 4277-4281 (1989). 2224

[82] A. Peres, "Separability criterion for density matrices", *Phys. Rev. Lett.* **77**, 1413-1415 (1996). 2224

[83] D.M. Greenberger, M.A. Horne and A. Zeilinger, "Going beyond Bell's theorem", pp. 69-72 in *Bell's theorem, quantum theory and conceptions of the Universe*, M. Kafatos (ed.), Kluwer Academic Publishers (1989). 2227

[84] D.M. Greenberger, M.A. Horne, A. Shimony, and A. Zeilinger, "Bell's theorem without inequalities", *Am. J. Phys.* **58**, 1131-1143 (1990). 2227

[85] D. Bouwmeester, J.W. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger, "Observation of three-photon Greenberger-Horne-Zeilinger entanglement", *Phys. Rev. Lett.* **82**, 1345-1349 (1999). 2231

[86] see for example § 5-1-2 of [68]. 2231

[87] J.W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, "Experimental entanglement swapping: entangling photons that never interacted", *Phys. Rev. Lett.* **80**, 3891-3894 (1998). 2235

[88] B. Hensen, H. Bernien, A.E. Dréau, A. Reiserer, N. Kalb, M.S. Blok, J. Ruitenberg, R.F. Vermeulen, R.N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M.W. Mitchell, M. Markham, D.J. Twitchen, D. Elkouss, S. Wehner, T.H. Taminiau and R. Hanson, "Experimental loophole-free violation of a Bell inequality using electron spins separated by 1.3 km", *Nature* **526**, 682-686 (2015). 2235

[89] M.R. Andrews, C.G. Townsend, H.J. Miesner, D.S. Durfee, D.M. Kurn and W. Ketterle, "Observation of interference between two Bose condensates", *Science* **275**, 637-641 (1997). 2238

[90] F. Laloë, "The hidden phase of Fock states, quantum non-local effects", *Eur. Phys. J. D* **33**, 87-97 (2005); F. Laloë and W.J. Mullin, "Non-local quantum effects with Bose-Einstein condensates", *Phys. Rev. Lett.* **99**, 150401 (2007). 2262, 2263, 2265

[91] R.P. Feynman and A.R. Hibbs, *Quantum mechanics and path integrals*, McGraw Hill (1965). 2267

[92] J. Zinn-Justin, *Intégrale de chemin en mécanique quantique: Introduction*, CNRS Editions et EDP Sciences (2003). 2267, 2280

[93] M. Le Bellac, *Physique quantique*, EDP Sciences et CNRS Editions (2013), tome II, chapitre 12. 2267

[94] D.M. Ceperley, "Path integrals in the theory of condensed helium", *Rev. Modern Physics*, **67**, 279-355 (1995). 2280

[95] B. Diu, C. Guthmann, D. Lederer et B. Roulet, *Physique statistique*, Hermann (1989). 2296

[96] K. Huang, *Statistical mechanics*, Wiley (1963). 2296

[97] F. Reif, *Fundamental of statistical and thermal physics*, McGraw-Hill (1965). 2296

[98] E. Wigner, "On the quantum correction for thermodynamic equilibrium", *Phys. Rev.* **40**, 749-759 (1932). 2298, 2323

[99] M. Hillery, R.F. O'Connell, M.O. Scully and E.P. Wigner, "Distribution functions in physics; fundamentals", *Physics Reports*, **106**, 121-167 (1984). 2298, 2315, 2323

[100] A. Perelomov, "Generalized coherent states and their applications", Springer (1986); see in particular Chap. 16. 2311, 2323

[101] C.K. Zachos, D. Fairlie and T.L. Cutright, *Quantum mechanics in phase space*, World Scientific, Singapore (2005); with the same title, *Asia Pacific Newsletters*, **01**, 37-46 (2012) or ArXiv:1104.5269v2. 2311, 2323

[102] G.G. Athanasiu and E.G. Fioratos, "Coherent states in finite quantum mechanics", *Nuclear Physics* **B 425**, 343-364 (1994). 2311

[103] L. Landau and E. Lifchitz, *Mechanics, Course of theoretical physics, Vol. I*, C § 42, Pergamon Press (1960) and Elsevier Butterworth-Heinemann (1976). 2314

[104] H. Goldstein, C.P. Poole et J.L. Safko, *Classical mechanics*, Addison-Wesley (2001). 2314

[105] N.L. Balazs and B.K. Jennings, "Wigner's function and other distribution functions in mock phase spaces", *Physics Reports*, **104**, 347-391 (1984). 2323

Absorption
and emission of photons, 2073
collision with, 971
of a quantum, a photon, 1311, 1353
of field, 2149
of several photons, 1368
rates, 1334
Acceptor (electron acceptor), 1495
Acetylene (molecule), 878
Action, 341, 1539, 1980
Addition
of angular momenta, 1015, 1043
of spherical harmonics, 1059
of two spins 1/2, 1019
Adiabatic
branching of the potential, 932
Adjoint
matrix, 123
operator, 112
Algebra (commutators), 165
Allowed energy band, 381, 1481, 1491
Ammonia (molecule), 469, 873
Amplitude
scattering amplitude, 929, 953
Angle (quantum), 2258
Angular momentum
addition of momenta, 1015, 1043
and rotations, 717
classical, 1529
commutation relations, 669, 725
conservation, 668, 736, 1016
coupling, 1016
electromagnetic field, 1968, 2043
half-integral, 987
of identical particles, 1497(ex.)
of photons, 1370
orbital, 667, 669, 685
quantization, 394
quantum, 667
spin, 987, 991
standard representation, 677, 691
two coupled momenta, 1091
Anharmonic oscillator, 502, 1135
Annihilation operator, 504, 513, 514, 1597
Annihilation-creation (pair), 1831, 1878
Anomalous
average value, 1828, 1852
dispersion, 2149
Zeeman effect, 987
Anti-normal correlation function, 1782, 1789
Anti-resonant term, 1312
Anti-Stokes (Raman line), 532, 752
Antibunching (photon), 2121
Anticommutation, 1599
field operator, 1754
Anticrossing of levels, 415, 482
Antisymmetric ket, state, 1428, 1431
Antisymmetrizer, 1428, 1431
Applications of the perturbation theory, 1231
Approximation
central field approximation, 1459
secular approximation, 1374
Argument (EPR), 2205
Atom(s), *see* helium, hydrogenoid
donor, 837
dressed, 2129, 2133
many-electron atoms, 1459, 1467
mirrors for atoms, 2153
muonic atom, 541
single atom fluorescence, 2121
Atomic
beam (deceleration), 2025
orbital, 869, 1496(ex.)
parameters, 41
Attractive bosons, 1747
Autler-Townes
doublet, 2144
effect, 1410
Autoionization, 1468
Average value (anomalous), 1828
Azimuthal
quantum number, 811

Band (energy), 381
Bardeen-Cooper-Schrieffer, 1889
Barrier (potential barrier), 68, 367, 373
Basis

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

change of bases, 174
characteristic relations, 101, 119
continuous basis in the space of states, 99
mixed basis in the space of states, 99
BCHSH inequalities, 2209, 2210
BCS, 1889
broken pairs and excited pairs, 1920
coherent length, 1909
distribution functions, 1899
elementary excitations, 1923
excited states, 1919
gap, 1894, 1896, 1923
pairs (wave function of), 1901
phase locking, 1893, 1914, 1916
physical mechanism, 1914
two-particle distribution, 1901
Bell's
inequality, 2208
theorem, 2204, 2208
Benzene (molecule), 417, 495
Bessel
Bessel-Parseval relation, 1507
spherical Bessel function, 944
spherical equation, 961
spherical function, 966
Biorthonormal decomposition, 2194
Bitter, 2059
Blackbody radiation, 651
Bloch
equations, 463, 1358, 1361
theorem, 659
Bogolubov
excitations, 1661
Hamiltonian, 1952
operator method, 1950
phonons, spectrum, 1660
transformation, 1950
Bogolubov-Valatin transformation, 1836, 1919
Bohr, 2207
electronic magneton, 856
frequencies, 249
magneton, *see* front cover pages
model, 40, 819
nuclear magneton, 1237
radius, 820
Boltzmann
constant, *see* front cover pages
distribution, 1630
Born
approximation, 938, 977, 1320
Born-Oppenheimer approximation, 528, 1177, 1190
Born-von Karman conditions, 1490
Bose-Einstein
condensation, 1446, 1638, 1940
condensation (repulsive bosons), 1933
condensation of pairs, 1857
distribution, 652, 1630
statistics, 1446
Bosons, 1434
at non-zero temperature, 1745
attractive, 1747
attractive instability, 1745
condensed, 1638
in a Fock state, 1775
paired, 1881
Boundary conditions (periodic), 1489
Bra, 103, 104, 119
Bragg reflection, 382
Brillouin
formula, 452
zone, 614
Broadband
detector, 2165
optical excitation, 1332
Broadening (radiative), 2138
Broken pairs and excited pairs (BCS), 1920
Brossel, 2059
Bunching of bosons, 1777

C.S.C.O., 133, 137, 153, 236
Canonical
commutation relations, 142, 223, 1984
ensemble, 2289
Hamilton-Jacobi canonical equations, 214
Hamilton-Jacobi equations, 1532
Cauchy principal part, 1517
Center of mass, 812, 1528
Center of mass frame, 814
Central

field approximation, 1459
potential, 1533
Central potential, 803, 841
scattering, 941
stationary states, 804
Centrifugal potential, 809, 888, 893
Chain (von Neumann), 2201
Chain of coupled harmonic oscillators, 611
Change
of bases, 124, 174, 1601
of representation, 124
Characteristic equation, 129
Characteristic relation of an orthonormal basis, 116
Charged harmonic oscillator in an electric field, 575
Charged particle
in an electromagnetic field, 1536
Charged particle in a magnetic field, 240, 321, 771
Chemical bond, 417, 869, 1189, 1210
Chemical potential, 1486, 2287
Circular quanta, 761, 783
Classical
electrodynamics, 1957
histories, 2272
Clebsch-Gordan coefficients, 1038, 1051
Closure relation, 93, 117
Coefficients
Clebsch-Gordan, 1038
Einstein, 1334, 2083
Coherences (of the density matrix), 307
Coherent length (BCS), 1909
Coherent state (field), 2008
Coherent superposition of states, 253, 301, 307
Collision, 923
between identical particles, 1454, 1497(ex.)
between identical particles in classical mechanics, 1420
between two identical particles, 1450
cross section, 926
scattering states, 928
total scattering cross section, 926
with absorption, 971
Combination
of atomic orbitals, 1172
Commutation, 1599
canonical relations, 142, 223
field operator, 1754
of pair field operators, 1861
relations, 1984
Commutation relations
angular momentum, 669, 725
field, 1989, 1996
Commutator algebra, 165
Commutator(s), 91, 167, 171, 187
of functions of operators, 168
Compatibility of observables, 232
Complementarity, 45
Complete set of commuting observables (C.S.C.O.), 133, 137, 236
Complex variables (Lagrangian), 1982
Compton wavelength of the electron, 825, 1235
Condensates
relative phase, 2237
with spins, 2254
Condensation
BCS condensation energy, 1917
Bose-Einstein, 1446, 1857, 1933
Condensed bosons, 1638
Conduction band, 1492
Conductivity (solid), 1492
Configurations, 1467
Conjugate momentum, 214, 323, 1531, 1983, 1987, 1995
Conjugation (Hermitian), 111
Conservation
local conservation of probability, 238
of angular momentum, 668, 736, 1016
of energy, 248
of probability, 237
Conservative systems, 245, 315
Constants of the motion, 248, 317
Contact term, 1273
Contact term (Fermi), 1238, 1247
Contextuality, 2231
Continuous
spectrum, 133, 219, 264, 1316
variables (in a Lagrangian), 1984
Continuum of final states, 1316, 1378, 1380
Contractions, 1802

Convolution product of two functions, 1510
Cooling
Doppler, 2026
down atoms, 2025
evaporative, 2034
Sisyphus, 2034
sub-Doppler, 2155
subrecoil, 2034
Cooper model, 1927
Cooper pairs, 1927
Cooperative effects (BCS), 1916
Correlation functions, 1781, 1804
anti-normal, 1782, 1789
dipole and field, 2113
for one-photon processes, 2084
normal, 1782, 1787
of the field, spatial, 1758
Correlations, 2231
between two dipoles, 1157
between two physical systems, 296
classical and quantum, 2221
introduced by a collision, 1104
Coulomb
field, 1962
gauge, 1965
Coulomb potential
cross section, 979
Coupling
between angular momenta, 1016
between two angular momenta, 1091
between two states, 412
effect on the eigenvalues, 438
spin-orbit coupling, 1234, 1241
Creation and annihilation operators, 504, 513, 514, 1596, 1990
Creation operator (pair of particles), 1813, 1846
Critical velocity, 1671
Cross section
and phase shifts, 951
scattering cross section, 926, 933, 953, 972
Current
metastable current in superfluid, 1667
of particles, 1758
of probability, 240
probability current in hydrogen atom, 851
Cylindrical symmetry, 899(ex.)

Darwin term, 1235, 1279
De Broglie
relation, 10
wavelength, *see* front cover pages, 11, 35
Decay of a discrete state, 1378
Deceleration of an atomic beam, 2025
Decoherence, 2199
Decomposition (Schmidt), 2193
Decoupling (fine or hyperfine structure), 1262, 1291
Degeneracy
essential, 811, 825, 845
exchange degeneracy, 1423
exchange degeneracy removal, 1435
lifted by a perturbation, 1125
rotation invariance, 1072
systematic and accidental, 203
Degenerate eigenvalue, 127, 203, 217, 260
Degereracy
lifted by a perturbation, 1117
parity, 199
Delta Dirac function, 1515
potential well and barriers, 83–85(ex.)
use in quantum mechanics, 97, 106, 280
Density
Lagrangian, 1986
of probability, 264
of states, 389, 1316, 1484, 1488
operator, 449, 1391
operator and matrix, 299
particle density operator, 1756
Density functions
one and two-particle, 1502(ex.)
Depletion (quantum), 1940
Derivative of an operator, 169
Detection probability amplitude (photon), 2166
Detectors (photon), 2165
Determinant
Slater determinant, 1438, 1679
Deuterium, 834, 1107(ex.)
Diagonalization

of a 2×2 matrix, 429
of an operator, 128
Diagram (dressed-atom), 2133
Diamagnetism, 855
Diatomic molecules
rotation, 739
Diffusion (momentum), 2030
Dipole
-dipole interaction, 1142, 1153
-dipole magnetic interaction, 1237
electric dipole transition, 863
electric moment, 1080
Hamiltonian, 2011
magnetic dipole moment, 1084
magnetic term, 1272
trap, 2151
Dirac, *see* Fermi
delta function, 97, 106, 280, 1515
equation, 1233
notation, 102
Direct
and exchange terms, 1613, 1632, 1634, 1646, 1650
term, 1447, 1453
Discrete
bases of the state space, 91
spectrum, 132, 217
Dispersion (anomalous), 2149
Dispersion and absorption (field), 2147
Distribution
Boltzmann, 1630
Bose-Einstein, 1630
Fermi-Dirac, 1630
function (bosons), 1629
function (fermions), 1629
functions, 1625, 1733
functions (BCS), 1899
Distribution law
Bose-Einstein, 652
Divergence (energy), 2007
Donor atom, 837, 1495
Doppler
cooling, 2026
effect, 2022
effect (relativistic), 2022
free spectroscopy, 2105
temperature, 2033
Double
condensate, 2237
resonance method, 2059
spin condensate, 2254
Doublet (Autler-Townes), 2144
Down-conversion (parametric), 2181
Dressed
states and energies, 2133
Dressed-atom, 2129, 2133
diagram, 2133
strong coupling, 2141
weak coupling, 2137

E.P.R., 1225(ex.)
Eckart (Wigner-Eckart theorem), *see* Wigner
Effect
Autler-Townes, 2144
Mössbauer, 2040
photoelectric, 2110
Effective Hamiltonian, 2141
Ehrenfest theorem, 242, 319, 522
Eigenresult, 9
Eigenstate, 217, 232
Eigenvalue, 11, 25, 176, 216
degenerate, 217, 260
equation, 126, 429
of an operator, 126
Eigenvector, 176
of an operator, 126
Einstein, 2110
coefficients, 1334, 1356, 2083
EPR argument, 297, 1104
model, 534, 653
Planck-Einstein relations, 3
temperature, 659
Einstein-Podolsky-Rosen, 2204, 2261
Elastic
scattering, 925
scattering (photon), 2086
scattering, form factor, 1411(ex.)
total cross section, 972
Elastically bound electron model, 1350
Electric
conductivity of a solid, 1492
Electric dipole
Hamiltonian, 2011
interaction, 1342

matrix elements, 1344
moment, 1080
selection rules, 1345
transition and selection rules, 863
transitions, 2056
Electric field (quantized), 2000, 2005
Electric polarisability
NH_3, 484
Electric polarizability
of the 1s state in Hydrogen, 1299
Electric quadrupole
Hamiltonian, 1347
moment, 1082
transitions, 1348
Electric susceptibility
bound electron, 577
of an atom, 1351
Electrical
susceptibility, 1223(ex.)
Electrodynamics
classical, 1957
quantum, 1997
Electromagnetic field
and harmonic oscillators, 1968
and potentials, 321
angular momentum, 1968, 2043
energy, 1966
Lagrangian, 1986, 1992
momentum, 1967, 2019
polarization, 1970
quantization, 631, 637
Electromagnetic interaction of an atom with a wave, 1340
Electromagnetism
fields and potentials, 1536
Electron spin, 393, 985
Electron(s)
configurations, 1463
gas in solids, 1491
in solids, 1177, 1481
mass and charge, *see* front cover pages
Electronic
configuration, 1459
paramagnetic resonance, 1225(ex.)
shell, 827
Elements of reality, 2205
Emergence of a relative phase, 2248, 2253
Emission
of a quantum, 1311
photon, 2080
spontaneous, 2081, 2135
stimulated (or induced), 2081
Energy, *see* Conservation, Uncertainty
and momentum of the transverse electromagnetic field, 1973
band, 381
bands in solids, 1177, 1481
conservation, 248
electromagnetic field, 1966
Fermi energy, 1772
fine structure energy levels, 986
free energy, 2290
levels, 359
levels of harmonic oscillator, 509
levels of hydrogen, 823
of a paired state, 1869
recoil energy, 2023
Ensemble
canonical, 2289
grand canonical, 2291
microcanonical, 2285
statistical ensembles, 2295
Entanglement
quantum, 2187, 2193, 2203, 2242
swapping, 2232
Entropy, 2286
EPR, 2204, 2261
elements of reality, 2205
EPRB, 2205
paradox/argument, 1104
Equation of state
ideal quantum gas, 1640
repulsive bosons, 1745
Equation(s)
Bloch, 1361
Hamilton-Jacobi, 1982, 1983, 1988
Lagrange, 1982, 1993
Lorentz, 1959
Maxwell, 1959
Schrödinger, 11, 12, 306
von Neumann, 306
Essential degeneracy, 811, 825
Ethane (molecule), 1223
Ethylene (molecule), 536, 881

Evanescent wave, 29, 67, 70, 78, 285
Evaporative cooling, 2034
Even operators, 196
Evolution
  field operator, 1765
  of quantum systems, 223
  of the mean value, 241
  operator, 313, 2069
  operator (expansion), 2070
  operator (integral equation), 2069
Exchange, 1611
  degeneracy, 1423
  degeneracy removal, 1435
  energy, 1469
  hole, 1774
  integral, 1474
  term, 1447, 1451, 1453
Excitations
  BCS, 1923
  Bogolubov, 1661
  vacuum, 1623
Excited states (BCS), 1919
Exciton, 838
Exclusion principle (Pauli), 1437, 1444, 1463, 1484
Extensive (or intensive) variables, 2292

Fermi
  contact term, 1238
  energy, 1445, 1481, 1486, 1772
  gas, 1481
  golden rule, 1318
  level, 1486, 1621
  radius, 1621
  surface (modified), 1914
  , *see* Fermi-Dirac
Fermi level
  and electric conductivity, 1492
Fermi-Dirac
  distribution, 1486, 1630, 1717
  statistics, 1446
Fermions, 1434
  in a Fock state, 1771
  paired, 1874
Ferromagnetism, 1477
Feynman
  path, 2267
  postulates, 341
Fictitious spin, 435, 1359
Field
  absorption, 2149
  commutation relations, 1989, 1996
  dispersion and absorption, 2147
  intense laser, 2126
  interaction energy, 1764
  kinetic energy, 1763
  normal variables, 1971
  operator, 1752
  operator (evolution), 1763, 1765
  pair field operator, 1861
  potential energy, 1764
  quantization, 1765, 1999
  quasi-classical state, 2008
  spatial correlation functions, 1758
Final states continuum, 1378, 1380
Fine and hyperfine structure, 1231
Fine structure
  constant, *see* front cover pages, 825
  energy levels, 1478
  Hamiltonian, 1233, 1276, 1478
  Helium atom, 1478
  Hydrogen, 1238
  of spectral lines, 986
  of the states $1s$, $2s$ et $2p$, 1276
Fletcher, 2111
Fluctuations
  boson occupation number, 1633
  intensity, 2125
  vacuum, 644, 2007
Fluorescence (single atom), 2121
Fluorescence triplet, 2144
Fock
  space, 1593, 2004
  state, 1593, 1614, 1769, 2103
Forbidden, *see* Band
  energy band, 381, 390, 1481
  transition, 1345
Forces
  van der Waals, 1151
Form factor
  elastic scattering, 1411(ex.)
Forward scattering (direct and exchange), 1874
Fourier

series and transforms, 1505
Fragmentation (condensate), 1654, 1776
Free
electrons in a box, 1481
energy, 2290
particle, 14
quantum field (Fock space), 2004
spherical wave, 941, 944, 961
spherical waves and plane waves, 967
Free particle
stationary states with well-defined angular momentum, 959
stationary states with well-defined momentum, 19
wave packet, 14, 57, 347
Frequency
Bohr, 249
components of the field (positive and negative), 2072
Rabi's frequency, 1325
Friction (coefficient), 2028
Function
of operators, 166
periodic functions, 1505
step functions, 1521
Fundamental state, 41

Gap (BCS), 1894, 1896, 1923
Gauge, 1343, 1536, 1960, 1963
Coulomb, 1965
invariance, 321
Lorenz, 1965
Gaussian
wave packet, 57, 292, 2305
Generalized velocities, 214, 1530
Geometric quantization, 2311
Gerlach, *see* Stern
GHZ state, 2222, 2227
Gibbs-Duhem relation, 2296
Golden rule (Fermi), 1318
Good quantum numbers, 248
Grand canonical, 1626, 2291
Grand potential, 1627, 1721, 2292
Green's function, 337, 936, 1781, 1786, 1789
evolution, 1785
Greenberger-Horne-Zeilinger, 2227
Groenewold's formula, 2315
Gross-Pitaevskii equation, 1643, 1657
Ground state, 363
harmonic oscillator, 509, 520
Hydrogen atom, 1228(ex.)
Group velocity, 55, 60, 614
Gyromagnetic ratio, 396, 455
orbital, 860
spin, 988

H_2^+ molecular ion, 85(ex.), 417, 1189
Hadronic atoms, 840
Hall effect, 1493
Hamilton
function, 1532
function and equations, 1531
Hamilton-Jacobi canonical equations, 214, 1532, 1982, 1983, 1988
Hamiltonian, 223, 245, 1527, 1983, 1988, 1995
classical, 1531
effective, 2141
electric dipole, 1342, 2011
electric quadrupole, 1347
fine structure, 1233, 1276
hyperfine, 1237, 1267
magnetic dipolar, 1347
of a charged particle in a vector potential, 1539
of a particle in a central potential, 806, 1533
of a particle in a scalar potential, 225
of a particle in a vector potential, 225, 323, 328
Hanbury Brown and Twiss, 2120
Hanle effect, 1372(ex.)
Hard sphere
scattering, 980, 981(ex.)
Harmonic oscillator, 497
in an electric field, 575
in one dimension, 527, 1131
in three dimensions, 569
in two dimensions, 755
infinite chain of coupled oscillators, 611
quasiclassical states, 583
thermodynamic equilibrium, 647

three-dimensional, 841, 899(ex.)
two coupled oscillators, 599
Hartree-Fock
approximation, 1677, 1701
density operator (one-particle), 1691
equations, 1686, 1731
for electrons, 1695
mean field, 1677, 1693
potential, 1706
thermal equilibrium, 1711, 1733
time-dependent, 1701, 1708
Healing length, 1652
Heaviside step function, 1521
Heisenberg
picture, 317, 1763
relations, 19, 39, 41, 45, 55, 232, 290
Helicity (photon), 2051
Helium
energy levels, 1467
ion, 838
isotopes, 1480
isotopes ^{3}He and ^{4}He, 1435, 1446
solidification, 535
Hermite polynomials, 516, 547, 561
Hermitian
conjugation, 111
matrix, 124
operator, 115, 124, 130
Histories (classical), 2272
Hole
creation and annihilation, 1622
exchange, 1774
Holes, 1621
Hybridization of atomic orbitals, 869
Hydrogen, 645
atom, 803
atom in a magnetic field, 853, 855, 862
atom, relativistic energies, 1245
Bohr model, 40, 819
energy levels, 823
fine and hyperfine stucture, 1231
ionisation energy, *see* front cover pages
ionization energy, 820
maser, 1251
molecular ion, 85(ex.), 417, 1189
quantum theory, 41
radial equation, 821
Stark effect, 1298
stationary states, 851
stationary wave functions, 830
Hydrogen-like systems in solid state physics, 837
Hydrogenoid systems, 833
Hyperfine
decoupling, 1262
Hamiltonian, 1237, 1267
Hyperfine structure, *see* Hydrogen, muonium, positronium, Zeeman effect, 1231
Muonium, 1281

Ideal gas, 1625, 1787, 1791, 1804
correlations, 1769
Identical particles, 1419, 1591
Induced
emission, 1334, 1366, 2081
emission of a quantum, 1311
emission of photons, 1355
Inequality (Bell's), 2208
Infinite one-dimensional well, 271
Infinite potential well, 74
in two dimensions, 201
Infinitesimal unitary operator, 178
Insulator, 1492
Integral
exchange integral, 1474
scattering equation, 935
Intense laser fields, 2126
Intensive (or extensive) variables, 2292
Interaction
between magnetic dipoles, 1141
dipole-dipole interaction, 1141, 1153
electromagnetic interaction of an atom with a wave, 1340
field and particles, 2009
field and atom, 2010
magnetic dipole-dipole interaction, 1237
picture, 353, 1393, 2070
tensor interaction, 1141
Interference
photons, 2167
two-photon, 2170, 2183
Ion H_2^+, 1189

Ionization
photo-ionization, 2109
tunnel ionization, 2126
Isotropic radiation, 2079

Jacobi, *see* Hamilton

Kastler, 2059, 2062
Ket, *see* state, 103, 119
for identical particles, 1436
Kuhn, *see* Thomas

Lagrange
equations, 1530, 1982, 1993
fonction and equations, 214
multipliers, 2281
Lagrangian, 1530, 1980
densities, 1986
electromagnetic field, 1986, 1992
formulation of quantum mechanics, 339
of a charged particle in an electromagnetic field, 1538
particle in an electromagnetic field, 323
Laguerre-Gaussian beams, 2065
Lamb shift, 645, 1245, 1388, 2008
Landau levels, 771
Landé factor, 1072, 1107(ex.), 1256, 1292
Laplacian, 1527
of $1/r$, 1524
of $Y_l^m(\theta,\varphi)/r^{l+1}$, 1526
Larmor
angular frequency, 857
precession, 394, 396, 410, 455, 857, 1071
Laser, 1359, 1365
Raman laser, 2093
saturation, 1370
trap, 2151
Lattices (optical), 2153
Least action
principle of, 1539
Legendre
associated function, 714
polynomial, 713
Length (healing), 1652
Level
anticrossing, 415, 482
Fermi level, 1621
Lifetime, 343, 485, 645
of a discrete state, 1386
radiative, 2081
Lifting of degeneracy by a perturbation, 1125
Light
quanta, 3
shifts, 1334, 2138, 2151, 2156
Linear, *see* operator
combination of atomic orbitals, 1172
operators, 90, 108, 163
response, 1350, 1357, 1364
superposition of states, 253
susceptibility, 1365
Local conservation of probability, 238
Local realism, 2209, 2230
Longitudinal
fields, 1961
relaxation, 1400
relaxation time, 1401
Lorentz equations, 1959
Lorenz (gauge), 1965

Magnetic
dipole term, 1272
dipole-dipole interaction, 1237
effect of a magnetic field on the levels of the Hydrogen atom, 1251
hyperfine Hamiltonian, 1267
interactions, 1232, 1237
quantum number, 811
resonance, 455
susceptibility, 1224, 1487
Magnetic dipole
Hamiltonian, 1347
transitions and selection rules, 1084, 1098, 1348
Magnetic dipoles
interactions between two dipoles, 1141
Magnetic field
and vector potential, 321
charged particle in a, 240, 771
effects on hydrogen atom, 853, 855
harmonic oscillator in a, 899(ex.)
Hydrogen atom in a magnetic field, 1263, 1289

multiplets, 1074
quantized, 2000, 2005
Magnetism (spontaneous), 1737
Many-electron atoms, 1459
Maser, 477, 1359, 1365
hydrogen, 1251
Mass correction (relativistic), 1234
Master equation, 1358
Matrice(s), 119, 121
diagonalization of a 2×2 matrix, 429
Pauli matrices, 425
unitary matrix, 176
Maxwell's equations, 1959
Mean field (Hartree-Fock), 1693, 1708, 1725
Mean value of an observable, 228
evolution, 241
Measurement
general postulates, 216, 226
ideal von Neumann measurement, 2196
of a spin 1/2, 394
of observables, 216
on a part of a physical system, 293
state after measurement, 221, 227
Mendeleev's table, 1463
Metastable superfluid flow, 1671
Methane (molecule), 883
Microcanonical ensemble, 2285
Millikan, 2111
Minimal wave packet, 290, 520, 591
Mirrors for atoms, 2153
Mixing of states, 1121, 1137
Model
Cooper model, 1927
Einstein model, 534
elastically bound electron, 1350
vector model of atom, 1071
Modes
vibrational modes, 599, 611
Modes (radiation), 1974, 1975
Molecular ion, 417
Molecule(s)
chemical bond, 417, 869, 873, 878, 883, 1189
rotation, 796
vibration, 527, 1137
vibration-rotation, 885
Mollow, 2144
Moment
quadrupole electric moment, 1225(ex.)
Momentum, 1539
conjugate, 214, 323, 1983, 1987, 1995
diffusion, 2030
electromagnetic field, 1967, 2019
mechanical momentum, 328
Monogamy (quantum), 2221
Mössbauer effect, 1415, 2040
Motional narrowing, 1323
condition, 1323, 1398, 1408
Multiphoton transition, 1368, 2040, 2097
Multiplets, 1072, 1074, 1467
Multipliers (Lagrange), 2281
Multipolar waves, 2052
Multipole
moments, 1077
Multipole operators
introduction, 1077, 1083
parity, 1082
Muon, 527, 541, 1281
Muonic atom, 541, 839
Muonium, 835
hyperfine structure, 1281
Zeeman effect, 1281

Narrowing (motional), 1323, 1408
condition, 1398
Natural width, 345, 1388
Need for a quantum treatment, 2118, 2120
Neumann
spherical function, 967
Neutron mass, *see* front cover pages
Non-destructive detection of a photon, 2159
Non-diagonal order (BCS), 1912
Non-locality, 2204
Non-resonant excitation, 1350
Non-separability, 2207
Nonlinear
response, 1357, 1368
susceptibility, 1369
Norm
conservation, 238
of a state vector, 104, 237
of a wave function, 13, 90, 99

Normal
correlation function, 1782, 1787
variables, 602, 616, 631, 633
variables (field), 1971
Nuclear
multipole moments, 1088
Bohr magneton, 1237
Nucleus
spin, 1088
volume effect, 1162, 1268
Number
occupation number, 1439, 1593
photon number, 2135
total number of particles in an ideal gas, 1635

Observable(s), 130
C.S.C.O., 133, 137
commutation, 232
compatibility, 232
for identical particles, 1429, 1441
mean value, 228
measurement of, 216, 226
quantization rules, 223
symmetric observables, 1441
transformation by permutation, 1434
whose commutator is $i\hbar$, 187, 289
Occupation number, 1439, 1593
operator, 1598
Odd operators, 196
One-particle
Hartree-Fock density operator, 1691
operators, 1603, 1605, 1628, 1756
Operator(s)
adjoint operator, 112
annihilation operator, 504, 513, 514, 1597
creation and annihilation, 1990
creation operator, 504, 513, 514, 1596
derivative of an operator, 169
diagonalization, 126, 128
even and odd operators, 196
evolution operator, 313, 2069
field, 1752
function of, 166
Hermitian operators, 115
linear operators, 90, 108, 163
occupation number, 1598
one-particle operator, 1603, 1605, 1628, 1756
parity operator, 193
particle density operator, 1756
permutation operators, 1425, 1430
potential, 168
product of, 90
reduced to a single particle, 1607
representation, 121
restriction, 165
restriction of, 1125
rotation operator, 1001
symmetric, 1628, 1755
translation operator, 190
two-particle operator, 1608, 1610, 1631, 1756
unitary operators, 173
Weyl operator, 2300
Oppenheimer, *see* Born, 1177, 1190
Optical
excitation (broadband), 1332
lattices, 2153
pumping, 2062, 2140
Orbital
angular momentum (of radiation), 2052
atomic orbital, 1496(ex.)
hybridization, 869
linear combination of atomic orbitals, 1172
quantum number, 1463
state space, 988
Order parameter for pairs, 1851
Orthonormal basis, 91, 99, 101, 133
characteristic relation, 116
Orthonormalization
and closure relations, 101, 140
relation, 116
Oscillation(s)
between two discrete states, 1374
between two quantum states, 418
Rabi, 2134
Oscillator
anharmonic, 502
harmonic, 497
strength, 1352

Pair(s)

annihilation-creation of pairs, 1831, 1874, 1887
BCS, wave function, 1909
Cooper, 1927
of particles (creation operator), 1813, 1846
pair field (commutation), 1861
pair field operator, 1845
pair wave function, 1851
Paired
bosons, 1881
fermions, 1874
state energy, 1869
states, 1811
states (building), 1818
Pairing term, 1878
Paramagnetism, 855
Parametric down-conversion, 2181
Parity, 2106
degeneracy, 199
of a permutation operator, 1431
of multipole operators, 1082
operator, 193
Parseval
Parseval-Plancherel equality, 20
Parseval-Plancherel formula, 1511, 1521
Partial
reflection, 79
trace of an operator, 309
waves in the potential, 948
waves method, 941
Particle (current), 1758
Particles and holes, 1621
Partition function, 1626, 1627, 1717
Path
integral, 2267
space-time path, 339
Pauli
exclusion principle, 1437, 1444, 1463, 1481
Hamiltonian, 1009(ex.)
matrices, 425, 991
spin theory, 986
spinor, 993
Penetrating orbit, 1463
Penrose-Onsager criterion, 1776, 1860, 1947
Peres, 2212
Periodic
boundary conditions, 1489
classification of elements, 1463
functions, 1505
potential (one-dimensional), 375
Permutation operators, 1425, 1430
Perturbation
applications of the perturbation theory, 1231
lifting of a degeneracy, 1125
one-dimensional harmonic oscillator, 1131
random perturbation, 1320, 1325, 1390
sinusoidal, 1311
stationary perturbation theory, 1115
Perturbation theory
time dependent, 1303
Phase
locking (BCS), 1893, 1916
locking (bosons), 1938, 1944
relative phase between condensates, 2237, 2248
velocity, 37
Phase shift (collision), 951, 1497(ex.)
with imaginary part, 971
Phase velocity, 21
Phonons, 611, 626
Bogolubov phonons, 1660
Photodetection
double, 2172, 2184
single, 2169, 2171
Photoelectric effect, 1412(ex.), 2110
Photoionization, 2109, 2165
rate, 2115, 2124
two-photon, 2123
Photon, 3, 631, 651, 2004, 2005, 2110
absorption and emission, 2067
angular momentum, 1370
antibunching, 2121
detectors, 2165
non-destructive detection, 2159
number, 2135
scattering (elastic), 2086
scattering by an atom, 2085
vacuum, 2007
, *see* Absorption, Emission
Picture

Heisenberg, 317, 1763
interaction, 1393, 2070
Pitaevskii (Gross-Pitaevskii equation), 1643, 1657
Plancherel, *see* Parseval
Planck
constant, *see* front cover pages, 3
law , 2083
Planck-Einstein relations, 3, 10
Plane wave, 14, 19, 95, 943
Podolsky (EPR argument), 297, 1104
Pointer states, 2199
Polarizability
of the 1s state in Hydrogen, 1299
Polarization
electromagnetic field, 1970
of Zeeman components, 1295
space-dependent, 2156
Polynomial method (harmonic oscillator), 555, 842
Polynomials
Hermite polynomials, 516, 547, 561
Position and momentum representations, 181
Positive and negative frequency components, 2072
Positron, 1281
Positronium, 836
hyperfine structure, 1281
Zeeman effect, 1281
Postulate (von Neumann projection), 2202
Postulates of quantum mechanics, 215
Potential
adiabatic branching, 932
barrier, 26, 68, 367, 373
centrifugal potential, 809, 888, 893
Coulomb potential, cross section, 979
cylindrically symmetric, 899(ex.)
Hartree-Fock, 1706
infinite one-dimensional well, 74
operator, 168
scalar and vector potentials, 1536, 1960, 1963
scattering by a, 923
self-consistent potential, 1461
square potential, 63
square well, 29
step, 28, 65, 75, 284
well, 71, 367
well (arbitrary shape), 359
well (infinite one-dimensional), 271
well (infinite two-dimensional, 201
Yukawa potential, 977
Precession
Larmor precession, 396, 1071
Thomas precession, 1235
Preparation of a state, 235
Pressure (ideal quantum gas), 1640
Principal part, 1517
Principal quantum number, 827
Principle
of least action, 1539, 1980
of spectral decomposition, 11, 216
of superposition, 237
Probability
amplitude, 11, 253, 259
conservation, 237
current, 240, 283, 333, 349, 932
current in hydrogen atom, 851
density, 11, 264
fluid, 932
of photon absorption, 2076
of the measurement results, 9, 11
transition probability, 439
Process (pair annihilation-creation), 1878, 1887
Product
convolution product of functions, 1510
of matrices, 122
of operators, 90
scalar product, 101, 141, 149, 161
state (tensor product), 311
tensor product, 147
tensor product, applications, 441
Projection theorem, 1070
Projector, 109, 133, 165, 218, 222, 1108(ex.)
Propagator
for the Schrödinger equation, 335
of a particle, 2267, 2272
Proper result, 9
Proton
mass, *see* front cover pages
spin and magnetic moment, 1237, 1274
Pumping, 1358

Pure (state or case), 301

Quadrupolar electric moment, 1082, 1225(ex.)
Quanta (circular), 761, 783
Quantization
  electrodynamics, 1997
  electromagnetic field, 631, 637, 1997
  of a field, 1765
  of angular momentum, 394, 677
  of energy, 3, 11, 71, 359
  of measurement results, 9, 216, 398
  of the measurement results, 405
  rules, 11, 223, 226, 2274
Quantum
  angle, 2258
  electrodynamics, 1245, 1282, 1997
  entanglement, 2187, 2193
  monogamy, 2221
  number
    orbital, 1463
    principal quantum number, 827
  numbers (good), 248
  resonance, 417
  treatment needed, 2118, 2120
Quasi-classical
  field states, 2008
  states, 765, 791, 801
  states of the harmonic oscillator, 583
Quasi-particles, 1736, 1840
  Bogolubov phonons, 1954
  Quasi-particle vacuum, 1836

Rabi
  formula, 440, 460, 1324, 1376
  formula), 419
  frequency, 1325
  oscillation, 2134
Radial
  equation, 842
  equation (Hydrogen), 821
  equation in a central potential, 808
  integral, 1277
  quantum number, 811
Radiation
  isotropic, 2079
  pressure, 2024
Radiative
  broadening, 2138
  cascade of the dressed atom, 2145
Raman
  effect, 532, 740, 1373(ex.)
  laser, 2093
  scattering, 2091
  scattering (stimulated), 2093
Random perturbation, 1320, 1325, 1390
Rank (Schmidt), 2196
Rate (photoionization), 2115, 2124
Rayleigh
  line, 752
  scattering, 532, 2089
Realism (local), 2205, 2209
Recoil
  blocking, 2036
  effect of the nucleus, 834
  energy, 1415, 2023
  free atom, 2020
  suppression, 2040
Reduced
  density operator, 1607
  mass, 813
Reduction of the wave packet, 221, 279
Reflection on a potential step, 285
Refractive index, 2149
Reiche, *see* Thomas
Relation (Gibbs-Duhem), 2296
Relative
  motion, 814
  particle, 814
  phase between condensates, 2248, 2258
  phase between spin condensates, 2253
Relativistic
  corrections, 1233, 1478
  Doppler effect, 2022
  mass correction, 1234
Relaxation, 465, 1358, 1390, 1413, 1414(ex.)
  general equations, 1397
  longitudinal, 1400
  longitudinal relaxation time, 1401
  transverse, 1403
  transverse relaxation time, 1406
Relay state, 2086, 2098, 2106
Renormalization, 2007
Representation(s)
  change of, 124
  in the state space, 116

of operators, 121
position and momentum, 139, 181
Schrödinger equation, 183–185
Repulsion between electrons, 1469
Resonance
magnetic resonance, 455
quantum resonance, 417, 1158
scattering resonance, 69, 954, 983(ex.)
two resonnaces with a sinusoidal excitation, 1365
width, 1312
with sinusoidal perturbation, 1311
Restriction of an operator, 165, 1125
Rigid rotator, 740, 1222(ex.)
Ritz theorem, 1170
Root mean square deviation
general definition, 230
Rosen (EPR argument), 297, 1104
Rotating frame, 459
Rotation(s)
and angular momentum, 717
invariance and degeneracy, 734
of diatomic molecules, 739
of molecules, 796, 885
operator(s), 720, 1001
rotation invariance, 1478
rotation invariance and degeneracy, 1072
Rotator
rigid rotator, 740, 1222(ex.)
Rules
quantization rules, 2274
selection rules, 197
Rutherford's formula, 979
Rydberg constant, *see* front cover pages

Saturation
of linear response, 1368
of the susceptibility, 1369
Scalar
and vector potentials, 321, 1536
interaction between two angular momenta, 1091
observable, operator, 732, 737
potential, 225
product, 89, 92, 101, 141, 149, 161
product of two coherent states, 593
Scattering
amplitude, 929, 953
by a central potential, 941
by a hard sphere, 980, 981(ex.)
by a potential, 923
cross section, 933, 953, 972
cross section and phase shifts, 951
inelastic, 2091
integral equation, 935
of particles with spin, 1102
of spin 1/2 particles, 1108(ex.)
photon, 2086
Raman, 2091
Rayleigh, 532, 2089
resonance, 954, 983(ex.)
resonant, 2089
stationary scattering states, 951
stationary states, 928
stimulated Raman, 2093
Schmidt
decomposition, 2193
rank, 2196
Schottky anomaly, 654
Schrödinger, 2190
equation, 11, 12, 223, 306
equation in momentum representation, 184
equation in position representation, 183
equation, physical implications, 237
equation, resolution for conservative systems, 245
picture, 317
Schwarz inequality, 161
Second
quantization, 1766
harmonic generation, 1368
Secular approximation, 1316, 1374
Selection rules, 197, 863, 2014, 2056
electric quadrupolar, 1348
magnetic dipolar, 1098, 1348
Self-consistent potential, 1461
Semiconductor, 837, 1493
Separability, 2207, 2223
Separable density operator, 2223
Shell (electronic), 827
Shift

light shift, 2138
of a discrete state, 1387
Singlet, 1024, 1474
Sinusoidal perturbation, 1311, 1374
Sisyphus
cooling, 2034
effect, 2155
Slater determinant, 1438, 1679
Slowing down atoms, 2025
Solids
electronic bands, 1177
energy bands of electrons, 1491
energy bands of electrons in solids, 381
hydrogen-like systems in solid state physics, 837
Space (Fock), 1593
Space-dependent polarization, 2156
Space-time path, 339, 1539
Spatial correlations (ideal gas), 1769
Specific heat
of an electron gas, 1484
of metals, 1487
of solids, 653
two level system, 654
Spectral
decomposition principle, 7, 11, 216
function, 1795
terms, 1469
Spectroscopy (Doppler free), 2105
Spectrum
BCS elementary excitation, 1923
continuous, 219, 264
discrete, 132, 217
of an observable, 126, 216
Spherical
Bessel equation, 961
Bessel function, 944, 966
free spherical waves, 961
free wave, 944
Neumann function, 967
wave, 941
waves and plane waves, 967
Spherical harmonics, 689, 705
addition of, 1059
expression for $l = 0, 1, 2$, 709
general expression, 707
Spin
and magnetic moment of the proton, 1237
angular momentum, 987
electron, 985, 1289
fictitious, 435
gyromagnetic ratio, 396, 455, 988
nuclear, 1088
of the electron, 393
Pauli theory, 986, 988
quantum description, 985, 991
rotation operator, 1001
scattering of particles with spin, 1102
spin 1 and radiation, 2044, 2049, 2050
system of two spins, 441
Spin 1/2
density operator, 449
ensemble of, 1358
fictitious, 1359
interaction between two spins, 1141
preparation and measurement, 401
scattering of spin 1/2 particles, 1108(ex.)
Spin-orbit coupling, 1018, 1234, 1241, 1279
Spin-statistics theorem, 1434
Spinor, 993
rotation, 1005
Spontaneous
emission, 343, 645, 1301, 2081, 2135
emission of photons, 1356
magnetism of fermions, 1737
Spreading of a wave packet, 59, 348
Square
barrier of potential, 26, 68
potential, 26, 63, 75, 283
potential well, 71, 271
spherical well, 982(ex.)
Standard representation (angular momentum), 677, 691
Stark effect in Hydrogen atom, 1298
State(s), *see* Density operator
density of, 389, 1316, 1484, 1488
Fock, 1593, 1614, 1769, 2103
ground state, 363
mixing of states by a perturbation, 1121
orbital state space, 988
paired, 1811

pointer states, 2199
quasi-classical states, 583, 765, 791, 801
relay state, 2086, 2098, 2106
stable and unstable states, 485
state after measurement, 221
state preparation, 235
stationary, 63, 359, 375
stationary state, 24, 246
stationary states in a central potential, 804
unstable, 343
vacuum state, 1595
vector, 102, 215
Stationary
perturbation theory, 1115
phase condition, 18, 54
scattering states, 928, 951
states, 24, 63, 246, 359
states in a periodic potential, 375
states with well-defined angular momentum, 944, 959
states with well-defined momentum, 943
Statistical
entropy, 2217
mechanics (review of), 2285
mixture of states, 253, 299, 304, 450
Statistics
Bose-Einstein, 1446
Fermi-Dirac, 1446
Step
function, 1521
potential, 28, 65, 75, 284
Stern-Gerlach experiment, 394
Stimulated
(or induced) emission, 1334, 1366, 2081
Raman scattering, 2093
Stokes Raman line, 532, 752
Stoner (spontaneous magnetism), 1737
Strong coupling (dressed-atom), 2141
Subrecoil cooling, 2034
Sum rule (Thomas-Reiche-Kuhn), 1352
Superfluidity, 1667, 1674
Superposition
of states, 253
principle, 7, 237
principle and physical predictions, 253
Surface (modified Fermi surface), 1914
Susceptibility, *see* Linear, nonlinear, tensor
electric susceptibility of an atom, 1351
electrical susceptibility, 577, e1223
electrical susceptibility of NH_3, 484
magnetic susceptibility, 1224
tensor, 1224, 1410(ex.)
Swapping (entanglement), 2232
Symmetric
ket, state, 1428, 1431
observables, 1429, 1441
operators, 1603, 1605, 1608, 1610, 1628, 1631, 1755
Symmetrization
of observables, 224
postulate, 1434
Symmetrizer, 1428, 1431
System
time evolution of a quantum system, 223
two-level system, 435
Systematic
and accidental degeneracies, 203
degeneracy, 845

Temperature (Doppler), 2033
Tensor
interaction, 1141
product, 147, 441
product of operators, 149
product state, 295, 311
product, applications, 201
susceptibility tensor, 1224
Term
direct and exchange terms, 1613, 1632, 1634, 1646, 1650
pairing, 1878
spectral terms, 1467, 1469
Theorem
Bell, 2204, 2208
Bloch, 659
projection, 1070
Ritz, 1170
Wick, 1799, 1804

Wigner-Eckart, 1065, 1085, 1254
Thermal wavelength, 1635
Thermodynamic equilibrium, 308
harmonic oscillator, 647
ideal quantum gas, 1625
spin 1/2, 452
Thermodynamic potential (minimization), 1715
Thomas precession, 1235
Thomas-Reiche-Kuhn sum rule, 1352
Three-dimensional harmonic oscillator, 569, 841, 899(ex.)
Three-level system, 1409(ex.)
Three-photon transition, 1370
Time evolution of quantum systems, 223
Time-correlations (fluorescent photons), 2145
Time-dependent
Gross-Pitaevskii equation, 1657
perturbation theory, 1303
Time-energy uncertainty relation, 250, 279, 345, 1312, 1389
Torsional oscillations, 536
Torus (flow in a), 1667
Total
elastic scattering cross section, 972
reflection, 67, 75
scattering cross section (collision), 926
Townes
Autler-Townes effect, 1410
Trace
of an operator, 163
partial trace of an operator, 309
Transform (Wigner), 2297
Transformation
Bogolubov, 1950
Bogolubov-Valatin, 1836, 1919
Gauge, 1960
of observables by permutation, 1434
Transition, *see* Probability, Forbidden, Electric dipole, Magnetic dipole, Quadrupole electric dipole, 2056
magnetic dipole transition, 1098
probability, 439, 1308, 1321, 1355
probability per unit time, 1319
probability, spin 1/2, 460
three-photon transition, 1370
two-photon, 2097
virtual, 2100
Translation operator, 190, 579, 791
Transpositions, 1431
Transverse
fields, 1961
relaxation, 1403
relaxation time, 1406
Trap
dipolar, 2151
laser, 2151
Triplet, 1024, 1474
fluorescence triplet, 2144
Tunnel
effect, 29, 70, 365, 476, 540, 1177
ionization, 2126
Two coupled harmonic oscillators, 599
Two-dimensional
harmonic oscillator, 755
infinite potential well, 201
wave packets, 49
Two-level system, 393, 411, 435, 1357
Two-particle operators, 1608, 1610, 1631, 1756
Two-photon
absorption, 1373(ex.)
interference, 2170, 2183
transition, 1409(ex.), 2097

Uncertainty
relation, 19, 39, 41, 45, 232, 290
time-energy uncertainty relation, 1312
Uniqueness of the measurement result, 2201
Unitary
matrix, 125, 176
operator, 173, 314
transformation of operators, 177
Unstable states, 343

Vacuum
electromagnetism, 644, 2007
excitations, 1623
fluctuations, 2007
photon vacuum, 2007
quasi-particule vacuum, 1836
state, 1595
Valence band, 1493

Van der Waals forces, 1151
Variables
intensive or extensive, 2292
normal variables, 602, 616, 631, 633
Variational method, 1169, 1190, 1228(ex.)
Vector
model, 1091
model of the atom, 1071, 1256
observable, operator, 732
operator, 1065
potential, 225
potential of a magnetic dipole, 1268
Velocity
critical, 1671
generalized velocities, 214, 1530
group velocity, 23, 614
phase velocity, 21, 37
Vibration(s)
modes, 599, 611
modes of a continuous system, 631
of molecules, 885, 1137
of nuclei in a crystal, 534, 611, 653
of the nuclei in a molecule, 527
Violations of Bell's inequalities, 2210, 2265
Virial theorem, 350, 1210
Virtual transition, 2100
Volume effect, 544, 840, 1162, 1268
Von Neumann
chain, 2201
equation, 306
ideal measurement, 2196
reduction postulate, 2202
statistical entropy, 2217
Vortex in a superfluid, 1667

Water (molecule), 873, 874
Wave (evanescent), 67
Wave function, 88, 140, 226
BCS pairs, 1901, 1909
Hydrogen, 830
norm, 90
pair wave functions, 1851
particle, 11
Wave packet(s)
Gaussian, 57, 2305
in a potential step, 75
in three dimensions, 53
minimal, 290, 520, 591
motion in a harmonic potential, 596
one-photon, 2168
particle, 13
photon, 2163
propagation, 20, 57, 242, 398
reduction, 221, 227, 265, 279
spreading, 57, 59, 347, 348(ex.)
two-dimension, 49
two-photons, 2181
Wave(s)
de Broglie wavelength, 10, 35
evanescent, 29
free spherical waves, 961
multipolar, 2052
partial waves, 948
plane, 14, 19, 943
wave function, 11, 88, 140, 226
Wave-particle duality, 3, 45
Wavelength
Compton wavelength, 1235
de Broglie, 10
Weak coupling (dressed-atom), 2137
Well
potential square well, 29
potential well, 367
Weyl
operator, 2300
quantization, 2311
Which path type of experiments, 2202
Wick's theorem, 1799, 1804
Wigner transform, 2297
Wigner-Eckart theorem, 1065, 1085, 1254

Young (double slit experiment), 4
Yukawa potential, 977

Zeeman
components, polarizations, 865
effect, 855, 862, 987, 1251, 1253, 1257, 1261, 1281
polarization of the components, 1295
slower, 2025
Zeeman effect
Hydrogen, 1289
in muonium, 1281
in positronium, 1281
Muonium, 1284

Zone (Brillouin zone), 614

WILEY-VCH

This new, third volume of Cohen-Tannoudji's groundbreaking textbook covers advanced topics of quantum mechanics such as uncorrelated and correlated identical particles, the quantum theory of the electromagnetic field, absorption, emission and scattering of photons by atoms, and quantum entanglement. Written in a didactically unrivalled manner, the textbook explains the fundamental concepts in seven *chapters,* each followed by *complements* that provide more detailed discussions, examples and applications.

- Completing the success story: the third and final volume of the quantum mechanics textbook written by 1997 Nobel laureate Claude Cohen-Tannoudji and his colleagues Bernard Diu and Franck Laloë
- As easily comprehensible as possible: all steps of the physical background and its mathematical representation are spelled out explicitly
- Comprehensive: in addition to the fundamentals themselves, the books comes with a wealth of elaborately explained examples and applications

Claude Cohen-Tannoudji *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris where he also studied and received his PhD in 1962. In 1973 he became Professor of atomic and molecular physics at the Collège des France. His main research interests were optical pumping, quantum optics and atom-photon interactions. In 1997, Claude Cohen-Tannoudji, together with Steven Chu and William D. Phillips, was awarded the Nobel Prize in Physics for his research on laser cooling and trapping of neutral atoms.*

Bernard Diu *was Professor at the Denis Diderot University (Paris VII). He was engaged in research at the Laboratory of Theoretical Physics and High Energy where his focus was on strong interactions physics and statistical mechanics.*

Franck Laloë *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris. His first assignment was with the University of Paris VI before he was appointed to the CNRS, the French National Research Center. His research was focused on optical pumping, statistical mechanics of quantum gases, musical acoustics and the foundations of quantum mechanics.*

Cover Design: Wiley
Cover Image: © antishock/Getty Images

ISBN 978-3-527-34555-7

www.wiley-vch.de

Also available as an e-book